sciendo

AI and AI-powered tools for pronunciation training

Hana Vančová
Trnava University in Trnava, Slovakia
hana.vancova@truni.sk

Abstract: Artificial intelligence (AI) is taking over many spheres of human life, including language pedagogy. While some areas need to find their ground with AI and resolve ethical issues arising from its use, other spheres of education, such as pronunciation, may benefit from the system's ability to communicate with learners and provide them with implicit feedback while carrying out different communication tasks. This technical advancement of AI then opens doors for further educational opportunities that have not existed in the past. The study aims to explore the existing use of AI-powered tools in foreign language pronunciation training by meta-analysis of 15 research papers benefiting from using AI and AI-powered tools (mobile and web applications, chatbots, intelligent virtual assistants) and provide suggestions for their future applications in educational practice. The research results also indicate that this field of study is still underrepresented in language pedagogy. The existing experience with AI-powered tools confirms a relatively good experience in developing intelligibility, increasing motivation and addressing the speaking anxiety of foreign language learners in formal and informal learning settings.

Key words: artificial intelligence, AI-powered tools, pronunciation, ELT, education technology

Introduction

Bringing technology to language pedagogy has undeniable benefits (Neri et al., 2002; Benzies, 2017; O'Brien et al., 2018; Yoshida, 2018; Pokrivčáková, 2019; Rogerson-Revell, 2021; Vančová, 2021). Especially for pronunciation technology has always been present in some form in its training. Literature generally identifies CAPT (computer-assisted pronunciation training) as a subfield of CALL (computer-assisted language learning), with ASR (automatic speech recognition) and AI (artificial intelligence) technology being one of the most significant components of its use. Technologies bring tailor-made, customised training with increased opportunities to exposure and practice of the target language, help develop learner autonomy, can provide a suitable model when a native speaker is not available, and many other advantages. The role of technology in the current pedagogical framework varies. While some believe technology should be integrated into the current paradigm, others call for a shift and necessary changes (Pokrivčáková, 2019). Rogerson-Revell (2019) is critical towards technologies in pronunciation practice, as their implementation is regressive and narrows the techniques to drilling and repetition, which contradicts commonly used methods used in the communicative approach in ELT. Digital tools can, in a specific case, provide incomplete information, which can even be inaccurate, or presents the same exercises for all learners, irrespective of their needs.

AI and AI-powered tools in pronunciation training

AI and AI-powered tools present the latest technological advancements in language teaching. As a part of computer science, AI is based on a set of technologies, methods and approaches that are used for performing tasks typical for humans. This machinery uses data mining, machine learning, and natural language processing and is influential in the period of "Education.4.0" (Almelhes, 2023, p. 1259). Chen, Chen and Lin (2022) traced the application of AI tools from the beginning. While in the early stages, AI took the form of computers or humanoid robots, now it is available as assistants, teacher colleagues, sensors, chatbots or independent instructors. Rogerson-Revell (2021) sees robots, talking heads and other embodied AI devices as the future of pronunciation training.

To characterise AI in general, Li (2022, p. 4) views it as "a new technical science that studies and develops theories, methods, technologies and application systems for simulating, extending and expanding human intelligence. It refers to making the machines produced by people have the cognitive behaviour, thinking ability and learning activities of the outside world like people, and simulate human thinking". Hockly (2023) distinguishes between the "weak" and "strong" AI tools – creating the strong AI which would be able to perform mental operations as humans is the ultimate goal of AI developers. However, at this stage, only weak AI is available now in education,

as it can perform relatively simple tasks (translation, error correction, structured practice) within a given field of their application.

The potential of such systems in education is virtually endless because "AI could also change people's fundamental understanding and practices of teaching and learning" (An et al., 2023, p. 188). AI systems are based on the system's ability to imitate human behaviour (Almelhes, 2023) and even learn from interactions with humans and other systems. In language education, AI is based on natural language processing, speech synthesis and recognition (Shufang, 2021). In practice, AI tools analyse human linguistic behaviour, create their own language model and apply it in communication with human speakers. Lee and Lim (2023) introduce an example of a teachable agent who analyses a learner's pronunciation and then adapts it to the learner's developing skills.

Such tools also require new approaches and responsibilities in their use. Ng et al. (2023, p. 142) summarise the overview of AI competencies that would allow people simply "critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool online, at home, and in the workplace". The authors suggest a deeper understanding of the workings of AI, recognising its implementation in various devices, ability to critically evaluate its function, ability to reason with the technology, etc. Eventually, they relate the AI competencies to Bloom's taxonomy of cognitive domains (Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation).

The educational community has yet to embrace AI in its used and must resolve ethical issues (Ng et al., 2023; Hockly, 2023) surrounding it. One of the critical issues related to the use of AI is the storage of information and ownership of the data collected by intelligent tools. The authors summarise the most frequently cited concerns and challenges in using AI, such as the fear of teachers being replaced by AI tools, the incorrect information and guidance of AI tools, and privacy concerns related to data collection by the tools. In addition, academic misconduct related to machine-generated texts is another concern the educational field may face (Gao et al., 2022; Khalil & Er, 2023). Iphofen and Kritikos (2021) appeal to policymakers and designers to closely collaborate when developing new intelligent technologies as their creative abilities grow, but this brings potentially malicious events which should be prevented from the earliest stages. AI technology can even be labelled "disruptive" (Sumakul, Hamied & Sukyadi, 2022). However, pronunciation training can benefit from the systems' ability to recognise elements of human speech and thus train speakers to communicate in the target language.

Pokrivčáková (2019), Ng et al. (2023), and An et al. (2023), among others, point to the importance of equipping teachers with new competencies on technical, instructional and ethical levels. The teachers should get support in implementing the AI technologies, which have been fast-forwarded primarily into distance education, whether regular or emergency. It is necessary to point out that using technologies should align with the existing educational framework and always help learners achieve the expected learning outcomes. The selection of any digital tool should not compromise the achievement of the learning outcomes (Yoshida, 2018). It is equally important to identify the instrument's quality and the users' attitudes towards its use. Davis (1989) developed a Technology Acceptance Model (TAM) where the perceptions of the users regarding the perceived usefulness and ease of use influence their future intentions to use the tool in various contexts, pointing at the paradoxical nature of AI tools in education: on the one hand, they allow creating learning courses available to almost an unlimited number of learners worldwide, on the other hand, they provide learners with almost personal guidance due to the level of automation in the process.

