



**STUDENT'S PERCEPTIONS TOWARDS  
PRONUNCIATION FEEDBACKS IN ENGLISH SPEAKING  
ACTIVITY THROUGH ELSASPEAK TASK**

**Sindi Nurhayati Hermina<sup>1</sup>, Setia Muljanto<sup>2</sup>, Pipih Setiawati<sup>3</sup>, Mohammad Soni<sup>4</sup>,  
R. Muhammad Satria Gyas Mustagis<sup>5</sup>**  
Institut Pendidikan Indonesia<sup>1,2,3,4,5</sup>  
e-mail: [sindinurhayatih@gmail.com](mailto:sindinurhayatih@gmail.com)

Diterima: 23/1/2026; Direvisi: 10/3/2026; Diterbitkan: 15/3/2026

**ABSTRAK**

Pelafalan bahasa Inggris merupakan aspek penting dalam komunikasi lisan, namun masih menjadi tantangan bagi siswa sekolah menengah pertama di Indonesia, terutama terkait ketepatan fonologis dan kepercayaan diri berbicara. Perkembangan teknologi kecerdasan buatan, khususnya *Computer-Assisted Pronunciation Training (CAPT)*, menawarkan peluang baru melalui penyediaan umpan balik pelafalan yang bersifat langsung dan personal. Penelitian kualitatif deskriptif ini bertujuan untuk mengeksplorasi persepsi siswa kelas IX terhadap umpan balik pelafalan yang diberikan oleh aplikasi ELSA Speak serta mengidentifikasi alasan di balik persepsi tersebut. Penelitian ini dilaksanakan di SMPN 6 Garut dengan melibatkan 34 siswa kelas IX yang dipilih melalui teknik *purposive sampling*. Data dikumpulkan melalui jurnal reflektif mingguan, observasi kelas selama latihan pelafalan, serta dokumentasi skor pelafalan dari aplikasi. Data dianalisis menggunakan teknik analisis tematik untuk mengidentifikasi pola pengalaman siswa. Hasil penelitian menunjukkan adanya polarisasi persepsi siswa. Sebagian siswa merasakan manfaat dari ELSA Speak karena lingkungan latihan yang aman, umpan balik visual yang jelas, dan dukungan terhadap pembelajaran mandiri. Sebaliknya, siswa lain mengalami frustrasi teknis, keterbatasan penjelasan artikulatoris, serta kelelahan digital akibat latihan yang repetitif dan minim interaksi manusia. Penelitian ini menyimpulkan bahwa ELSA Speak bukan solusi tunggal pembelajaran pelafalan dan lebih efektif digunakan sebagai alat pendukung dengan bimbingan guru. Penelitian selanjutnya disarankan untuk mengkaji peran mediasi guru dan penggunaan jangka panjang teknologi CAPT dalam konteks pembelajaran yang beragam.

**Kata Kunci:** *CAPT; ELSA Speak; pembelajaran mandiri; persepsi siswa; umpan balik pelafalan*

**ABSTRACT**

English pronunciation is essential for effective oral communication, yet it remains a persistent challenge for Indonesian junior high school students, particularly in terms of accuracy and speaking confidence. Recent advances in artificial intelligence, especially *Computer-Assisted Pronunciation Training (CAPT)*, have introduced new opportunities for providing immediate and personalized pronunciation feedback. This descriptive qualitative study explored ninth-grade students' perceptions of pronunciation feedback provided by the ELSA Speak application and examined the reasons underlying these perceptions. The study was conducted at SMPN 6 Garut and involved 34 ninth-grade students selected through *purposive sampling*. Data were collected through weekly reflective journals, classroom observations during pronunciation practice, and documentation of pronunciation scores generated by the application. The data



were analyzed using thematic analysis to identify recurring patterns in students' experiences. The findings reveal a clear polarization in students' perceptions. Some students benefited from ELSA Speak due to its safe practice environment, visual feedback, and support for autonomous learning, while others experienced technical frustration, limited articulatory explanations, and digital fatigue caused by repetitive practice and reduced human interaction. The study concludes that ELSA Speak is not a one-size-fits-all solution and is most effective when used as a supplementary tool supported by teacher guidance. Further research is recommended to examine the role of teacher mediation and long-term use of AI-based pronunciation tools in diverse learning contexts.

**Keywords:** *autonomous learning; CAPT; ELSA Speak; pronunciation feedback; student perception*

## INTRODUCTION

Pronunciation constitutes a fundamental component of oral communication in English because it directly influences intelligibility, comprehensibility, and learners' confidence in conveying meaning (Amrate & Tsai, 2025). Inaccurate pronunciation may lead to communication breakdowns and contributes to increased speaking anxiety as well as reduced classroom participation (Al-khresheh, 2024). Consequently, the development of pronunciation competence is particularly crucial at the junior high school level, as this stage provides the foundational basis for students' oral communication skills in subsequent stages of education.

