



Integrating AI-Based Pronunciation App ELSA Speak in English Lessons: A Case Study at SMP Negeri 4 Batanghari Nuban

Erry Raihan¹ Mai Zuniati¹, Umar Al Faruq Al Hasyim¹, Suhono¹

¹ Universitas Ma'arif Lampung, Indonesia

Corresponding Author erryraihaan@icloud.com*

ABSTRACT

The development of digital technology has significantly influenced the way English is taught and learned, particularly in pronunciation practice. One of the applications that can assist students in improving their pronunciation is ELSA Speak, which provides automatic feedback on spoken English. This study aimed to examine the implementation of ELSA Speak in English pronunciation learning and to explore students' responses to its use in the classroom. This study used a qualitative case study design conducted at SMP Negeri 4 Batanghari Nuban. The participants consisted of one English teacher and 31 eighth-grade students, with five students selected for in-depth interviews. The data were obtained through interviews and analyzed using descriptive qualitative analysis. The results showed that the application supported students in practicing pronunciation more regularly and helped them recognize and correct their pronunciation errors through instant feedback. In addition, students responded positively because the application made the learning process more interactive and engaging. However, several obstacles were also identified, such as unstable internet connections and students' difficulties in pronouncing certain English sounds. The findings also indicate that the teacher plays an important role in guiding students in using the application effectively during the learning process. The teacher's guidance helps students understand how to use the feedback provided by the application to improve their pronunciation. Therefore, the integration of technology-based applications such as ELSA Speak can be considered as an alternative tool to support pronunciation learning in English classrooms. This study also suggests that the use of digital pronunciation applications can enhance students' engagement and provide additional opportunities for independent pronunciation practice.

Keywords: *ELSA Speak, Pronunciation Practice, Students' Perception, English Learning.*

ARTICLE INFO

Article history:

Received

March 20, 2026

Revised

May 13, 2026

Accepted

June 25, 2026

Journal Homepage

<https://ojs.staialfurqan.ac.id/IJoASER/>

This is an open access article under the CC BY SA license

<https://creativecommons.org/licenses/by-sa/4.0/>

PENDAHULUAN

English is widely recognized as an international language that plays a significant role in global communication, education, science, and technology (Romayanti & Zuniati, 2021). In the process of learning English, students are required to master four main language skills, namely listening, speaking, reading, and writing. Among these skills, speaking is often considered one of the most challenging abilities because it requires not only knowledge of vocabulary and grammar but also clear and accurate pronunciation to convey meaning effectively. According to Gilakjani (2021), pronunciation is a crucial

component of speaking ability because it directly influences how well a speaker's message can be understood by listeners.

Pronunciation plays an important role in helping learners communicate their ideas clearly. Even when students possess sufficient vocabulary and grammatical knowledge, inaccurate pronunciation may lead to misunderstandings in communication. According to Derwing and Munro (2020), pronunciation accuracy significantly affects the intelligibility of spoken language in communication. Similarly, Levis (2020) explains that pronunciation instruction is essential in English language learning because it helps learners develop clearer and more understandable speech.

For many learners of English as a Foreign Language (EFL), mastering English pronunciation is often challenging due to differences between the sound system of English and their first language. These differences may cause learners to mispronounce certain vowels, consonants, or stress patterns. According to Sharma (2021), pronunciation difficulties often occur because learners tend to transfer pronunciation patterns from their native language when speaking English. As a result, students may experience difficulties in expressing their ideas clearly during spoken communication.

In many classroom contexts, students also have limited opportunities to practice pronunciation regularly. Teachers may face challenges in providing individual feedback to each student, particularly in classes with a large number of learners. According to Richards (2021), the lack of individualized feedback in language classrooms can slow down students' progress in improving their speaking skills. Therefore, integrating technology into language learning has become increasingly important to support students' learning needs.

With the rapid development of educational technology, various digital tools have been introduced to support language learning. According to Hidayatullah, Hasyim, and Tobing (2024), the importance of continuously updating and integrating new technologies is emphasized in order to improve the overall learning experience. According to Nguyen (2021), technology-based learning tools can provide additional opportunities for students to practice language skills independently outside the classroom. One example of such technology is the ELSA Speak application (English Learning Speech Assistant), which is designed to assist learners in improving their English pronunciation through artificial intelligence.