Intelligent systems can provide learners with limitless practice according to their proficiency level (also levelled learning, Lee & Lim, 2023) without the often-cited anxiety related to training with a human teacher (El Sazly, 2021; Hsu, Chen & Todd, 2021; Moussali & Cardoso, 2020), and freeing teacher's attention from time-consuming correction of errors the learners make and administration, or increasing teaching efficiency through introducing a high level of automation and personalisation in education (Ng et al., 2023). The teacher then becomes a facilitator in the classroom (Jin, Han, Ko, 2023). The learning environments, often virtual, can be realistic and change and develop with time.

Willingness to communicate based on the learner's perceived linguistic competence and low anxiety level in communication (Ayedoun, Hayashi & Seta, 2019) can be directly promoted through the use of embodied conversational agents, i.e., computer-generated characters which to a large extent resemble human-human interaction and add more naturalistic dimension in this interaction. AI can resolve learning and evaluating pronunciation in two ways. For example, from the perspective of training, they can replace native-speaking teachers who appear to be preferred in a pronunciation classroom. From the evaluation perspective, AI technology can improve learners' pronunciation accuracy.

The origins of using technology to calculate deviations from a model pronunciation trace back to the 1990s (Shufang, 2021) when the then-used system evaluated features of segments and sentences that were not pre-specified. However, in the early 2000s, the systems based on ASR evaluated pronunciation differently from human

raters across more prolonged stretches of spoken texts, while the evaluation of single phonemes was similar. Native speakers of the language collected the results – the added layer of difficulty appears to be a foreign or second language pronunciation.

Lee et al. (2010) maintain that being exposed to comprehensible input promotes learners' semantic processing, allowing learners to produce output supports developing their syntactic processing. Thus, in communication, the speaking participants get feedback on their production based on the reaction of their listening counterparts. In this respect particularly robots can raise motivation to improve pronunciation by correcting erroneous production of young learners.

However, although much research has been done into the system's functionality, relatively little in comparison has been done to investigate pedagogical opportunities and challenges it brings in pronunciation training. Shufang (2021) refers to this as human-computer interaction (or human-machine communication), which comprises elements of speech recognition, natural language understanding, and speech synthesis. Speech synthesis, in particular, is the bridging element between this type of communication and artificial intelligence.

ASR is an inherent part of AI systems. In pronunciation training, ASR-based systems are direct predecessors of AI-powered systems. Bashori et al. (2022) identified two primary advantages of using ASR in the classroom: providing a better pronunciation model than the one supplied by non-native teachers. The second one is the ability of learners from the FL context to overcome the obstacle of finding a communication partner. However, their results have proven that a careful selection of materials and teacher preparation is essential for improving FL learners' pronunciation. Bajorek (2017) came to a similar conclusion and recommends using ASR technology as an additional instrument to traditional classroom pronunciation training. There may need to be more than indirect implicit feedback for FL learners who must be explained and guided towards desirable pronunciation.

Numerous studies have proven the effectiveness of using ASR for accuracy (Wang & Young, 2015; Sigdi & Shaari, 2017; McCrocklin, 2019). However, with the new adaptive systems imitating human interaction, AI-powered tools could also be used for developing intelligibility, as the tools respond to the recognised speech of FL learners. There are many forms in which AI-powered tools enter language learning (see Pokrivčáková, 2019). The most popular for pronunciation practice are intelligent personal assistants, chatbots and learning applications integrated with chatbots.

The difference between "automatic" and "intelligent" seems to lie in the degree of manipulation of the gathered data. Automated tools compare the data against the model and give feedback based on the articulators' incorrect position or quantity of the phoneme (Dai, 2022). The ASR system's main modules are oral language assessment, pronunciation error detection, and corrective feedback. Intelligent tools, on the other hand, make further decisions about the use of the data regarding the response to the user or learner. Thus, while automatic tools follow a predetermined path, intelligent tools allow for a certain degree of spontaneity in task-solving. AI-powered tools do not provide learners with completely original or unexpected reactions but give a more significant variation in the interaction between the system and the learner.

Korzekwa et al. (2022) argue that one of the main shortcomings of intelligent pronunciation assessment tools is rooted in the lack of erroneous pronunciation models because these systems tend to generate pronunciation models from native speakers, who are often trained in elocution (Bogach et al., 2021; Rogerson-Revell, 2021).

The success of spoken communication between a digital tool ("artificial agent", Foster & Stuart-Smith, 2023) and a human speaker relies on various factors, including the ability of the tool to recognise the speaker's voice, speaker's preference for a female/male voice of the tool, their attitude to computers in general and artificial agents in particular, etc. The authors also note extending the range from traditionally used Southern English and General American accents.

There are different types of AI-powered technologies for language learning (see Pokrivčáková, 2019). In pronunciation practice, Intelligent Personal Assistants, chatbots and learning applications are the most frequently used and cited tools in research studies.

Intelligent Personal Assistants (IPAs) integrated into widely available home appliances and devices (smartphones, home assistants, chatbots) commercially known as Amazon Alexa or Echo, Siri for iOS, Google Home, etc.) break the traditional learning school setting. Using IPAs gives the context of cooperation or communication to achieve another goal. Their typical use is to assist people while carrying out various tasks, typically looking for factual information accessible online or performing various service tasks. They transform human speech into text, identify and locate the information online, and transform it into human-like speech. Pronunciation training is then an added benefit of using IPAs. However, their ability to recognise non-native speech can be challenging (Dizon, 2021).

Moussali and Cardoso (2020) summarise the key features of IPAs as tools for developing learners' autonomy, self-access language learning (SALL), motivation, and informal tools. They allow learners to communicate with them

during native speakers' absence; however, while human speakers can overcome the obstacles of non-canonical pronunciation of foreign speakers, these tools require a high degree of accuracy. When API misunderstands the speaker, it can ask an additional question, comment, no response or incorrect answer as indirect feedback and an opportunity to try alternative pronunciation. However, the study revealed that learners might abandon an original goal instead of applying a compensation strategy (using a synonymous expression, providing the command in another form, or asking indirectly using another piece of knowledge). This experience contradicts the earlier research results with young learners (Underwood, 2017), who were willing to rephrase their commands when they were misunderstood.