English pronunciation remains one of the most challenging aspects of English as a Foreign Language (EFL) learning for junior high school students in Indonesia. Previous studies have shown that these challenges arise from substantial differences between Indonesian and English phonological systems, limited exposure to authentic spoken English, and affective factors such as fear of making mistakes and low self-confidence (Harahap et al., 2025; Wang, 2025). Similar learning conditions were observed during preliminary classroom observations in ninth-grade English classes at SMPN 6 Garut, where many students demonstrated difficulty pronouncing English words accurately and confidently. These difficulties were reflected in students' hesitation, avoidance of oral participation, and heightened anxiety during speaking activities that required them to pronounce English words or sentences aloud.

The urgency of this issue is further intensified by the predominant pronunciation teaching practices in the classroom, which largely rely on teacher modeling and brief repetition drills. Such practices provide limited opportunities for individualized practice and detailed corrective feedback. As a result, students frequently remain uncertain about their pronunciation accuracy and how to correct their errors. This uncertainty negatively affects students' confidence and reduces their willingness to engage in English-speaking activities. If pronunciation difficulties and negative learning experiences are not addressed at an early stage, they are likely to persist and become more deeply rooted at higher levels of education (Al-khresheh, 2024).

Alongside these challenges, rapid advancements in technology and artificial intelligence have introduced alternative approaches to pronunciation learning that can complement the limitations of conventional classroom instruction. One prominent approach is *Computer-Assisted Pronunciation Training* (CAPT), which employs *Automatic Speech Recognition* (ASR) technology to analyze learners' spoken output and provide immediate, personalized, and repeated feedback (Amrate & Tsai, 2025). CAPT environments allow learners to practice



pronunciation independently without the pressure of public performance, thereby potentially reducing speaking anxiety and increasing learner engagement (Yang & Chung, 2025).

ELSA Speak has emerged as a widely used mobile-based pronunciation learning application within the framework of Computer-Assisted Pronunciation Training (CAPT). The application provides real-time pronunciation analysis, visual and auditory feedback, and detailed evaluation of both segmental and suprasegmental features, such as vowel and consonant sounds, word stress, and intonation (Karimah, 2025). Through continuous interaction with this feedback, learners are expected to develop greater phonetic awareness and gradually improve their pronunciation accuracy. A growing body of research conducted in Indonesia has demonstrated the effectiveness of ELSA Speak in improving pronunciation accuracy, speaking fluency, and learners' confidence (Liunokas, 2025; Rosalina, 2025). However, most of these studies have adopted quantitative research designs and focused primarily on measurable learning outcomes, providing limited insight into learners' subjective experiences in interpreting and utilizing AI-generated pronunciation feedback. Moreover, many of these studies have been conducted at the senior high school or tertiary level, leaving the junior high school context relatively underexplored. Studies on technology acceptance in language learning also tend to emphasize perceived usefulness and ease of use in general, without placing AI-generated pronunciation feedback as the core learning mechanism (Han et al., 2025; H.-T. Hsu & Lin, 2022). Previous studies on pronunciation anxiety have largely focused on affective factors in isolation, leaving insufficient empirical evidence on the role of technology-based feedback in shaping learners' self-efficacy and confidence (Woreta et al., 2025; Zhang, 2022).

Based on the discussion above, several research gaps can be identified. First, there is a lack of qualitative research exploring junior high school students' lived experiences and perceptions of using ELSA Speak for pronunciation learning. Second, limited attention has been given to examining the effectiveness of AI-generated pronunciation feedback from students' perspectives, particularly in terms of clarity, usefulness, and its impact on confidence. Third, there is a shortage of integrative studies that combine technology acceptance, autonomous learning, and social cognitive perspectives in investigating AI-assisted pronunciation learning at the junior high school level (Chen & Zhao, 2022). In response to these gaps, the present study offers scientific novelty by reconceptualizing AI-based pronunciation feedback not merely as a corrective tool but as a multidimensional learning experience. This study seeks to understand how students interpret and utilize the feedback provided by ELSA Speak, how the feedback supports their autonomous pronunciation practice, and how it influences affective factors such as confidence, self-efficacy, and willingness to speak English.

The study is grounded in an integrated theoretical framework that draws upon the Technology Acceptance Model (TAM), autonomous learning theory, and Social Cognitive Theory. TAM is employed to explain how students' perceptions of the usefulness and ease of use of ELSA Speak influence their acceptance of the application as a pronunciation learning tool (Pham & Pham, 2025). Autonomous learning theory provides a lens for understanding how students plan, monitor, and evaluate their pronunciation practice independently through repeated interaction with AI-generated feedback (Mohebbi, 2025). Social Cognitive Theory, particularly the concept of *self-efficacy*, is used to examine how pronunciation feedback shapes students' confidence and engagement in speaking activities (Khuhro, 2024; Woreta et al., 2025). These theoretical perspectives are operationalized through key indicators presented in the theoretical framework table in the subsequent section.