ELSA Speak is an AI-based mobile application that uses speech recognition technology to analyze learners' pronunciation and provide instant feedback on their speaking performance. According to Tran and Vu (2024), AI-based pronunciation applications can help learners improve pronunciation accuracy because they provide immediate feedback that allows learners to identify and correct their pronunciation errors.

A number of earlier studies have explored the use of AI-based pronunciation applications, particularly ELSA Speak, within English language learning contexts. For instance, Kholis (2021) found that the use of the ELSA Speak application helped students develop more accurate pronunciation, as the immediate feedback feature guided them in refining their word production. In a similar context, Indrayani et al. (2025) reported that the integration of ELSA Speak in classroom practice was perceived to support students' speaking performance, particularly in terms of pronunciation and fluency. Moreover, Yosintha and Rekha (2022) examined learners' perceptions of using ELSA Speak in online pronunciation classes and found that students generally responded positively, as the application provided broader opportunities for practice and encouraged self-directed improvement. Furthermore, Aulia and Santosa (2025), through a systematic review, highlighted that ELSA Speak was found to play an important role in supporting learners' pronunciation development, increasing their motivation, and fostering independent learning across various EFL settings.

However, previous studies have primarily focused on measuring learning outcomes, with limited attention given to the actual classroom implementation of AI-based pronunciation tools. In addition, there is still a lack of in-depth exploration of students' learning experiences and challenges in face-to-face classroom contexts, particularly in Indonesian junior high schools.

Therefore, this study presents a distinct contribution by examining the integration of the ELSA Speak application in English instruction through a qualitative case study design. In contrast to earlier research that largely focuses on learning outcomes, this study aims to explore the process of classroom implementation, students' responses, and their learning experiences while using the application. Furthermore, it emphasizes the teacher's role in guiding the use of ELSA Speak and investigates the challenges that emerge during its application in an actual classroom setting. Conducted at SMP Negeri 4 Batanghari Nuban, this research offers in-depth and contextual insights into the use of AI-based pronunciation tools in a junior high school EFL setting in Indonesia, an area that is still limited in the existing literature.

Based on this background, this study is guided by three main research questions. It examines how the ELSA Speak application is implemented in English pronunciation learning, explores students' responses toward its use, and investigates the challenges encountered during its implementation in classroom practice.

METHOD

This study employed a qualitative approach using a case study design to explore the implementation of the ELSA Speak application in supporting students' English pronunciation learning. A qualitative method was chosen as the study aims to gain a deeper understanding of the learning process, students' experiences, and their responses to the use of technology in language learning. As stated by Creswell, John W. (2014), qualitative research is suitable for exploring and interpreting the meanings that individuals or groups assign to social or human issues. Furthermore, Yin, Robert K. (2018) emphasizes that a case study design is appropriate for examining a contemporary phenomenon within its real-life context.

The research took place at SMP Negeri 4 Batanghari Nuban, Lampung, during the 2025/2026 academic year. This site was selected because the use of technology-based media, particularly for pronunciation practice, was still limited in English teaching. The participants consisted of one English teacher and 31 eighth-grade students who were involved in learning activities using the ELSA Speak application. They were selected through purposive sampling, as they were directly engaged in the instructional process where the application was implemented. Sugiyono (2017) explains that purposive sampling is used to select participants based on specific criteria aligned with the research objectives.

Data were collected through classroom observation, interviews, and documentation to obtain comprehensive information about the use of ELSA Speak in pronunciation learning. Observations were conducted during English lessons to identify how the teacher incorporated the application and to examine students' participation and engagement throughout the learning process. Interviews with the teacher and selected students were carried out to gain deeper insights into their experiences, perceptions, and challenges in using the application. In addition, supporting documents such as lesson plans, instructional materials, and screenshots of classroom activities were gathered to complement the data.

