

AI-powered tools in English pronunciation classroom: An action research report

Hana Vančová

Department of English Language and Literature, Trnava University in Trnava, , email: hana.vancova@truni.sk

How to cite: Vančová, H. (2024). AI-powered tools in English pronunciation classroom: An action research report. In Y. Choubsaz, P. Díez-Arcón, A. Gimeno-Sanz, J. Hriňák, X. Liashuk, S. Pokrivčáková & H. Vančová (Eds.), *CALL for Humanity - EUROCALL 2024 Short Papers*. <https://doi.org/10.4995/EuroCALL2024.2024.19082>

Abstract

The constant development of computer-assisted language learning (CALL) has been recently enriched by implementing artificial intelligence (AI) technology. The ability of the AI to process the learner's input and interactively move it forward towards accomplishing a communicative goal distinguishes it from automatic speech recognition (ASR) used thus far. In this way, AI-powered tools help develop learners' foreign language proficiency across all language layers, including pronunciation. The presented study aims to report recent experiences implementing AI-powered tools in a pronunciation classroom EFL context gained through action research. The study participants were first-year English students who were asked to explore the potential of AI-powered tools in improving their pronunciation through tasks imitating real-life communication and accuracy-aimed tasks. At the end, the participants were invited to complete a questionnaire focusing on their perceptions and experience with the tools. Their answers provide insight into their attitudes and allow for formulating recommendations for their implementation in English pronunciation classrooms. In particular, the less proficient participants appreciated conversation with a chatbot because it allowed them to practice communication in different scenarios. On the other hand, more advanced participants preferred tools providing explicit and guided feedback.

Keywords: AI-powered tools; pronunciation; English; pronunciation training; CALL; EFL.

1. Introduction

Pronunciation training has historically relied on technology, predominantly using various tools and programs that provide a model for EFL learners, promoting independent practice. Within the umbrella term computer-assisted language learning (CALL), computer-assisted pronunciation training (CAPT) encompasses a diverse group of tools that address the needs of most learners both in and out of the classroom (Golonka et al., 2014; Rogerson-Revell, 2021). Recent research has explored and confirmed the growing use of AI in language classrooms (Park & Doo, 2024; Pokrivčáková, 2019; Wang et al., 2024), particularly in enhancing pronunciation and overall oral competence. AI-powered tools typically include mobile apps, chatbots, virtual assistants, and AI-based modules within more complex learning programs or applications (Vančová, 2023). These tools have proven to be comparable in quality to human communicators and are being considered for implementation in the academic sphere as official evaluation tools (Isbell et al., 2024).

AI-powered tools significantly contribute to improving pronunciation accuracy, increasing motivation, and reducing speaking anxiety (Vančová, 2023). Notably, EFL and ESL learners using AI-based applications have shown greater improvements in pronunciation compared to those using traditional training methods (Mohammadkarimi, 2024; Shafiee Rad & Roohani, 2024). While learners generally express positive attitudes

towards AI-based pronunciation tools, challenges such as addressing individual pronunciation differences must be acknowledged (Mohammadkarimi, 2024). Overall, AI shows great potential for enhancing pronunciation training in both formal and informal learning settings (Vančová, 2023). Teachers must reflect on the latest technological advancements used in the classroom to better manage various aspects of teaching and organization, providing tailored feedback to all learners and creating uniquely designed input with a high level of individualization and automation.

Automatic Speech Recognition (ASR) technology plays a crucial role in providing feedback and evaluating learner performance (Vančová, 2021) by statistically analyzing pronunciation deviations and comparing the speaker's performance to a model integrated into the system. However, AI goes beyond acoustic analysis by simulating human-like responses to vocal input based on typical human behavior (Shafiee Rad & Roohani, 2024). Therefore, AI-powered tools can demonstrate the communicative value of pronunciation to EFL learners rather than merely providing drilling and imitation exercises.