Before presenting the research design, it is important to note that this study adopts a descriptive qualitative approach to gain an in-depth understanding of students' experiences, perceptions, and subjective meanings in using ELSA Speak. This approach is considered appropriate for exploring how students perceive the effectiveness of AI-generated pronunciation feedback and the reasons underlying their perceptions, which cannot be fully captured through quantitative measures.

**Table 1. Theoretical Framework of the Study**

Theoretical Framework	Key Terms	Observed Indicators
<b>Technology Acceptance Model (TAM)</b>	Perceived Usefulness (PU)	- Students' perceptions of whether ELSA Speak helps improve pronunciation accuracy - Students' views on the effectiveness of AI-generated feedback for learning pronunciation
	Perceived Ease of Use (PEOU)	- Students' perceptions of ease in using the ELSA Speak application - Ability to navigate features (feedback display, scoring, repetition tools) without difficulty
	Technology Acceptance	- Willingness to continue using ELSA Speak for pronunciation practice - Positive attitudes toward AI-assisted pronunciation learning
<b>Autonomous Learning Theory</b>	Independent Practice	- Frequency of self-directed pronunciation practice using ELSA Speak - Students practicing outside classroom requirements
	Self-Monitoring	- Students' ability to notice pronunciation errors through AI feedback - Awareness of specific pronunciation aspects (sounds, stress, intonation)
	Self-Evaluation	- Students' reflections on pronunciation improvement - Ability to judge progress based on feedback and scores provided by the application
<b>Social Cognitive Theory</b>	Self-Efficacy	- Students' confidence in pronouncing English words and sentences - Increased belief in their ability to improve pronunciation
	Affective Factors (Anxiety & Confidence)	- Reduction of speaking anxiety during pronunciation tasks - Increased comfort practicing pronunciation without fear of negative evaluation
	Engagement & Willingness to Speak	- Willingness to participate in oral activities - Increased motivation to practice pronunciation



<b>AI-Based Pronunciation Feedback (CAPT / ELSA Speak)</b> (Am et al., 2025)	Segmental Features	- Accuracy of vowel and consonant pronunciation as perceived by students - Awareness of phoneme-level corrections
	Suprasegmental Features	- Students' understanding of word stress, intonation, and rhythm - Ability to apply feedback to connected speech
	Feedback Clarity & Usefulness	- Students' perceptions of clarity, detail, and relevance of AI feedback - Perceived helpfulness of visual and auditory feedback

Based on the background of the study and the identified research gaps, this study aims to explore ninth-grade students' perceptions of AI-based pronunciation feedback provided by the ELSA Speak application in English pronunciation learning. This study also seeks to examine how AI-generated pronunciation feedback shapes students' pronunciation awareness, learning autonomy, and self-efficacy in pronunciation practice. Accordingly, the study is guided by the following research questions:

1. What are ninth-grade students' perceptions of the pronunciation feedback provided by the ELSA Speak application in English pronunciation learning?
2. How does AI-generated pronunciation feedback through ELSA Speak influence students' pronunciation awareness, learning autonomy, and self-efficacy?

Through this investigation, the study is expected to contribute theoretically to the growing body of research on AI-assisted pronunciation learning by providing a deeper, learner-centered understanding of how pronunciation feedback functions in real learning contexts. Practically, the findings are anticipated to offer pedagogically relevant insights for English teachers and educational practitioners in Indonesian junior high schools, particularly in designing pronunciation instruction that effectively integrates technology-based feedback to support students' autonomous learning, confidence development, and willingness to communicate in English.

## RESEARCH METHOD

This study utilizes a descriptive qualitative design to investigate how ninth-grade students at SMPN 6 Garut perceive AI-based pronunciation feedback. This specific design allows for a profound exploration of individual interpretations and emotional reactions toward technological interventions in language learning (Hasumi & Chiu, 2024). The participants involved 34 students from a single intact class, selected through *purposive sampling* to ensure they possessed relevant experience with the ELSA Speak application. By selecting learners aged 14 to 15 with varying levels of digital literacy and English proficiency, the researcher gathered a broad spectrum of perspectives on the effectiveness of automated feedback. This approach prioritizes the subjective viewpoints of the learners rather than statistical generalizations, aiming to uncover the specific mechanisms influencing their engagement with mobile-assisted language learning tools (Judijanto et al., 2024). Consequently, the research setting provides a naturalistic context to examine affective dimensions like confidence and anxiety in a technology-dependent environment (Lee, 2022).

To ensure data credibility through *triangulation*, three primary qualitative instruments were employed: reflective journals, classroom observations, and documentation (Maulida et al.,



2024). The students used the ELSA Speak application during pronunciation practice both in class and through guided independent study, receiving immediate feedback on accuracy, stress, and intonation. Reflective journals served as the central tool for capturing personal learning processes and emotional responses, utilizing weekly open-ended questions based on established reflective learning principles (Desky et al., 2025). Simultaneously, non-participant classroom observations were conducted to record verbal and non-verbal behaviors, such as repeated attempts or signs of frustration during the practice sessions. Additionally, screenshots of pronunciation scores were collected not for quantitative measurement, but to contextualize student progress and the perceived accuracy of the AI feedback. These systematic procedures facilitated a comprehensive documentation of students' interactive experiences with the AI tool throughout the implementation phase.