The data in this study consisted of both primary and secondary sources. Primary data were obtained from direct observations and interviews with the English teacher and students, providing detailed information about the implementation of the ELSA Speak application and students' responses. Meanwhile, secondary data were derived from

documentation, including lesson plans, learning materials, and screenshots of classroom activities, which served to support and validate the primary data.

The collected data were analyzed using qualitative data analysis procedures, including data reduction, data display, and conclusion drawing. According to Miles, Matthew B. and Huberman, A. Michael (2014), qualitative data analysis involves these three main steps. To ensure the credibility of the findings, data triangulation was applied by comparing information obtained from observations, interviews, and documentation. Denzin, Norman K. (2012) states that triangulation is used to enhance the validity and reliability of qualitative research results.

RESULTS AND DISCUSSION

Implementation of ELSA Speak in English Pronunciation Learning

The interview with the English teacher revealed that the use of the ELSA Speak application in the classroom began with an explanation of the purpose of integrating the application into the learning process. The teacher informed the students that the application was intended to help them practice English pronunciation more regularly and to provide immediate feedback on their speaking performance. The integration of technology in language learning is considered beneficial because it can increase students' engagement and motivation during the learning process. According to Jeremy Harmer (2021), the use of technological tools in language learning can make learning activities more interactive and support students in developing their language skills. After explaining the purpose of the application, the teacher demonstrated how to download, install, and use the ELSA Speak application so that students could clearly understand how to operate it during the lesson.



Figure 1. Students practicing pronunciation using the ELSA Speak application.

During the teaching and learning process, the teacher first introduced the pronunciation material that would be practiced in the session. The teacher also provided examples of correct pronunciation of several words and sentences before asking the students to practice using the application. The students then listened to the pronunciation model provided in ELSA Speak and repeated the words or sentences according to the instructions. This activity helped students compare their pronunciation with the correct model and improve their pronunciation accuracy. Pronunciation practice is important because it helps learners produce English sounds more clearly and effectively in communication. As stated by Abbas Pourhosein Gilakjani (2021), continuous pronunciation practice supported by appropriate learning tools can help learners improve their speaking performance. While the students were practicing, the teacher monitored their activities by walking around the classroom and assisting students who experienced

difficulties in using the application or pronouncing certain words. After the practice session, the teacher discussed several common pronunciation mistakes and provided additional explanations to help students improve their pronunciation skills.

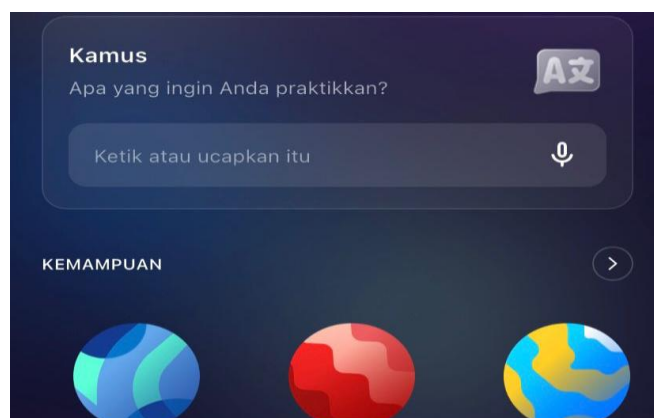


Figure 2. The Implementation of ELSA Speak in the Classroom

The findings indicate that the implementation of the ELSA Speak application follows a structured and guided learning process in which the teacher plays a central role in facilitating technology integration. Although the application provides automated feedback, the teacher's guidance remains essential in helping students interpret the feedback and apply it effectively. This suggests that technology functions as a complementary tool rather than a replacement for the teacher in the learning process.

The ELSA Speak application offers various features that assist students in improving their English pronunciation. Through the application, students are able to listen to pronunciation examples and repeat the words or sentences by using the microphone feature. After the pronunciation practice, the application automatically evaluates the students' speech and provides immediate feedback. This feature enables students to recognize their pronunciation errors and motivates them to practice repeatedly until they achieve better pronunciation. By engaging in continuous practice, students can gradually enhance their pronunciation accuracy and develop greater confidence when speaking English. This finding also reflects the role of AI-based applications in promoting self-directed learning, as students are given opportunities to practice independently while receiving instant corrective feedback.