Annamalai et al. (2023) perceive *chatbots* as potential educational tools that interact and simulate human interaction. Lim et al. (2022) argue that they may be embodied and have faces or avatars. They also perceive them as communication tools that are a step below personal assistants. Using keywords, chatbots simulate human conversation through voice or text (Ji, Han, y Ko, 2023) (Rogerson-Revell, 2021). They have proven their qualities in educational settings, but other spheres must also be explored. Hockly (2023) and Ji, Han and Ko (2023) recognise their relevance for pronunciation training based on repetition and drills. Huang, Hew and Fryer (2022) add a social aspect to their use and other advantages. However, Lee and Lim (2023) highlight the importance of chatbots being responsive and adaptive to learners' developing linguistic competence and providing varied content and responses to reflect the communication with the learner.

AI-powered learning applications tend to be available as mobile applications for general and particular groups of learners (e.g. Bussuu for beginner learners; Duolingo effective for gifted children, Hazar, 2022). Mobile applications provide a multimodal practice and presentation of language with an unlimited amount of exercise, respecting students' proficiency and allowing them to progress at their own pace in formal and informal settings. Alfuhaid (2021) identified that Duolingo, in particular, can enhance learners' self-confidence in oral communication because its users are used to communicating in the target language. The practice is self-directed, and the applications benefit from data mining and learning analysis (Portnoff et al., 2021). They can also be used in academic settings to prepare students for examinations and tests (Brick & Cervi-Wilson, 2019).

Developing the pronunciation component is currently carried out within the broader context of speaking skills. Speaking is inherently correlated to pronunciation. Pronunciation instruction does not only include training of particular sounds to achieve a total accuracy of individual sounds (imitation of native speakers, nativeness principle; Levis, 2005, 2020; Vančová, 2019) but also includes prosody (intonation, rhythm, stress) as a mark of meaning change in words. However, achieving accuracy is a relatively difficult yet non-essential task for most non-native learners of foreign languages. Thus, achieving speaking comprehensibility or intelligibility is the current pronunciation training paradigm. Accentedness, i.e. the degree of difference between native speech, is relatively easily measured through digital devices; the degree of comprehensibility (the amount of effort the listener must put into understanding someone and intelligibility (how much the listener actually understood) are rated by human raters and are not directly related, as a person with a strong accent can be easily understood. Accentedness evaluation may be stricter than the comprehensibility rate. Comprehensibility and intelligibility depend on various factors, including the listener's willingness to understand a non-native speaker (Derwing & Munro, 1997, 2009; Vančová, 2021). Systematic deviations from native speech are easy to overcome; however, prosody and lexical stress, in particular, play an important role in comprehension (Field, 2005). To achieve intelligibility, Jenkins (2006) presented Lingua Franca Core for pronunciation, which comprises a set of core and non-core pronunciation features promoting intelligibility among speakers based on pronunciation features of various English accents.

In this regard, Vančová (2020) reported views of EFL teachers that while professional interpreters and, to a certain extent, translators should be proficient in more than one variety of English due to the need to work with speakers of various backgrounds of English including non-native speakers, teachers are expected to master a particular native variant of English they teach and be consistent in its use in their classroom. Thus, teachers are expected to achieve a high level of accuracy. At the same time, other professional users of English should focus on achieving high levels of listening comprehensibility.

The currently available research provides theoretical and practical challenges in teaching pronunciation to non-native speakers and identifies those features the AI-powered systems need to provide. However, a systematic study putting an originally designed tool into practice with real foreign language learners is relatively rare. Thus, pronunciation training and the subsequent evaluation through AI-powered tools are most frequently carried out using commercially available learning applications and programs. Non-specified systems are being tested for accuracy by a non-defined group of speakers (Jiao et al., 2021; Shufang, 2021; Cheng & Wang, 2022) and indicate future improvements of the tools.

Research objectives

Due to the increasing presence of AI in language pedagogy, the study aims to answer the following three research questions:

1. Which AI-powered tools are used in pronunciation training?
2. What is the focus of the pronunciation training using AI-powered tools?
3. What are the results of such pedagogical treatment?

Methodology

A method of meta-analysis of selected research papers published in scholarly journals and conference proceedings was selected to fulfil the formulated objectives of the study. The main criteria for including analysing papers were search results using key words "artificial intelligence" and "pronunciation", with alternatives including "AI-powered", "speech error", and additional terms such as "ELT", "EFL", "ESL", "chatbot", and "intelligent personal assistant". The alternative terms were selected due to the immersion of both focal aspects, artificial intelligence and pronunciation, into technological and language systems. As mentioned above, both aspects are an inherent part of their respective systems and cannot be separated from the larger context they are used in. For this present review, pronunciation training is understood as (1) an activity improving speaker's pronunciation accuracy, (2) an activity improving speaker's comprehensibility,

To identify the papers suitable for this review, databases *Web of Science, Springer Link, Elsevier, SAGE, ERIC, Taylor & Francis, Scopus* and *Wiley Online Library*, and additionally, *Google Scholar* were searched for keywords selected. Other criteria included the availability of the full-text version online, the type of the study and the language of the study (English). Review papers, technical descriptions of systems and final theses were excluded from the selection. No time limit for paper publishing was set due to the assumption that the AI technology is relatively new to its application to pronunciation training, and observing a more extended period of its use beyond the state-of-the-art pedagogical practices could be beneficial for creating a context and observing application trends.

The abstracts of papers meeting the criteria in the first stage of the selection process were then brought upon closer inspection for their significance in the analysis. Studies not relevant to the objective of this paper were eliminated. It allowed forming a representative sample of 15 articles revolving around the fundamental concept of the study, which is the use of AI for pronunciation training. Although the primary aim was to analyse training in English pronunciation, this criterion was later abandoned due to the interest in the application of the tools in an educational context; training pronunciation in languages other than English (Japanese) was also included in the process of analysis as potential samples of good practice. The participants of the selected studies did not involve only pre-service and in-service teachers or students of English but foreign language users in general.

After carrying out both stages of the selection process, a sample of 15 papers was created. The sample includes articles found in journals *Computers and Education: Artificial Intelligence, Expert Systems, Indonesian EFL Journal, Interactive Learning Environments, International Journal Of All Research Writings, Journal of Asia TEFL, Journal of English Language Teaching in Foreign Language Context, Journal of Mekong Societies, Language Testing, LEARN Journal: Language Education and Acquisition Research Network, Proceedings of the 4th Conference on Conversational User Interfaces, Speech Communication, Sustainability* and *Tesol Journal*. This selection of publications indicates the growing interest in education publications in AI and the shift from computer science to other spheres of life.