The collected data underwent a rigorous thematic analysis following the reflexive framework established by Braun and Clarke (2024). This analytical process initiated with the *data familiarization* phase, involving repeated engagement with journals and observation notes to gain a holistic understanding of the learners' experiences. Subsequently, the researcher generated initial codes to categorize meaningful units of data regarding feedback clarity and user autonomy. These codes were systematically clustered into broader themes, such as anxiety reduction and technical challenges, which were reviewed for internal coherence across the entire data set. Each theme was clearly defined to articulate how the AI feedback influenced the pronunciation learning process. Finally, the findings were interpreted by linking emerging patterns to theories of autonomous learning and self-efficacy. This systematic method provided a balanced narrative, explaining the nuanced differences between students who felt empowered by the application and those who encountered significant obstacles during their independent practice sessions.

## RESULTS AND DISCUSSION

### Results

#### Increased Awareness of Pronunciation Errors through Visual Feedback

One prominent finding indicates that students who perceived ELSA Speak as helpful experienced increased awareness of their pronunciation errors, particularly at the segmental level. Reflective journals revealed that the color-coded visual feedback enabled students to notice specific phonemes that were inaccurately pronounced, making abstract pronunciation errors more concrete and observable. This finding aligns with (H.-W. Hsu, 2022), who argues that visualized feedback enhances learners' phonological awareness by transforming auditory input into multimodal representations. Similarly, (Amrate & Tsai, 2025) note that explicit visual feedback in CAPT environments helps learners detect fossilized pronunciation errors that often go unnoticed in traditional classroom settings.

Student Code	Reflective Statement
S3	"I never knew that the <i>s</i> in <i>island</i> is silent. The app kept marking it red, and then I realized I always said it wrong."
S11	"When the color turned yellow or red, I knew exactly which sound was wrong. It made pronunciation clearer for me."
S19	"Before using ELSA, I thought my pronunciation was okay, but the feedback showed small mistakes I never noticed."



The findings support the Technology Acceptance Model (TAM), particularly the construct of perceived usefulness. Students regarded the feedback as useful because it provided clear and actionable information about their pronunciation errors (Pham & Pham, 2025). The scientific novelty of this study lies in revealing how junior high school students interpret visual AI-generated feedback as a learning cue rather than merely a corrective mechanism, thereby extending previous outcome-based studies.

### Reduced Anxiety and Perceived Safety in Pronunciation Practice

Another important theme concerns students' perception of ELSA Speak as a psychologically safe environment for pronunciation practice. Many students reported feeling less anxious when practicing with the application compared to speaking in front of teachers or peers. The absence of social judgment reduced fear of embarrassment and encouraged repeated attempts. This finding resonates with (Zhang, 2022), who emphasizes that pronunciation anxiety often stems from fear of negative evaluation. AI-based feedback, perceived as neutral and non-judgmental, allows learners to focus on improvement rather than performance. (Yang & Chung, 2025) further argue that private, technology-mediated practice environments can reduce affective barriers in speaking tasks.

Student Code	Reflective Statement
S7	"I feel more comfortable practicing with the app because nobody laughs if I make mistakes."
S14	"With ELSA, I can repeat many times without being shy. In class, I'm afraid to speak."
S22	"The app doesn't judge me, so I feel calmer and more confident."

From Social Cognitive Theory, these experiences contribute to increased *self-efficacy*, as repeated low-anxiety practice strengthens students' belief in their pronunciation ability (Khuhro, 2024; Woreta et al., 2025). The novelty here lies in demonstrating how AI feedback indirectly shapes affective outcomes, not merely pronunciation accuracy.

### Enhanced Learning Autonomy and Motivation through Gamified Feedback

Several students perceived the scoring and progress features of ELSA Speak as motivating and supportive of autonomous learning. Rather than viewing scores as evaluation, students interpreted them as personal challenges, encouraging repeated independent practice. This finding aligns with Autonomous Learning Theory, which emphasizes learner control, self-monitoring, and self-evaluation (Mohebbi, 2025). Gamified feedback supports these processes by enabling learners to track progress and set personal goals. (Han et al., 2025) also note that gamification increases learner engagement when feedback is perceived as informative rather than punitive.

Student Code	Reflective Statement
S5	"I want to practice again when my score is low. It feels like a challenge."
S18	"I can practice anytime at home and see my progress by myself."
S27	"I feel proud when my score improves, even a little."

The scientific contribution of this finding lies in illustrating how AI-generated feedback functions as a self-regulatory tool for junior high school learners, a population often assumed to be dependent on teacher guidance.