AI-based tools offer significant advantages to learners by not only statistically evaluating input but also tailoring the adaptive nature of the output to the speaker, allowing them to implicitly receive feedback that mimics real-life communication. In a systematic review of AI-powered tools for pronunciation practice, Vančová (2023) concluded that different tools can be used for different learners, one-fits-all learners tool currently does not exist. Both off-the-shelf and specifically designed tools are employed for intentional or non-intentional pronunciation training. Many freely available learning apps integrate AI-powered features, enabling chatbots and virtual assistants to communicate with language learners in all environments and promote ubiquitous learning. Tailor-made tools can fine-tune specific pronunciation features for university students, while freely available tools help workers in tourism enhance their pronunciation to achieve specific communication goals.

In the classroom context, despite its limitations, AI offers a wide range of opportunities to improve teaching practices, carry out administrative tasks, and optimize workload, classroom management, and assessment (Ahmad et al., 2022; Celik et al., 2022; Lamas & Arnab, 2022; Mollick & Mollick, 2023). However, AI will likely not fully replace human teachers in the near future (Chan & Tsi, 2023).

In pronunciation training, EFL learners must address various pronunciation mistakes, including phonetic, phonemic, allophonic, and distributional errors (Moulton, 1962; Vančová, 2016) at both segmental and suprasegmental levels. Typical pronunciation training activities are based on sound imitation and sound discrimination (Couper, 2021; Pennington, 2021). The pronunciation goals can be defined in terms of accuracy, intelligibility and comprehensibility (Kennedy & Trofimovich, 2008; Thomson, 2017).

Based on the features of AI-powered tools for pronunciation practice, this study aims to answer the question: *How can teachers benefit from using AI-powered tools in the English pronunciation classroom to enhance their teaching practice?*

2. Method

2.1. Context and participants

Teachers who completed their formal education often face challenges in their own classrooms once they start teaching. The problems may be unique to each teacher's situation and conditions. Therefore, teachers need to improve their own teaching practices by educating themselves and implementing new techniques into their teaching. In action research as a research method, the teacher becomes a researcher and can solve the challenges by reflecting on their own actions and practices while teaching (Vančová, 2021). The typical steps in action research include identification and observation of classroom practices, investigation of the potential techniques they apply in teaching, practical implementation of new techniques in the classroom, evaluation of the achieved results and reflection on the teaching experience (for reference, see Nazari, 2022; Tobin et al., 2024).

The presented action research was conducted within a practical English phonetics course. This course primarily focuses on contextualizing information about the specifics of the English pronunciation system while also offering

opportunities to practice the articulation of theoretically described features in spoken texts. Over the years, the course techniques related to pronunciation practice have evolved. Initially, learners modeled their pronunciation on audio recordings using the traditional listen-and-repeat technique, with feedback provided individually through written comments by the teacher. Due to the range of pronunciation deviations from the model and the teacher's limited time, only the target sounds of the lesson were addressed in the feedback. In subsequent years, students benefited from ASR technology, which automatically indicated the specific pronunciation of individual speakers. Recently, with the availability of AI-based technology for educational purposes, there has been an opportunity to explore its potential to improve teaching practices.

The participants in this action research were two groups of first-year students of English language teaching and English language and culture (N=39). In the previous semester, they attended a course on English language phonetics and phonology, which included a module on practical pronunciation training.

2.2. The AI-powered pronunciation practice instruments observed

Action research aims to identify areas where the teacher-researcher can improve their teaching practices. Consequently, this study will focus on potential teaching improvements based on the type of tool and the suggestions of study participants, rather than evaluating the quality of a specific instrument. The three research tools will remain unnamed, but their nature will be described for context. All three observed tools are freely available online and were related to the pronunciation focus of the lesson.

Tool A was designed as a chatbot. Its interface allows users to choose an accent and a scenario to explore. Users speak into a microphone, and the chatbot holds a conversation with them. Participants were asked to complete at least three scenarios, each lasting a minimum of five minutes.