It must be noted that the number of identified papers was lower than expected before the search. This result may be supported by the findings of Huang et al. (2023, p. 112), who analysed 516 papers published between 2000 and 2019. Pronunciation development using "intelligent tutoring and assessment systems for pronunciation and speech training" was the ninth most researched area identified by authors, confirming the often-cited claim that pronunciation is the overlooked component of language proficiency. In addition, a comprehensive review of developing interactional competence through AI dialogue systems in university students comprised 28 papers, and only four dealt with pronunciation (Zhai & Wibowo, 2023). In 2020, Moussali and Cardoso identified only three other studies on using intelligent personal assistants for education. Research studies into grammar and vocabulary teaching through AI-powered tools occur more frequently, which may be caused by higher predictability of grammatical and lexical patterns than the predictability of unique pronunciation patterns of individual learners.

A systematic review of putting AI-powered tools into pronunciation classrooms presents a gap in the current knowledge and is vital for foreign language pedagogy.

Results

The following section of the paper provides insight into specific conditions in which AI and AI-powered tools are used for pronunciation training. More specifically, research participants, pronunciation focus, tools used, and

concrete conditions for their implementation will be presented. The overview of the studies is presented in Table 1 (Appendix 1).

Research participants

A closer look at the participant sample is necessary to understand who can benefit from using AI-powered tools in pronunciation practice. One of the critical features of the presented study is the variety of participants involved in the analysed studies.

Regarding sample size, the smallest research sample size consisted of two participants who attended a junior high school in Indonesia (Citrayasa, 2019), allowing for a more qualitative investigation into the participants' experiences. They were also the youngest participants in the research studies. On the contrary, the largest sample of a single study (N=304) consisted of speakers who were being evaluated and stratified groups of human raters involved in comparing the accuracy of automatically administered speaking tests (Isbell, Crowther & Nishizawa, 2023). This number of participants allowed a statistical evaluation of the collected data and possible extrapolation of results to other spheres of the use of AI-powered tools.

As for the educational background, the majority of participants came from the context of higher education (in the process of admission, students, and teachers). Khampusaen, Chanprasopchai & Lao-un (2023) reported the most varied group of learners, who trained local tourism professionals with different educational backgrounds, and their ages spanning from 16 to 49 years old.

While the majority of studies focused on learners or users in terms of their performance and attitudes, one study (Sumakul, Hamied, & Sukyadi, 2022) investigated teachers' experiences using AI in their classrooms.

This brief overview of the participant structure suggests a potentially universal and beneficial use of AI-powered tools across different groups of learners.

Pronunciation focus

As suggested in the previous parts of the study, there are two main domains of pronunciation training: accuracy (or its counterpart accentedness) or comprehensibility (in some contexts interchangeable with intelligibility). As the AI systems are more advanced and inherently contain ASR technology, the primary goal of most studies was not speaking accuracy'. However, three studies are exceptions. Using a mixed-method pretest-posttest design, Noviyanti (2020) used a spell-checker as an alternative tool for independent learning in a pandemic situation. The researcher asked learners to use a spell checker and found out that, on average, the participants' pronunciation improved on average 33 points in the posttest. The researcher concentrated on silent letters s, k, h, b and t. In addition to pronunciation changes, the respondents expressed generally favourable opinions on value, effectiveness and attitudes towards the tool and promotion of independent learning in the subsequently administered questionnaire.

Chuyen et al. (2021) used Duolingo in an experimental study carried out with forty-one 16-year-old Thai students to improve particular pronunciation features (using dental and dental fricatives, alveolar plosives in verb endings, silent h sound on the segmental level and using word stress in 2- and 3-syllabic words and falling and rising intonation). The experimental group used the application for six weeks. After the treatment, participants in the experimental group proved improvements in the pronunciation mentioned above.

Chung and Bong (2022) focused on improving Lingua Franca core and non-core features (see Jenkins, 2006), as the authors identify them as key to the intelligibility of Asian EFL learners. In particular, Korean-accented English was central to the study. The researchers carried out two experiments. Participants completed an intelligibility cloze test based on listening to 100 sentences in the first part. In the second experiment, participants listened to 100 sentences pronounced by a Google Assistant. The results revealed that human raters were more successful in recognising Korean-accented speech than the AI application. However, with the increasing difficulty of recognising sounds by the listeners, the difficulty of AI applications also grew. The threshold for intelligibility was set at 60% of the recognition rate.

The remaining studies benefited from AI in developing aspects of general speaking competence, with pronunciation being one of its components. More specifically, AI-powered tools were used to improve fluency (Hou, Chen & Todd, 2021; Zou et al., 2023), intelligibility (Chung & Bong, 2021), motivation (Khampusaen, Chanprasopchai & Lao-un, 2021), overcoming FL speaking anxiety (Dizon, 2017; Hou, Chen & Todd, 2021), using pronunciation for specific purposes (tourism in Thailand, Khampusaen, Chanprasopchai & Lao-un, 2023; administering large quantities of speaking tests, Isbell, Crowther & Nishizawa, 2023), or analysis of emotional connection to chatbots and acceptance of the AI-powered tools (Cohn et al., 2021; Zhu et al., 2023).

AI-powered tools used and observed

Three main types of AI-powered tools were used and analysed within the selected sample of papers.

The most dominant group of papers (N=9) comprised studies based using *learning apps with AI-powered modules including a chatbot* (Citrayasa, 2019; Chuyen., Linh & Phuc, 2021, El Shazly, 2021; Sumakul, Hamied & Sukyadi, 2022; Zhu et al., 2022, Annamalai et al, 2023; Isbell, Crowther & Nishizawa, 2023; Khampusaen, Chanprasopchai, & Lao-un, 2023; Zou et al. 2023).

The second largest group of papers (N=5) focused on the use of *IPAs Alexa and Google Assistant* (Dizon, 2017; Cohn et al., 2021; Hsu, Chen & Todd, 2021; Chung & Bong, 2022; Dizon, Tang & Yamamoto, 2022).

A single study used *an AI-powered spell-checker* (Noviyanti, 2020) during remote learning conditions.

Framework for using AI-powered tools in pronunciation training

The most valuable information the presented analysis brought was the description of the procedures and conditions in which the tools were used. For this paper, two primary conditions were recognised (1) using AI-powered tools for educational purposes and (2) using AI-powered tools for non-educational purposes.