Based on the interview with the English teacher, the implementation of the ELSA Speak application in the classroom was carried out through several stages during the learning process. These stages were designed to help students understand how to use the application and practice English pronunciation effectively.

Students' Responses toward the Use of ELSA Speak

The results of the interviews with five students showed that most of them gave positive responses toward the use of the ELSA Speak application in English pronunciation learning. The students stated that the application made the learning process more interesting and different from the usual classroom activities. They also mentioned that using the application allowed them to practice pronunciation more frequently and helped them become more engaged in the learning process. This result supports the findings of Yosintha and Rekha (2022), who found that students responded positively to the use of ELSA Speak, as it offered greater opportunities for practice and promoted independent learning. In a similar vein, Aulia and Santosa (2025) reported that the use of ELSA Speak in learning activities was able to foster students' motivation and encourage greater engagement in practicing pronunciation.

Some students explained that they initially felt shy when they had to speak into the microphone while using the application. However, after practicing several times, they gradually became more comfortable and confident. The students also reported that listening to the pronunciation model and repeating the words helped them understand how to pronounce English words more accurately. As a result, they became more aware of their pronunciation mistakes and were motivated to practice more regularly. This finding aligns with Kholis (2021), who indicated that ELSA Speak assists learners in developing more accurate pronunciation by providing immediate feedback. Furthermore, Indrayani et al. (2025) noted that the use of ELSA Speak contributes to students' speaking performance, especially in enhancing pronunciation and fluency. This finding further reinforces earlier studies suggesting that AI-based applications encourage autonomous learning. Liakin et al. (2015) explained that mobile-assisted pronunciation learning allows students to engage in independent practice while also enhancing their involvement in the language learning process.

These findings also highlight the importance of affective factors in language learning. The increase in students' confidence and the reduction of speaking anxiety suggest that the use of AI-based applications creates a less intimidating learning environment. Compared to traditional classroom speaking activities, practicing with the application allows students to learn at their own pace, which contributes to a more positive learning experience.

Overall, the students' responses indicate that the use of the ELSA Speak application has a positive influence on their pronunciation learning. The application not only helps students improve their pronunciation skills but also increases their motivation, confidence, and engagement in practicing English speaking.

Table 1. Students' Responses toward the Use of ELSA Speak

Student	Students' Responses
S1	The student felt happy because the learning process was different and more interesting when using the ELSA Speak application.
S2	The student stated that the application helped in practicing English pronunciation more frequently.
S3	The student mentioned that the application helped distinguish the correct pronunciation of English words.
S4	The student enjoyed using the application although initially felt shy when speaking into the microphone.
S5	The student enjoyed using the application although initially felt shy when speaking into the microphone.

Benefits of Using ELSA Speak for Pronunciation Practice

The results of this study indicate that the use of the ELSA Speak application provides several advantages in learning English pronunciation. One of the most important features of the application is the immediate feedback given after students pronounce words or sentences. Through this feedback, students are able to identify their pronunciation errors and make improvements through repeated practice. This immediate feedback mechanism is particularly important because it allows learners to correct errors in real time, which is often difficult to achieve in traditional classroom settings.

The integration of technology in language learning can also enhance students' motivation and participation in classroom activities. According to Jack C. Richards (2021), digital learning tools can support language development by providing opportunities for learners to practice skills interactively and repeatedly. In this case, the ELSA Speak application allows students to listen to pronunciation models and compare their pronunciation with the correct one.

In addition, the use of the application encourages students to be more actively involved in pronunciation practice. Rather than only listening to explanations from the

teacher, students can directly practice speaking using the application. This interactive activity helps students gradually improve their pronunciation accuracy and build greater confidence when speaking English. Furthermore, Sa'diah et al. (2023), explain that when students are interested in what they do, they tend to enjoy the learning process and understand the material better. This suggests that the use of engaging digital tools such as ELSA Speak can create a more student-centered learning environment.