Tool B allowed learners to type in sentences they wanted to practice and drill. The program analyzed spoken answers on various language layers, including pronunciation. Alternatively, the software could generate sentences for speakers to practice and provide alternative training options.

Tool C involved asking users simple questions. After each recording, users received feedback regarding their accuracy on language levels or exam scores. Users could choose their practice targets, such as exam preparation or job interviews.

2.3. Data collection and analysis

Throughout the action research, students worked individually at home within the blended course. This work involved the practical implementation of pronunciation issues discussed in the in-class seminar and the pronunciation training exercises. The research encompassed three lessons, each incorporating an AI-powered tool for pronunciation practice. After each lesson, students completed a questionnaire with open-ended questions to analyze their attitudes and experiences with the instruments. This action research will present their reflections and recommendations from the perspective of teaching practices. Students were informed of the anonymous and voluntary nature of the questionnaires. The number of collected responses varied for different tools: 30 for Tool A, 24 for Tool B, and 15 for Tool C. After collecting all digital responses, they were downloaded and ordered according to the order of their submission. Each participant (hereinafter referred to as P) received a number that labeled their responses across all tools observed. All responses were read and analyzed for their content in terms of the topics they addressed. Then, codes were developed according to the topics the study focused on and the participants addressed (i.e., perceived ease of use, perceived accuracy of the tool, and engagement of the learners while using the tools).

3. Results

3.1. Tool A

The first observed tool was a chatbot with various scenarios and role-plays. For some students, this was the first experience with the chatbot (*"It was something new for me, I have never tried to make a conversation with AI before. I don't see any minuses only great opportunity for improving speaking and reading skills,"* participant 1; *"I have never done anything like that before,"* P6). The students mostly appreciated the opportunity to practice their speaking skills (*"a good language practice, engaging with the speaker than automatically they get feedback,"* P8). Other recognized positives was the engaging and useful nature of the tool in improving speaking skills (*"I found it useful to have different scenarios to choose from and the AI responses were helpful in keeping the conversation going,"* P12). Due to the adaptability of AI-powered tools, the conversation skills of students also improved (*"I improved my creative thinking in terms of creating new topics to talk about,"* P16; *"the questions AI asked were relevant to the real world, which I find very useful and I will definitely use it again to improve my speaking skills,"* P21; *"I can imagine a person can improve not only with language but also with social interactions,"* P15). In addition, learners can use the chatbot for preparation of real-life conversations (*"it could be useful for practicing before going to an English-speaking country as the scenario I tried would be a common thing on a vacation/trip,"* P27). In particular, students improved pronunciation of specific words, such as *sourdough bread*, *dietary*, *aligns* (P8), *psychology* (P7) or general speaking fluency (P3, P12, P28, P34).

The participants would suggest using the tool for younger or beginner learners (*"I think it would be very good for younger students as I we are already able to deal with most of these situations pretty easily... throughout our studies we have had roleplays with several of the same topic that are available with the AI,"* P27).

However, the participants could also see the artificial nature of the conversation (*"I don't think we can compare a bot chatting with a real-life communication, where the person has feelings and opinions. Chatting/talking with a bot seems very strange and unnatural to me,"* P28; *"It was a bit strange to have a human-like conversation with an AI. The tool contradicted itself a lot. Some it misheard a word I said and the conversation seemed to be going in circles,"* P5; *"I want to talk to a person not an AI,"* P17). The participants also reported minor technical problems related to the quality of the technical equipment available to them or acoustic conditions.