(1) Using AI-powered tools for educational purposes

Several types of research focus must be recognised within this group of papers. One of the first groups was the **theoretical framework** of the application of AI in education. The first is *the Technology Acceptance Model (TAM),* which considers users' general attitudes towards the technology and their subsequent use of technology based on perceived usefulness and ease of use of the tool. The first was a unique study studying the perception of Ai-powered tools from the teachers' perspective carried out by Sumakul, Hamied and Sakyadi (2022). The study analysed teachers' attitudes to using AI based on the Technology Acceptance Model. The teachers used ElsaSpeak to train learners' pronunciation at home, with weekly in-class sessions reflecting the training. The teachers viewed AI technology as generally positive for their FL classroom, and they also identified increased motivation and benefited from improved technological and educational skills. The second study taking into consideration TAM and the only one using AI-powered tools for languages other than English, was carried out by Dizon, Tang and Yamamoto (2022). They analysed Alexa in the context of foreign learners of Japanese. The researchers collected data from communication with Alexa by the participants and then administered a questionnaire handed to the participants. The questionnaire was based on the Technology Acceptance Model and focused on usefulness, effectiveness, and satisfaction. The results revealed that using IPA varied from participant to participant. However, the system could recognise up to 83% of the commands. The participants typically did not interact with the API after the lack of recognition; however, the system can help develop learner autonomy. Annamalai et al. (2023) carried out a study arising from Self-Determination Theory as an exploratory case study involving chatbots in teaching English. Within the study, learners used different chatbots to improve their English proficiency. A questionnaire then collected the perceptions of learners. Within the overall competence, the pronunciation was developed by providing feedback on the mispronunciation of individual words by repetition.

The second category of studies investigated **speaking anxiety** as one of the critical factors of oral proficiency; Dizon (2017) used a mixed-methods case study designed to examine the experience of four EFL learners in Japan with Alexa after a 20-minute interaction consisting of two tasks – giving five set commands to Alexa and using Earplay Interaction to participate in a story (deciding the story continuation). The participants were allowed to provide commands a maximum of three times. The researcher analysed the interactions between participants and tools observed in the first task on the range of understood – wake word error (a specific word not detected) and understood – no response (by the participant) in the second task. The results have shown overall better results in the second task (90% comprehension). One of the possible explanations for why participants performed weakly in the first task (50% comprehension) was that Alexa supported only three languages at the time of the research (American and British English, German). Dizon (2017) used a semi-structured interview to collect participants' opinions on six major aspects of their experience. The participants reported that indirect feedback and the ability to converse improved participant's effectiveness in using the tool, but the lack of language support did not develop the effectiveness of learners. The overall conclusion of the study was that with the increasing quality of the tools used, the opportunity to use them in improving comprehensibility would grow. On the contrary, the study by Hsu, Chen and Todd (2021) reported that anxiety can be lowered by using Alexa.

Three years later, El Shazly (2020) carried out a case study examining the impact of AI on speaking anxiety and performance. The study was designed as an eight-week quasi-experimental pretest-posttest design. The speaking proficiency was measured in pretest and posttest by an enhanced version of the public IELTS speaking rubric with a special rubric measuring the time and type of interaction added by the author. After the pretest, participants were exposed to 8-week lessons with AI-driven applications and chatbots. The participants were evaluated on precision, coherence, and appropriateness of speaking when communicating with AI-powered tools. In addition, the participants received pronunciation training using the tools for twenty minutes in class and five hours a week at

home. However, El Shazly's (2020) study did not confirm that AI would decrease the speaking anxiety of FL learners, which the author possibly explains as a result of the formal setting of the study.

Another category of papers focused on **developing language proficiency in general**. Hsu, Chen and Todd (2021) used Alexa to develop L2 listening and speaking skills in an experiment with seven lessons planned. Pronunciation, in general, was developed in the process. The participants were tested by TOEIC pre- and posttests on listening and speaking. The experimental group received movie-watching/text-reading lessons with vocabulary instruction and Alexa sessions. After the treatment, statistically significant differences between the control and experimental group were identified. The participants with mispronunciations reported a lack of responses by Alexa. Zou et al. (2023) carried out a 5-week experiment using AI-powered applications in two groups of Chinese learners of English. However, the experimental group was expected and reminded to use social networking applications to carry out interactive tasks. This condition created a fundamental difference resulting in the conclusion that although AI cannot create a flexible enough environment for learning languages, practising them in social networks can support the informal development of communicative competence. The differences in results between both groups in the posttest included improvement in speaking fluency, accuracy and pronunciation.

Isbel, Crowther and Nishizawa (2023) used Duolingo English Test as a benchmark for **evaluating speaking performances in an academic environment**. Different listener groups (faculty members, administration staff, graduate, and undergraduate students) aimed to evaluate comprehensibility and acceptability in the graduate and undergraduate admission process. The study aimed to find the relationship between humans and AI automated evaluation of speaking skills, which should help estimate the future academic success of the candidates. The results revealed that undergraduate students tended to be the strictest judges. At the same time, other groups of judges were the most tolerant of their speaking performance. On the other hand, there was a strong correlation between human and digital evaluation, thus suggesting that the DET results could be extrapolated to academic environments.

Citrayasa (2019) conducted a phenomenological qualitative study in an informal setting with two junior high schools using an English learning mobile application Busuu and focused on identifying users' **"lived experiences"**.

(2) Non-educational purpose

A small sample of analysed papers benefited from testing pronunciation for other than educational purposes; however, the results could be directly applicable in pronunciation classrooms. The first is a study by Cohn et a. (2021), who analysed **vocal alignment** of speakers shadowing Amazon Alexa voice actors and human speakers pronouncing interjections in an emotionally expressive or neutral manner in two repletion blocks. The measurements showed longer duration, variation and pitch when shadowing the emotionally expressive interjections. However, the difference in pronouncing longer words was detected when shadowing the human speaker. Another study focusing on including prosody in chatbots and the consequent responses of their users was carried out by Zhu et al. (2022). The authors predicted that more emotional expressions would increase positive perceptions of chatbots. The authors designed two tasks based on interaction and perception to confirm the prediction. The first task was designed around four combinations of choosing expressive or baseline prosody and/or expressive or baseline vocabulary in combination and interaction of the first group of speakers with chatbots. In the second task, different participants listened to the conversation samples collected from the first part of the study and rated the interactions. The results from both parts revealed that while interactions benefit from emotional vocabulary and prosody, independent listeners find prosody more critical. The second one is the study by Khampusaen, Chanprasopchai and Lao-un (2023), who designed an AI-powered course within **English for Specific Purposes**. In this particular context, the authors developed a pronunciation training program for local homestay owners in Thailand. The study was carried out in two stages – the first stage involved the identification of learners' needs and developing classes with ten new vocabulary items in the application Line. The second stage was based on using the application as a learning instrument. During the third stage, the researchers collected the results. The posttest results indicated that all participants improved in pronunciation; however, the results varied according to the age and education of all participants.