Tabel 2. Benefits of Using ELSA Speak for Pronunciation Practice

Benefit	Description
Instant feedback	The application provides immediate feedback after students pronounce words or sentences.
More pronunciation practice	Students can repeat the pronunciation practice several times until they achieve better results.
Increased motivation	The use of the application makes the learning process more interesting and interactive.
Pronunciation awareness	Students become more aware of their pronunciation mistakes and try to improve them.

Challenges in Using ELSA Speak in Pronunciation Learning

Even though the ELSA Speak application offers several advantages for pronunciation learning, a number of challenges were also encountered during its use in the classroom. Based on the interview data obtained from the English teacher and the students, one of the most common difficulties was related to the internet connection. In certain situations, an unstable network caused the application to operate more slowly, which sometimes disrupted the learning process. When the connection was not stable, the application required more time to process students' pronunciation and provide feedback, which occasionally interrupted the flow of classroom activities.

Another challenge was associated with the technical limitations of the devices used by students. Some students experienced problems related to microphone sensitivity and device performance while using the application. These issues sometimes affected the system's ability to recognize students' voices clearly, which resulted in inaccurate pronunciation detection. As a consequence, students sometimes had to repeat their pronunciation several times before the application could provide proper feedback.

Furthermore, several students reported that they found it difficult to understand the instructions provided in the application because most of the explanations were presented in English. For students who were still developing their English proficiency, this situation occasionally created confusion when they tried to follow the learning tasks in the application. Therefore, the teacher needed to provide additional guidance and clarification to ensure that students understood how to complete each activity.

Another difficulty was related to the students' pronunciation ability itself. Some English sounds were still challenging for students to pronounce correctly even after multiple attempts. The differences between English pronunciation patterns and the students' first language sometimes influenced the way they produced certain sounds. Because of this, students needed more time and repeated practice in order to improve their pronunciation accuracy.

These challenges indicate that the effectiveness of AI-based learning tools is influenced not only by the features of the application but also by external factors such as technological infrastructure, students' language proficiency, and teacher support. Therefore, successful implementation requires a combination of appropriate technology, adequate facilities, and active teacher involvement.

To overcome these challenges, the teacher provided continuous support throughout the learning process. The teacher gave additional explanations, repeated demonstrations, and guided students whenever they faced difficulties in using the

application. In addition, students were encouraged to practice regularly so that they could gradually become more confident and familiar with the pronunciation activities provided in the ELSA Speak application.

Table 3. Challenges in Using ELSA Speak

Challenge	Explanation
Internet connection	Unstable internet connection sometimes slowed down the application during the learning process.
Device limitations	Some students experienced technical problems related to the performance of their devices.
Microphone issues	The microphone on some devices could not capture students' voices clearly.
Pronunciation difficulties	Some students still found it difficult to pronounce certain English sounds correctly.

Conclusion

In conclusion, the findings of this study reveal that the integration of the ELSA Speak application in English pronunciation learning provides significant benefits for students. The application offers opportunities for repeated pronunciation practice and delivers immediate feedback, which helps students identify and correct their pronunciation errors more effectively. As a result, students gradually improve their pronunciation accuracy and develop greater confidence in speaking English.

In addition, the use of the ELSA Speak application contributes to increasing students' motivation and engagement in the learning process. The interactive features of the application create a more interesting and student-centered learning environment, allowing learners to actively participate in pronunciation practice. Furthermore, the findings indicate that AI-based applications can support autonomous learning by enabling students to practice independently beyond classroom activities.

These findings are consistent with previous studies which highlight the role of artificial intelligence in language learning. According to Tran and Vu (2024), the ELSA Speak application allows learners to practice pronunciation repeatedly while receiving instant feedback that can help improve pronunciation accuracy. Similarly, Nguyen (2021) states that technology-based pronunciation applications can enhance students' confidence and provide more effective opportunities for pronunciation practice.