From a pedagogical perspective, a scenario-based chatbot appears to be a useful tool for students who are open to using technology in language learning but have limited opportunities to speak in English. This tool can also be beneficial in the classroom context. The wide range of scenarios allows for exploring different topics across various proficiency levels. Learners can improve the pronunciation of particular words, as well as segmental and suprasegmental features, and appreciate the communicative value of pronunciation by observing how mispronunciation can divert the direction of a conversation within a context. The feedback collected from participants was predominantly neutral to positive, with no fully negative reactions toward the chatbot. However, future use should address the limitations of chatbots in pronunciation practice and provide more structured instructions from the teacher. This will enable students to test their pronunciation skills and overall speaking proficiency across different scenarios in a meaningful way.

3.2. Tool B

The second tool was accuracy focused and overall generated more critical responses by the participants due to the perceived sensitivity to pronunciation (*"that program sometimes didn't recognize my pronunciation, especially didn't recognize such easy words like "hey" or "okay." P2).* Due to the use of the decontextualized practice of sentences and lack of implicit feedback, the participants did not recognize the mistakes they made (*"When I tried to speak and correct myself multiple times it usually didn't even pick up what was wrong and still didn't like the way I said it. Even after trying for an hour, it was still the same. It didn't really tell what to fix in the word that had 0 points as well,"* P3). The lack of context in pronunciation practice was not engaging for all readers (*"I disliked that the generated sentences were repeating – dull,"* P23; *"It gave same phrases. Or it just changed some words*

in the sentence,” P29). However, some participants could see this as an opportunity to independently explore their pronunciation (“Also the accents percentage might have been wrong and I would improve it by adding more feature and deeper analysis of the speech and different variations of sentences,” P5; *I’ve noticed how small changes in pronunciation can increase as well as decrease my English level and now I may have a better understanding of distinguishing features in American and Britain English,*” P21). To confirm this assumption, P9 finetuned pronunciation of selected words (*particularly, comfortable, nuclear*). However, the tool was predominantly appreciated by more proficient participants (“*I have proficiency C1 overall... I did the test without any preparation and I did it fast. I did not overthink the answers, I simply pretended I was speaking spontaneously to a native speaker,*” P18).

From a pedagogical perspective, this tool appears to be most suitable for drilling exercises. Due to its ability to adapt text to the learner’s needs, any challenging sound can be finetuned using words selected for a particular learner. This feature was particularly appreciated by more proficient learners of English. However, applying this tool with younger or beginner learners may not encourage engagement and should only be used as a complementary, rather than a primary, source of activities.

3.3. Tool C

The last tool combined features of the previous two tools. It gave students open answers to respond to and then adapted the subsequent questions to their responses. However, the scenario was not modeled for learners as a context. This has proven to be an advantage for some participants (“*I would use it in future if I wanted to practice speaking skills on any given topic, since you do not know the question beforehand – advantage,*” P23; “*I improved my confidence and found out some of the new synonyms,*” P21) and disadvantage for other participants (“*However, I disliked the limited vocabulary and topics available for practice, which did not cater to my specific interests,*” P12). The overall evaluation of the tool was positive for students with little confidence and the possibility to speak with human speakers (“*good questions, nice and effective tool for learning for someone who has no one to speak to,*” P23; “*The only challenge I encountered was with my pronunciation. However, this tool helped me improve my vocabulary and conversational skills. ...I would use this tool in the future as it will help me enhance my speech and communicate freely on various topics,*” P24).

4. Discussion, conclusions and recommendations

Recent teaching practices have confirmed the growing use of AI in language classrooms (Pokrivcakova, 2019; Wang et al., 2024; Park & Doo, 2024). One of the primary applications of AI is enhancing EFL and ESL pronunciation. AI-powered tools, such as mobile apps, chatbots, virtual assistants, and AI-based modules, can be successfully implemented in teaching under specific conditions, including the meaningfulness of the tool (realistic scenarios), greater adaptability (sensitivity and responsiveness to the learner’s input), reliability, and detailed feedback for learners requiring guidance (visual or verbal evaluation of the input; Vančová, 2023). A carefully selected tool can be an effective aid for language learners (Pokrivcakova, 2022, 2024), but it appears that these tools will not replace teachers anytime soon (Chan & Tsi, 2023), even though their benefits in various spheres of teaching and learning have been proven (Ahmad et al., 2022; Celik et al., 2022; Lamas & Arnab, 2022; Mollick & Mollick, 2023).