Study limitations

Using AI, particularly in pronunciation training, is an underrepresented topic which should be addressed in the future, as its potential is limitless. The vast number of areas that benefit from using AI in pronunciation training (decreasing administrative load, overcoming psychological obstacles and creating new opportunities, improving AI-powered tools) inspires better pedagogical practices even now.

However, several limitations need to be identified and addressed in researching this particular topic. One of the issues is the need for more available full texts in academic databases and search engines, making the presented study more profound and allowing for a deeper understanding of the analysed issue.

Another limitation of the paper is the comparison of studies against other studies using AI in pronunciation training and not including studies dealing with similar issues more traditionally. This comparison could allow for more detailed identification of the benefits and drawbacks of using AI in pronunciation practice. However, it must be noted that the increased interest in the field will bring numerous opportunities to carry out other studies contributing to the area.

Discussion and recommendations

After presenting the theoretical background and analysis of selected research studies, answering research questions and recommendations for practice follow.

To answer the first research question based on the presented review, it can be concluded that AI-powered tools are present in pronunciation practice in relatively smaller quantities than for practising other layers of language (Moussali & Cardoso, 2020; Huang et al., 2023; Zhai & Wibowo, 2023). However, the relatively small sample of analysed papers (N=15) available for closer inspection showed strong dominance of some tools (learning apps with AI-powered modules including a chatbot, N = 9) on the one hand and tentative use of others (an AI-powered spell-checker in Noviyanti, 2020) on the other hand. The dominance of learning applications with integrated chatbots may be explained by their free accessibility and motivational character (Khampusaen, Chanprasopchai & Lao-un, 2021). The second largest group of papers (N=5) focused on the use of IPAs which are freely accessible in home appliances and support learning based on accuracy (Bogach et al., 2021; Rogerson-Revell, 2021; Foster & Stuart-Smith, 2023) in informal settings with the added benefit of pronunciation training while carrying out other tasks.

Regarding the second research question, the presented analysis results indicate that pronunciation is primarily perceived as a part of overall speaking proficiency and is developed in such parameters (fluency, rhythm; Zou et al., 2023). In such cases, intelligibility and comprehensibility (Levis, 2005, 2020; Derwing & Munro, 1997, 2009; Field, 2005) with human and AI-powered tools are being promoted. However, some studies focus on developing particular aspects of pronunciation, such as Lingua Franca core and non-core features (Jenkins, 2006;), which promote the overall intelligibility of non-native speakers of English or specific pronunciation issues arising from mother tongue interference.

Regarding the results achieved after the pedagogical treatment, the overall quality of language proficiency has increased on average, and learners tend to respond positively to the use of AI-powered tools and can relate to them reasonably.

The overall results of the pedagogical implementation of AI-powered tools show gains in other spheres of communication, such as decreasing speaking anxiety (Dizon, 2017; Hou, Chen & Todd, 2021), heightened motivation (Moussali & Cardoso, 2020) or use of compensation strategies when communicating with AI-powered tools has failed (Underwood, 2017). However, such conclusions are only limited to a particular group of learners, as each learner is an individual and a universal tool for solving all challenges in overcoming obstacles in language learning is yet to be found.

The results indicate a relatively positive development, and to ensure more teachers embrace and benefit from AI in their pedagogical practice, future research in this area should focus on the following points: (1) preparing teachers to use AI-powered tools in their pronunciation classroom (strengthening their pedagogical and technical competencies and level of information), (2) making more AI-powered tools available and accessible to the teachers, and (3) sharing more examples of good practice in the teacher community for inspiration and encouragement.

Conclusions

The undeniable impact of technology and AI-powered tools in language pedagogy has brought promising results in developing various aspects of language proficiency and increasing learners' motivation to learn, increasing attention to the meaningfulness of communication, learning, and understanding how technology benefits education. AI and AI-powered tools, as one of the prolific spheres of learning technology, have been slowly embraced in the educational community when learning vocabulary or grammar. Now, the attention is being focused on pronunciation improvement. This observation is evidenced by the volume of scientific reports on testing the quality of various systems; however, little systematic research has been done into putting AI-powered tools into pedagogical practice. Pronunciation is a critical component of speaking proficiency, and good pronunciation can overcome many problems in communication, be it a lack of comprehension or high levels of anxiety. The available AI-powered tools integrated into chatbots, language learning applications and virtual assistants allow learners to practice when the conditions allow them to focus on learning. The use of the systems for pronunciation training appears to be relatively limited compared to developing grammar or vocabulary; therefore, promoting their use should be a priority in different spheres of language learning.

References

AlDakhil, M., & AlFadda, H. (2022). EFL Learners' Perceptions Regarding the Use of Busuu Application in Language Learning: Evaluating the Technology Acceptance Model (TAM). *English Language Teaching*, 15(1), 1-15.

Alfuhaid, S. R. (2021). The Utilisation of Duolingo to Enhance the Speaking Proficiency of EFL Secondary School Students in Saudi Arabia. *English Language Teaching*, 14(11), 9-15.

Almelhes, S. A. (2023). A Review of Artificial Intelligence Adoption in Second-Language Learning. *Theory and Practice in Language Studies*, *13*(5), 1259-1269.

An, X., Chai, C. S., Li, Y., Zhou, Y., Shen, X., Zheng, C., & Chen, M. (2022). Modeling English teachers' behavioral intention to use artificial intelligence in middle schools. *Education and Information Technologies*, 1-22.

Annamalai, N., Eltahir, M. E., Zyoud, S. H., Soundrarajan, D., Zakarneh, B., & Al Salhi, N. R. (2023). Exploring English language learning via Chabot: A case study from a self determination theory perspective. *Computers and Education: Artificial Intelligence*, 100148.

Ayedoun, E., Hayashi, Y., & Seta, K. (2019). Adding communicative and affective strategies to an embodied conversational agent to enhance second language learners' willingness to communicate. *International Journal of Artificial Intelligence in Education*, *29*, 29-57.

Bajorek, J. P. (2017). L2 pronunciation in CALL: The unrealized potential of Rosetta stone, Duolingo, Babbel, and mango languages. *Issues and Trends in Educational Technology*, *5*(1), 24-51.