### Frustration with ASR Sensitivity and Technical Constraints

Despite positive perceptions, many students reported frustration related to technical limitations, particularly the sensitivity of the ASR system. Students felt discouraged when they



perceived their pronunciation as correct but continued to receive low scores. This finding echoes (Wang, 2025), who cautions that excessive ASR sensitivity may negatively affect learners' confidence if feedback is perceived as inaccurate. Similarly, (Li et al., 2025) argue that technical constraints such as background noise and microphone quality can distort feedback validity in mobile-assisted pronunciation learning.

Student Code	Reflective Statement
S9	"I said the word many times, but the score didn't change. It made me confused."
S16	"Sometimes the app doesn't hear my voice clearly, and I feel discouraged."
S30	"I started doubting myself because the app kept saying it was wrong."

These findings highlight an important nuance: AI feedback can undermine self-efficacy when learners perceive it as unreliable. This reinforces the study's novelty in showing that AI-based feedback is not inherently beneficial but contextually interpreted by learners.

### **Lack of Explanatory Feedback and Learning Fatigue**

Another negative perception concerned the lack of explicit articulatory explanations. Students expressed a preference for teacher modeling, particularly explanations involving tongue and mouth positioning. This supports (Tsang, 2025), who argues that effective pronunciation instruction should combine feedback with explicit phonetic guidance. Additionally, some students reported boredom and digital fatigue due to repetitive practice. (Zhang, 2022) notes that prolonged technology use without variation may reduce engagement, especially among younger learners.

Student Code	Reflective Statement
S12	"The app tells me it's wrong, but not how to fix it."
S21	"Sometimes I want the teacher to explain with mouth movement."
S33	"After many repetitions, I feel bored and tired."

This finding underscores the pedagogical implication that AI-based pronunciation tools should complement, not replace, teacher instruction. The novelty lies in highlighting learners' expectations for *human-mediated explanation* alongside AI feedback. Collectively, the findings demonstrate that AI-generated pronunciation feedback functions as a multidimensional learning experience, influencing cognitive awareness, affective responses, and autonomous learning behaviors. Unlike previous studies that primarily measured pronunciation improvement quantitatively, this study provides a learner-centered, qualitative perspective that reveals how junior high school students interpret, negotiate, and emotionally respond to AI-based feedback.

## **Discussion**

### **Students' Perceptions of AI-Based Pronunciation Feedback Provided by ELSA Speak**

The findings indicate that ninth-grade students held diverse and polarized perceptions of the AI-based pronunciation feedback provided by the ELSA Speak application. A substantial number of students perceived the feedback as helpful, as the color-coded visual indicators allowed them to identify specific pronunciation errors, particularly at the segmental level. These students viewed the feedback as clear, concrete, and actionable, allowing them to notice phonemes that were previously overlooked. As a result, the feedback was interpreted as a useful learning resource rather than merely an evaluative tool, supporting the construct of perceived usefulness in the Technology Acceptance Model (Pham & Pham, 2025).

However, other students perceived the feedback less positively due to technical constraints, especially the high sensitivity of the Automatic Speech Recognition (ASR) system.



These students reported confusion and frustration when their pronunciation attempts continued to receive low scores despite repeated efforts. Such experiences reduced their trust in the accuracy of the feedback and negatively affected their willingness to engage with the application. These contrasting perceptions demonstrate that students did not experience AI-generated pronunciation feedback uniformly, highlighting the importance of individual differences in technological readiness and learning preferences.

### **Influence of AI-Based Pronunciation Feedback on Pronunciation Awareness, Learning Autonomy, and Self-Efficacy**

Regarding pronunciation awareness, the findings show that AI-based visual feedback significantly supported students' ability to recognize and reflect on their pronunciation errors. The transformation of auditory input into visual representations helped students make abstract phonological errors more observable, which enhanced their awareness of specific pronunciation features. This finding is consistent with previous studies indicating that visualized feedback in CAPT environments promotes phonological awareness (Amrate & Tsai, 2025; H.-W. Hsu, 2022). This immediate and private feedback mechanism not only fosters a deeper understanding of phonetic intricacies but also cultivates learner autonomy by empowering students to independently identify and rectify their pronunciation deviations (Ginting & Hina, 2025; Qutub, 2023; Xodabande et al., 2025).

The results further reveal that the scoring and progress-tracking features of the ELSA Speak application fostered learning autonomy by encouraging independent and self-regulated practice. Many students interpreted pronunciation scores as personal challenges rather than external evaluations, motivating them to repeat tasks and monitor their own improvement. This pattern reflects key principles of autonomous learning theory, particularly learner control over monitoring and self-evaluation (Mohebbi, 2025). Notably, this finding challenges the common assumption that junior high school learners are highly dependent on teacher guidance. Instead, the AI-driven immediate and tailored feedback reduced reliance on instructors, thereby shifting the locus of control towards the learners (Huynh, 2024; Sariyani et al., 2022). This personalized and instantaneous feedback, characteristic of AI-powered tools, has been widely recognized for its efficacy in enhancing language acquisition and fostering learner autonomy across various educational contexts (Huynh, 2024).