The action research focused on three different AI-powered pronunciation practice tools presented to university learners of English: a scenario-based chatbot, a drilling adaptive program, and an open-ended conversation with a chatbot. The participants’ experiences, collected via questionnaire, revealed that these three tools can be used for various purposes for different groups of learners. The overall response to the tools was positive, although different tools generated varying levels of engagement. A comprehensive conclusion is provided in Table 1.

Table 1. Suggestions for recommended use of AI-powered tools in English pronunciation classrooms based on the classroom experience

Tool A (scenario-based chatbot)	Tool B (adaptive pronunciation practice)	Tool C (adaptive pronunciation checker)
<i>Pedagogical advantages</i> <ul style="list-style-type: none"> - Preferred by less proficient learners - Substitution of human interaction in EFL settings - Prepare learners for real-life experience - Relatively new for learners 	<i>Pedagogical advantages</i> <ul style="list-style-type: none"> - Adaptive to learners' pronunciation training needs - Addressing different goals (e.g. exams) - Detailed feedback - Preferred by most advanced learners 	<i>Pedagogical advantages</i> <ul style="list-style-type: none"> - Identification of learners' needs - Adapting the practice content - Detailed feedback
<i>Pedagogical disadvantages</i> <ul style="list-style-type: none"> - No explicit feedback required by proficient speakers - Limited number of scenarios - Unnatural 	<i>Pedagogical disadvantages</i> <ul style="list-style-type: none"> - Not engaging enough for less proficient learners - High demand for pronunciation accuracy 	Pedagogical disadvantages <ul style="list-style-type: none"> - Not engaging for less proficient learners

The chatbot design appears to be the most engaging for learners; however, more proficient learners found it only slightly challenging compared to the drilling adaptive tool. The latter might not always provide immediate explicit feedback but allows learners to independently explore the peculiarities of their pronunciation. On the other hand, chatbots prepare learners for authentic communication better than drilling exercises. Overall, each tool can be implemented at different stages of language learning for specific purposes. Encouraging students to engage with AI-powered technology can significantly benefit their pronunciation accuracy, comprehensibility, and overall oral competence.

Acknowledgements

I would like to express sincere gratitude to the research participants for their contributions to explaining their attitudes toward the AI-powered tools used in the presented study. Furthermore, I would like to thank two paper reviewers for their valuable and constructive feedback.

References

- Ahmad, S. F., Alam, M. M., Rahmat, M. K., Mubarik, M. S., & Hyder, S. I. (2022). Academic and administrative role of artificial intelligence in education. *Sustainability*, 14(3), 1101. <https://doi.org/10.3390/su14031101>
- Couper, G. (2021). Teacher cognition of pronunciation teaching: The techniques teachers use and why. *Journal of Second Language Pronunciation*, 7(2), 212-239. <https://doi.org/10.1075/jslp.20004.cou>
- Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends*, 66(4), 616-630. <https://doi.org/10.1007/s11528-022-00715-y>