Bashori, M., van Hout, R., Strik, H., & Cucchiarini, C. (2022). 'Look, I can speak correctly': learning vocabulary and pronunciation through websites equipped with automatic speech recognition technology. *Computer Assisted Language Learning*, 1-29.

Benzies, Y. J. C. (2017). Contributions of new technologies to the teaching of English pronunciation. *Language Value*, 9(1).

Bogach, N., Boitsova, E., Chernonog, S., Lamtev, A., Lesnichaya, M., Lezhenin, I., ... & Blake, J. (2021). Speech processing for language learning: A practical approach to computer-assisted pronunciation teaching. *Electronics*, 10(3), 235.

Brick, B., & Cervi-Wilson, T. (2019). Enhancing learners' professional competence via Duolingo classroom. Professional competencies in language learning and teaching, Research-publishing. net, 19-29.

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: a narrative overview. *Procedia Computer Science*, *136*, 16-24.

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *Ieee Access*, *8*, 75264-75278.

Cheng, Z., & Wang, Z. (2022). Automatic Scoring of Spoken Language Based on Basic Deep Learning. *Scientific Programming*, 2022.

Chung, B., & Bong, H. K. M. (2022). A study on the intelligibility of Korean-Accented English: Possibilities of implementing AI applications in English education. *Journal of Asia TEFL*, 19(1), 197.

Chuyen, N. T. H., Linh, H. T., & Phuc, N. T. H. (2021). Enhancing English Pronunciation for High School Students Through Duolingo Application. *International Journal of All Research Writings*, 3(1), 46-54.

Citrayasa, V. (2019). Junior high school students' lived experiences of learning English using Busuu. *Indonesian EFL Journal*, 5(2), 85-92. doi: 10.25134/ieflj.v5i2.1900.

Dai, Y. (2022). An Automatic Pronunciation Error Detection and Correction Mechanism in English Teaching Based on an Improved Random Forest Model. *Journal of Electrical and Computer Engineering*, *2022*.

Cohn, M., Predeck, K., Sarian, M., & Zellou, G. (2021). Prosodic alignment toward emotionally expressive speech: Comparing human and Alexa model talkers. *Speech Communication*, 135, 66-75.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340. https://doi.org/10.2307/249008

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, 19(1), 1–16.

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42, 276–490.

Dizon, G. (2017). Using intelligent personal assistants for second language learning: A case study of Alexa. *Tesol Journal*, 8(4), 811-830.

Dizon, G. (2021). Affordances and constraints of intelligent personal assistants for second-language learning. *RELC Journal*, 00336882211020548.

Dizon, G., Tang, D., & Yamamoto, Y. (2022). A case study of using Alexa for out-of-class, self-directed Japanese language learning. *Computers and Education: Artificial Intelligence*, 3, 100088.

El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, 38(3), e12667

Foster, M. E., & Stuart-Smith, J. (2023). Social robotics meets sociolinguistics: Investigating accent bias and social context in HRI.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL quarterly,* 39(3), 399-423.

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2022). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *BioRxiv*, 2022-12.

Hazar, E. (2022). Learning a Brand-New Language through Duolingo: A Case Study of a Gifted Student. *African Educational Research Journal*, 10(4), 447-453.

Hockly, N. (2023). Artificial Intelligence in English Language Teaching: The Good, the Bad and the Ugly. *RELC Journal*, 00336882231168504.

Hsu, H. L., Chen, H. H. J., & Todd, A. G. (2021). Investigating the impact of the Amazon Alexa on the development of L2 listening and speaking skills. *Interactive Learning Environments*, 1-14.

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. Journal of *Computer Assisted Learning*, 38(1), 237–257. https://doi.org/10.1111/jcal.12610

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, Research Issues and Applications of Artificial Intelligence in Language Education. *Educational Technology & Society*, *26*(1), 112-131. https://doi.org/10.30191/ETS.202301_26(1).0009

Iphofen, R., & Kritikos, M. (2021). Regulating artificial intelligence and robotics: ethics by design in a digital society. *Contemporary Social Science*, *16*(2), 170-184.

Isbell, D. R., Crowther, D., & Nishizawa, H. (2023). Speaking performances, stakeholder perceptions, and test scores: Extrapolating from the Duolingo English test to the university. *Language Testing*, 02655322231165984.

Jenkins, J. (2006). Current perspectives on teaching world Englishes and English as a lingua franca. *TESOL quarterly*, 40(1), 157-181.

Jiao, F., Song, J., Zhao, X., Zhao, P., & Wang, R. (2021). A spoken English teaching system based on speech recognition and machine learning. *International Journal of Emerging Technologies in Learning (iJET),* 16(14), 68-82.

Khalil, M., & Er, E. (2023). Will ChatGPT get you caught? Rethinking of plagiarism detection. arXiv preprint arXiv:2302.04335.

Khampusaen, D., Chanprasopchai, T., & Lao-un, J. (2023). Empowering Thai Community-based Tourism Operators: Enhancing English Pronunciation Abilities with AI-based Lessons. *Journal of Mekong Societies*, 19(1), 132-159.

Lee, K. A., & Lim, S. B. (2023). Designing a Leveled Conversational Teachable Agent for English Language Learners. *Applied Sciences*, *13*(11), 6541.

Lee, S., Noh, H., Lee, J., Lee, K., & Lee, G. G. (2010). Cognitive effects of robot-assisted language learning on oral skills. In *Second language studies: Acquisition, learning, education and technology* (pp. 1–4).

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL quarterly*, 39(3), 369-377.

Levis, J. (2020). Revisiting the intelligibility and nativeness principles. *Journal of Second Language Pronunciation*, 6(3), 310-328.

Li, Y. (2022). Teaching mode of oral English in the age of artificial intelligence. *Frontiers in Psychology*, *13*, 953482.

Lim, W. M., Kumar, S., Verma, S., & Chaturvedi, R. (2022). Alexa, what do we know about conversational commerce? Insights from a systematic literature review. *Psychology & Marketing*, *39*(6), 1129-1155.

McCrocklin, S. (2019). ASR-based dictation practice for second language pronunciation improvement. *Journal of Second Language Pronunciation*, 5(1), 98-118.

Moussalli, S., & Cardoso, W. (2020). Intelligent personal assistants: can they understand and be understood by accented L2 learners?. *Computer Assisted Language Learning*, *33*(8), 865-890.