However, this study also highlights that the successful implementation of technology-based learning tools is influenced by several factors, including internet connectivity, device performance, students' language proficiency, and teacher support. The teacher plays a crucial role in guiding students to effectively use the application and interpret the feedback provided. Therefore, technology should be viewed as a complementary tool that enhances, rather than replaces, the role of the teacher in the learning process. This study has several limitations that should be acknowledged. First, the research was conducted in a single school context with a limited number of participants, which may affect the generalizability of the findings. Second, the study employed a qualitative case study design without incorporating quantitative measurements to assess students' improvement in pronunciation accuracy. As a result, the findings mainly reflect students' experiences and perceptions rather than measurable learning outcomes.

In addition, the duration of the implementation was relatively limited, which may not fully capture the long-term impact of using the ELSA Speak application in pronunciation learning. Therefore, future studies are recommended to involve larger samples, longer implementation periods, and mixed-method approaches in order to obtain more comprehensive and generalizable findings.

REFERENCES

- Aulia, L., & Santosa, M. H. (2025). AI in language education: Systematic review of ELSA Speak's impact on English. *Jurnal Pendidikan Bahasa Inggris Indonesia*, 13(1), 1–10.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative and mixed methods approaches* (Vol. 54). United States of America: Sage Publications.
- Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research*, 6(2), 80–88.
- Derwing, T. M., & Munro, M. J. (2020). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.
- Dewi, R. (2021). Speaking students' worksheet based on e-book. *Journal of English Development*, 1(2), 35–44.
- Gilakjani, A. P. (2021). The significance of pronunciation in English language teaching and learning. *International Journal of Research in English Education*, 6(2), 1–15.
- Godwin-Jones, R. (2021). Emerging technologies: AI and language learning. *Language Learning & Technology*, 25(3), 5–19.
- Hidayatullah, R., Hasyim, U. A. A., & Tobing, D. D. (2024). Development of "Eng-Lab" application based on needs analysis to improve professional English in Metro City. *Journal of English Development*, 4(1), 182–192.
- Indrayani, L., Pratiwi, N. P. A., & Arthana, I. K. R. (2025). ELSA Speak as a speaking partner: A mixed-methods study in an Indonesian higher education context. *Journal of Education and Teaching (JET)*, 6(2), 518–533.
- Kholis, A. (2021). Elsa Speak app: Automatic speech recognition (ASR) for supplementing English pronunciation skills. *Pedagogy: Journal of English Language Teaching*, 9(1), 1–14.
- Levis, J. (2020). Revisiting the intelligibility principle in pronunciation teaching. *Journal of Second Language Pronunciation*, 6(1), 1–25.
- Li, S. (2022). Artificial intelligence and second language pronunciation learning: A review of recent research. *Computer Assisted Language Learning*, 35(8), 1795–1818.
- Liakin, D., Cardoso, W., & Liakina, N. (2015). Learning L2 pronunciation with a mobile speech recognizer: French. *CALICO Journal*, 32(1), 1–25.
- Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis*. Sage.
- Nguyen, T. T. (2021). Technology-assisted pronunciation learning in EFL classrooms. *International Journal of Emerging Technologies in Learning*, 16(4), 112–124.
- Richards, J. C. (2021). *Teaching and eeee digital age*. Cambridge University Press.
- Sa'diyah, H., Suhono, S., Pratiwi, W., Kushendar, K., & Sugiarto, S. (2023). Improving speaking skills of santri through drilling technique and cinema therapy in group counseling of Bakmin program. *Ghaidan: Jurnal Bimbingan Konseling Islam dan Kemasyarakatan*, 7(1), 65–74.
- Saito, K., & Akiyama, Y. (2020). Technology-mediated pronunciation training and second language learning. *Language Teaching Research*, 24(3), 1–18.
- Sugiyono, S. (2017). *Metode penelitian kualitatif, kuantitatif, dan R&D*. Yogyakarta: Auareta.
- Yin, R. K. (2018). *Case study research and applications* (Vol. 6). Sage Thousand Oaks, CA.
- Yosintha, R., & Rekha, A. (2022). "ELSA Speak" in an online pronunciation class: Students' voices. *IJEE (Indonesian Journal of English Education)*, 9(1), 160–175.
- Zou, D., Xie, H., & Wang, F. L. (2021). Future trends and research issues of technology-enhanced language learning: A bibliometric analysis. *Educational Technology Research and Development*, 69(1), 1–23.