With respect to self-efficacy, students who perceived the feedback as supportive reported reduced anxiety and increased confidence in pronunciation practice. The absence of social judgment created a psychologically safe environment that allowed students to practice without fear of embarrassment. Repeated low-anxiety practice strengthened students' belief in their ability to improve pronunciation, which aligns with Social Cognitive Theory emphasizing the role of mastery experiences in building self-efficacy (Khuhro, 2024; Woreta et al., 2025). Conversely, students who experienced persistent technical issues reported declining confidence and self-doubt, indicating that unreliable feedback can undermine self-efficacy rather than enhance it. Therefore, reliable AI-driven pronunciation tools are crucial for fostering a positive learning environment, promoting self-directed learning, and enhancing learners' overall speaking proficiency (Mohammadkarimi, 2024; Sariyani et al., 2022).

### **Pedagogical Implications of AI-Based Pronunciation Feedback**

The findings further reveal that AI-based pronunciation feedback alone is insufficient to address all learners' needs. Students' dissatisfaction with the lack of explicit articulatory explanations and learning fatigue caused by repetitive practice highlights the pedagogical limits of AI-based tools. Many students expressed a preference for teacher modeling and direct



explanation of mouth and tongue positioning, indicating that human mediation remains essential in pronunciation learning. Teachers play a crucial role in helping students interpret AI-generated feedback, clarify pronunciation mechanics, and provide affective support that technology cannot fully replicate (Jiang, 2022). This highlights a crucial need for a holistic approach that integrates AI-driven resources as a component within broader pedagogical strategies rather than a standalone solution (AbuSahyon et al., 2023; Godwin-Jones, 2024; Mohammadkarimi, 2024; Praveena & Anupama, 2025).

### **Implications for Integrating AI in Pronunciation Learning**

Taken together, the findings demonstrate that ELSA Speak functions most effectively as a supplementary pronunciation learning tool rather than a standalone solution. A balanced instructional approach that integrates AI-based feedback with teacher guidance can accommodate diverse learner needs, enhance pronunciation awareness, support autonomy, and protect students' self-efficacy. This integration represents the scientific novelty of the study by illustrating how technology acceptance, autonomous learning, and self-efficacy interact dynamically within real junior high school classroom contexts. Specifically, the synergy between AI's capacity for individualized feedback and the teacher's role in contextualizing linguistic nuances and providing motivational scaffolding can optimize pronunciation acquisition (Mohammadkarimi, 2024; Qiao & Zhao, 2023).

### **CONCLUSION**

This study demonstrates that AI-based pronunciation feedback provided through the ELSA Speak application does not function uniformly for all ninth-grade students at SMPN 6 Garut. The findings reveal a clear polarization in students' perceptions, indicating that while some students benefit from the application, others experience significant challenges that reduce its instructional value. Students who perceived ELSA Speak as helpful reported increased awareness of pronunciation errors through visual feedback, reduced anxiety resulting from the absence of social judgment, and greater learning autonomy supported by gamified feedback features. For these learners, the application facilitated independent practice and enhanced confidence in pronunciation learning.

Other students, however, experienced frustration due to the sensitivity of the Automatic Speech Recognition system, limited explanatory guidance, and the repetitive nature of pronunciation practice, which contributed to digital fatigue and reduced motivation. The absence of explicit articulatory explanations and meaningful human interaction caused pronunciation practice to be perceived as mechanical and less engaging. These findings suggest that AI-based pronunciation tools should not be regarded as a one-size-fits-all solution, particularly at the junior high school level, as their effectiveness is closely related to students' readiness for autonomous learning and their ability to interpret technical feedback.

Accordingly, the ELSA Speak application should be positioned as a supplementary tool rather than a replacement for teacher-led instruction. Teachers play a crucial role in mediating AI-generated feedback, explaining pronunciation mechanics, addressing technical limitations, and sustaining students' motivation. Effective pronunciation instruction at SMPN 6 Garut therefore requires a balanced pedagogical approach in which AI-based practice supports independent learning beyond the classroom, while in-class activities emphasize articulatory guidance, communicative practice, and affective support. Such an approach enables technology to enhance pronunciation learning without diminishing the essential human dimension of language education.