- Chan, C. K. Y., & Tsi, L. H. (2023). The AI revolution in education: Will AI replace or assist teachers in higher education?. *arXiv preprint arXiv:2305.01185*.
- García, C., Nickolai, D., & Jones, L. (2020). Traditional versus ASR-based pronunciation instruction. *Calico Journal*, 37(3), 213-232. doi: 10.1558/cj.40379
- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer assisted language learning*, 27(1), 70-105. <https://doi.org/10.1080/09588221.2012.700315>
- Isbell, D. R., Crowther, D., & Nishizawa, H. (2024). Speaking performances, stakeholder perceptions, and test scores: Extrapolating from the Duolingo English test to the university. *Language Testing*, 41(2), 233-262. <https://doi.org/10.1177/02655322231165984>
- Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review*, 64(3), 459-489. <https://doi.org/10.3138/cmlr.64.3.459>
- Lameras, P., & Arnab, S. (2022). Power to the teachers: An exploratory review on artificial intelligence in education. *Information*, 13(1), 14. <https://doi.org/10.3390/info13010014>
- Mohammadkarimi, E. (2024). Exploring the use of artificial intelligence in promoting English language pronunciation skills. *LLT Journal: A Journal on Language and Language Teaching*, 27(1), 98-115. <https://doi.org/10.24071/llt.v27i1.8151>
- Mollick, E. R., & Mollick, L. (2023). Using AI to implement effective teaching strategies in classrooms: Five strategies, including prompts. *The Wharton School Research Paper*. <http://dx.doi.org/10.2139/ssrn.4391243>
- Moulton, W. G. (1962). Toward a classification of pronunciation errors. *The Modern Language Journal*, 46(3), 101-109.
- Nazari, M. (2022). Plan, act, observe, reflect, identity: exploring teacher identity construction across the stages of action research. *RELC Journal*, 53(3), 672-685. <https://doi.org/10.1177/0033688220972456>
- Park, Y., & Doo, M. Y. (2024). Role of AI in Blended Learning: A Systematic Literature Review. *International Review of Research in Open and Distributed Learning*, 25(1), 164-196.
- Pennington, M. C. (2021). Teaching pronunciation: The state of the art 2021. *RELC Journal*, 52(1), 3-21. <https://doi.org/10.1177/00336882211002283>
- Pokrivcakova, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education*, 7(3), 135-153. <https://doi.org/10.2478/jolace-2019-0025>
- Pokrivcakova, S. (2022). Teacher trainees attitudes towards integrating chatbots into foreign language classes. In *INTED2022 Proceedings* (pp. 8294-8302). IATED. <https://doi.org/10.21125/inted.2022.2108>

- Pokrivcakova, S. (2024). Pre-service teachers' attitudes towards artificial intelligence and its integration into EFL teaching and learning. *Journal of Language and Cultural Education*, 11(3), 100-114. <https://doi.org/10.2478/jolace-2023-0031>
- Rogerson-Revell, P. M. (2021). Computer-assisted pronunciation training (CAPT): Current issues and future directions. *Relc Journal*, 52(1), 189-205. <https://doi.org/10.1177/0033688220977406>
- Shafiee Rad, H., & Roohani, A. (2024). Fostering L2 Learners' Pronunciation and Motivation via Affordances of Artificial Intelligence. *Computers in the Schools*, 1-22. <https://doi.org/10.1080/07380569.2024.2330427>
- Thomson, R. (2017). Measurement of accentedness, intelligibility, and comprehensibility. In *Assessment in second language pronunciation* (pp. 11-29). Routledge.
- Tobin, B., Farren, M., & Crotty, Y. (2024). Impacting teaching and learning through collaborative reflective practice. *Educational Action Research*, 1-21. <https://doi.org/10.1080/09650792.2024.2394933>
- Vancova, H. (2016). The pronunciation mistakes of Slovak learners of English. *Journal of International Scientific Publications: Educational Alternatives*, 14(1000021), 264-272.
- Vančová, H. (2021). Teaching English pronunciation using technology. Nümbrecht: KIRSCH-Verlag
- Vančová, H. (2023). AI and AI-powered tools for pronunciation training. *Journal of Language and Cultural Education*, 11(3), 12-24. <https://doi.org/10.2478/jolace-2023-0022>
- Wang, D., Tao, Y., & Chen, G. (2024). Artificial intelligence in classroom discourse: A systematic review of the past decade. *International Journal of Educational Research*, 123, 102275. <https://doi.org/10.1016/j.ijer.2023.102275>