Neri, A., Cucchiarini, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer assisted language learning*, 15(5), 441-467.

Ng, D. T. K., Leung, J. K. L., Su, J., Ng, R. C. W., & Chu, S. K. W. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Educational technology research and development*, *71*(1), 137-161.

Noviyanti, S. D. (2020). Artificial intelligence (AI)-Based pronunciation checker: an alternative for independent learning in pandemic situation. *Journal of English Language Teaching in Foreign Language Context*, 5(2), 162.

O'Brien, M., Derwing, T., Cucchiarini, C., Hardison, D., Mixxdorff, H., Thomosn, R., Strik, H., Levis, J., Munro, M., Foote, J., & Muller Levis, G. (2018). Directions for the future of technology in pronunciation research and teaching. *Journal of Second Language Pronunciation*, 4(2), 182-207.

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.

Portnoff, L., Gustafson, E., Rollinson, J., & Bicknell, K. (2021). Methods for Language Learning Assessment at Scale: Duolingo Case Study. International Educational Data Mining Society.

Rogerson-Revell, P. M. (2021). Computer-assisted pronunciation training (CAPT): Current issues and future directions. *Relc Journal*, *52*(1), 189-205.

Shufang, Z. (2021). Design of an Automatic English Pronunciation Error Correction System Based on Radio Magnetic Pronunciation Recording Devices. *Journal of Sensors*, 2021, 1-12.

Sidgi, L. F. S., & Shaari, A. J. (2017). The usefulness of automatic speech recognition (ASR) eyespeak software in improving Iraqi EFL students' pronunciation. *Advances in language and literary studies*, 8(1), 221-226.

Sumakul, D. T., Hamied, F. A., & Sukyadi, D. (2022). Artificial intelligence in EFL classrooms: Friend or foe? *LEARN Journal: Language Education and Acquisition Research Network, 15*(1), 232-256.

Underwood, J. (2017). Exploring AI language assistants with primary EFL students. *CALL in a climate of change: adapting to turbulent global conditions–short papers from EUROCALL*, *2017*, 317-321.

Vančová, H. (2019). Current issues in pronunciation teaching to non-native learners of English. *Journal of Language and Cultural Education*, 7(2), 140-155.

Vančová, H. (2020). *Pronunciation Practices in EFL Teaching and Learning*. University of Hradec Králové, Gaudeamus Publishing House.

Vančová, H. (2021). *Teaching English pronunciation using technology*. Kirsch-Verlag.

Yoshida, M. T. (2018). Choosing technology tools to meet pronunciation teaching and learning goals. *Catesol Journal*, *30*(1), 195-212.

Wang, Y. H., & Young, S. C. (2015). Effectiveness of feedback for enhancing English pronunciation in an ASR-based CALL system. *Journal of Computer Assisted Learning*, 31(6), 493-504.

Zhai, C. & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence*, 100134.

Zhu, Q., Chau, A., Cohn, M., Liang, K. H., Wang, H. C., Zellou, G., & Yu, Z. (2022). Effects of Emotional Expressiveness on Voice Chatbot Interactions. *Proceedings of the 4th Conference on Conversational User Interfaces*, 1-11.

Zou, B.; Guan, X.; Shao, Y.; Chen, P. (2023). Supporting Speaking Practice by Social Network-Based Interaction in Artificial Intelligence (AI)-Assisted Language Learning. *Sustainability*, 15, 2872. https://doi.org/10.3390/su15042872

Appendix 1: Overview of the analysed studies (ordered chronologically)

Study	AI-powered tool	Pronunciation focus/aim	Participants/country	Results
Dizon, G. (2017)	Alexa	speaking anxiety	4 EFL learners, Japan	the tool perceived as useful; anxiety not lowered
Citrayasa, V. (2019)	Busuu	participants' lived experiences	2 junior high school students, Indonesia	improvement of pronunciation
Noviyanti, S. D. (2020)	A spell checker	silent letters	30 students, Indonesia	improvement of pronunciation
Chuyen, N. T. H., Linh, H. T., & Phuc, N. T. H. (2021)	Duolingo	consonants, verb endings, word stress	41 students in grade 10 (16 years old), Thailand	improved pronunciation and confidence
Cohn, M. et al. (2021)	Alexa	vocal alignment	66 English native speakers (mean age = 20.64), USA	increased alignment with emotive expressions
El Shazly, R. (2021)	Mondly, Duolingo, FluentU, Glossika	speaking anxiety	48 EFL learners, Egypt (mean age = 18.9)	precision, coherence, and appropriateness
Hsu, H. L., Chen, H. H. J., & Todd, A. G. (2021)	Alexa	speaking skills, perceptions of IPAs	26 learners, Taiwan (mean age = 19)	a significant difference in speaking skills, reduced anxiety
Chung, B. & Bong, H. K. M. (2022)	Google Assistant	intelligibility in Lingua Franca Communication	30 NS college students from different countries, (mean age = 22)	AI-apps effective in KoE recognition
Dizon, G., Tang, D., & Yamamoto, Y. (2022)	Alexa	the AI-powered tool used in non-English context (Japan)	6 SL learners of Japanese	positive attitude, abandonment of the task
Sumakul, D. T., Hamied, F. A., & Sukyadi, D. (2022)	ElsaSpeak	teachers' perceptions of AI, TAM	4 EFL teachers, Indonesia	positive attitudes detected
Zhu, Q. et al. (2022)	chatbots	the degree of emotional expressiveness via prosody	50 and 56 participants (18 – 32 years old)	increased likeability with prosody
Annamalai, N., et al. (2023)	chatbots (Duolingo, Mondly & Andy),	autonomy; self-determination theory; competence and relatedness (psychological needs)	25 undergraduate students, Malaysia (26 to 35 years old)	support of psychological needs, lack of emotion, inaccuracy of information
Isbell, D. R., Crowther, D., & Nishizawa, H. (2023)	Duolingo English Test (DET)	a comparison of digital and human evaluation	100 DET test-takers, various linguistic backgrounds and 204 evaluators	correlation between comprehensibility and acceptability
Khampusaen, D., Chanprasopchai, T., & Lao-un, J. (2023)	AI-based lessons in application Line	improving pronunciation for tourism (ESP)	15 participants involved in tourism (16 to 49 years old)	effectivity and suitability of the application
Zou, B. P. et al. (2023)	Chatbots and WeChat (network-based interaction)	pronunciation, fluency, oral rhythm	70 students from Chinese universities and majors	positive attitudes, the app effective