## REFERENCES

- AbuSahyon, A. S. E., Alzyoud, A., AlShorman, O., & Al-Absi, B. (2023). AI-driven technology and chatbots as tools for enhancing English language learning in the context of second language acquisition: A review study. *International Journal of Membrane Science and Technology*, 10(1), 1209. <https://doi.org/10.15379/ijmst.v10i1.2829>
- Al-khresheh, M. H. (2024). Phonetic challenges in English: The impact of mispronunciation of the bilabial plosive/p/on communication among Saudi EFL learners. *Cogent Arts & Humanities*, 11(1), 2390777. <https://doi.org/10.1080/23311983.2024.2390777>
- Am, Asriati., Haryono, urwo, Karimuddin, Judijanto, L., Fatubun, R. R., Purwaningsih, N. K., & Kristanto. (2025). *Applied linguistics*. [https://www.google.co.id/books/edition/Applied\\_Linguistics/VBacEQAAQBAJ?hl=id&gbpv=1&dq=AIBased+Pronunciation+Feedback+\(CAPT++ELSA+Speak\)&pg=PA82&printsec=frontcover](https://www.google.co.id/books/edition/Applied_Linguistics/VBacEQAAQBAJ?hl=id&gbpv=1&dq=AIBased+Pronunciation+Feedback+(CAPT++ELSA+Speak)&pg=PA82&printsec=frontcover)
- Amrate, M., & Tsai, P. (2025). Computer-assisted pronunciation training: A systematic review. *ReCALL*, 37(1), 22–42. <https://doi.org/10.1017/S0958344024000181>
- Braun, V., & Clarke, V. (2024). Supporting best practice in reflexive thematic analysis reporting in Palliative Medicine: A review of published research and introduction to the Reflexive Thematic Analysis Reporting Guidelines (RTARG). *Palliative Medicine*, 38(6), 608–616. <https://doi.org/10.1177/02692163241234800>
- Chen, Y., & Zhao, S. (2022). Understanding Chinese EFL learners' acceptance of gamified vocabulary learning apps: An integration of self-determination theory and technology acceptance model. *Sustainability*, 14(18), 11288. <https://doi.org/10.3390/su141811288>
- Desky, R. A., Fatmawati, S. N., Kastur, A., Mustain, M., Julianto, J., & Rukmi, A. S. (2025). Strategi dan tantangan pengembangan berpikir reflektif dalam pembelajaran IPA sekolah dasar. *DIDAKTIKA: Jurnal Pemikiran Pendidikan*, 31(2), 366–379. <https://doi.org/10.30587/didaktika.v31i2.9845>
- Fatimah, S., Salsabila, D., Karim, U. A. A., & Nazhif, M. A. (2024). Eksplorasi pandangan siswa terhadap penggunaan teknologi dalam proses pembelajaran di era digital. *Edukasi Elita: Jurnal Inovasi Pendidikan*. <https://journal.lpkd.or.id/index.php/Edukasi/article/view/290>
- Ginting, D. A., & Hina, S. (2025). Integrating conversational AI tools to enhance pronunciation, listening skills, and learner autonomy in second language education. *Global Education International Journal of Educational Sciences and Languages*, 2(3), 35. <https://doi.org/10.70062/globaleducation.v2i3.265>
- Godwin-Jones, R. (2024). Distributed agency in second language learning and teaching through generative AI. *arXiv*. <https://doi.org/10.48550/arxiv.2403.20216>
- Han, Y., Yang, S., Han, S., He, W., Bao, S., & Kong, J. (2025). Exploring the relationship among technology acceptance, learner engagement and critical thinking in the Chinese college-level EFL context. *Education and Information Technologies*, 30(11), 14761–14784. <https://doi.org/10.1007/s10639-025-13375-1>
- Harahap, G. R., Wahyuni, D., Sarahseti, D. F., & Lubis, Y. (2025). A psycholinguistic analysis of students' speaking difficulties in English as a foreign language classroom. *MUDABBIR Journal Research and Education Studies*, 5(2), 3217–3225. <https://doi.org/10.56832/mudabbir.v5i2.1900>



- Hasumi, T., & Chiu, M.-S. (2024). Technology-enhanced language learning in English language education: Performance analysis, core publications, and emerging trends. *Cogent Education*, 11(1), 2346044. <https://doi.org/10.1080/2331186X.2024.2346044>
- Hsu, H.-T., & Lin, C.-C. (2022). Extending the technology acceptance model of college learners' mobile-assisted language learning by incorporating psychological constructs. *British Journal of Educational Technology*, 53(2), 286–306. <https://doi.org/10.1111/bjet.13165>
- Huynh, Q. Q. (2024). A literature review on EFL students' integration of AI-based tools in English learning and the perceived impact on autonomy. *International Journal of AI in Language Education*, 1(2), 30. <https://doi.org/10.54855/ijaile.24123>
- Jiang, R. (2022). How does artificial intelligence empower EFL teaching and learning nowadays? A review on artificial intelligence in the EFL context. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.1049401>
- Judijanto, L., Wibowo, G. A., Karimuddin, K., Samsuddin, H., Patahuddin, A., Anggraeni, A. F., Raharjo, R., & Simorangkir, F. M. A. (2024). *Research design: Pendekatan kualitatif dan kuantitatif*. PT. Sonpedia Publishing Indonesia.
- Karimah, S. A. (2025). Global trends of segmental phonemes assessment in EFL: A systematic literature review. *SENTRI: Jurnal Riset Ilmiah*, 4(11), 3340–3353. <https://doi.org/10.55681/sentri.v4i11.4873>
- Khuhro, N. (2024). Study of classroom climate, student engagement, self-efficacy and learning experiences: An analyses employing social cognitive theory. *Academy of Education and Social Sciences Review*, 4(1), 118–126. <https://doi.org/10.48112/aessr.v4i1.722>
- Lee, S.-M. (2022). A systematic review of context-aware technology use in foreign language learning. *Computer Assisted Language Learning*, 35(3), 294–318. <https://doi.org/10.1080/09588221.2019.1688836>
- Li, I. W., Jackson, D., & Koshy, P. (2025). Student's reported satisfaction at university: The role of personal characteristics and secondary school background. *Higher Education*, 89(6), 1513–1531. <https://doi.org/10.1007/s10734-024-01286-y>
- Liunokas, Y. (2025). The effectiveness of the Elsa speak application in enhancing speaking skills among students at university. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, 13(2), 8221–8235. <https://doi.org/10.24256/ideas.v13i2.6955>
- Maulida, N. S. S., Happy, N., & Sugiyono, E. (2024). Analisis kemampuan berpikir reflektif dalam menyelesaikan masalah ditinjau dari gaya belajar siswa. *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, 10(2), 677–687. <https://doi.org/10.36989/didaktik.v10i2.2854>
- Mohammadkarimi, E. (2024). Exploring the use of artificial intelligence in promoting English language pronunciation skills. *LLT Journal: A Journal on Language and Language Teaching*, 27(1), 98. <https://doi.org/10.24071/llt.v27i1.8151>
- Mohebbi, A. (2025). Enabling learner independence and self-regulation in language education using AI tools: A systematic review. *Cogent Education*, 12(1), 2433814. <https://doi.org/10.1080/2331186X.2024.2433814>



- Pham, V. T. T., & Pham, A. T. (2025a). English major students' satisfaction with ELSA Speak in English pronunciation courses. *PLOS ONE*, 20(1), e0317378. <https://doi.org/10.1371/journal.pone.0317378>
- Pham, V. T. T., & Pham, A. T. (2025b). English major students' satisfaction with ELSA Speak in English pronunciation courses. *PLOS ONE*, 20(1), e0317378. <https://doi.org/10.1371/journal.pone.0317378>
- Praveena, T. L., & Anupama, K. (2025). Machine learning meets language learning: The transformative potential of artificial intelligence in English language instruction. *Journal Human Research in Rehabilitation*, 15(1), 133. <https://doi.org/10.21554/hrr.042511>
- Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1255594>
- Qutub, H. (2023). The role of pronunciation-focused corrective feedback in learning second language pronunciation. *Journal for Educators Teachers and Trainers*, 14(6). <https://doi.org/10.47750/jett.2023.14.06.010>
- Rosalina, A. (2025). The utilization of the ELSA speak application as a learning medium to improve pronunciation and English speaking skills among students. *Journal of Education Science*, 11(2), 72–81. <https://doi.org/10.33143/jes.v11i2.5428>
- Sariani, S., Miladiyenti, F., Rozi, F., Haslina, W., & Marzuki, D. (2022). Incorporating mobile-based artificial intelligence to English pronunciation learning in tertiary-level students: Developing autonomous learning. *International Journal of Advanced Science Computing and Engineering*, 4(3), 220. <https://doi.org/10.30630/ijasce.4.3.92>
- Tsang, A. (2025). The relationships between EFL learners' anxiety in oral presentations, self-perceived pronunciation, and speaking proficiency. *Language Teaching Research*, 29(4), 1639–1659. <https://doi.org/10.1177/13621688221102522>
- Wang, Y. (2025). Reducing anxiety, promoting enjoyment and enhancing overall English proficiency: The impact of AI-assisted language learning in Chinese EFL contexts. *British Educational Research Journal*. <https://doi.org/10.1002/berj.4187>
- Woreta, G. T., Zewude, G. T., & Józsa, K. (2025). The mediating role of self-efficacy and outcome expectations in the relationship between peer context and academic engagement: A social cognitive theory perspective. *Behavioral Sciences*, 15(5), 681. <https://doi.org/10.3390/bs15050681>
- Xodabande, I., Shiri, S., & Zohrabi, M. (2025). Exploring the impacts of an AI-driven instructional intervention on Iranian EFL learners' pronunciation skill development. *Discover Education*, 4(1). <https://doi.org/10.1007/s44217-025-00782-2>
- Yang, J. C., & Chung, C.-J. (2025). Effects of anxiety levels on English pronunciation learning in digital game-based learning. *International Journal of Human-Computer Interaction*, 0(0), 1–18. <https://doi.org/10.1080/10447318.2025.2539457>
- Zhang, Y. (2022). The effect of educational technology on EFL learners' self-efficacy. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.881301>