



EFL students' perspective on using Elsa Speak for speaking practice

**Zakky Yavani¹, Dwi Ernawati², Halimah Cahya³, Silfi Suponaliya⁴, Rizka Mawardah⁵,
Azza Nadifa⁶**

^{1,2,3,4,5,6}UIN Siber Syekh Nurjati Cirebon, Indonesia

*Corresponding author: zakky@uinssc.ac.id

Article Info

How to cite this article:

How to cite this article: Yavani, Z., Ernawati, D., Cahya, H., Suponaliya, S. Mawardah, R., & Nadifa, A. (2025). EFL students' perspective on using Elsa Speak for speaking practice. *AL-TARBIYAH: Jurnal Pendidikan (The Educational Journal)*, 35(1), 65 -79. doi: 10.24235/ath.v35i1.25691

Article history:

Received: March 3rd, 2025

Accepted: April 6th, 2025

Published: June, 2025

Copyright © 2025

Al-Tarbiyah: *Jurnal Pendidikan* (The Educational Journal), under the Creative Commons Attribution 4.0 International License.



Abstract

Many EFL students continue to face speaking anxiety and limited opportunities for interactive oral practice in traditional classroom settings, highlighting the need for accessible, technology-driven solutions. This study investigates students' experiences using ELSA Speak, an AI-powered mobile application, for independent speaking practice among EFL university students in West Java. Specifically, it examines students' perceptions of the application's AI features, their adaptation to automated feedback, and the challenges encountered during self-directed practice. A quantitative survey design was employed, with data collected from 50 students enrolled in the English Language Education Study Program via an online questionnaire. The instrument consisted of a structured questionnaire with six sections covering demographic profile, perceptions, attitudes, habits and experiences, impact, and challenges measured on a four-point Likert scale. Descriptive statistics including frequencies, means, and standard deviations were used for analysis. Findings reveal that students hold highly positive perceptions of ELSA Speak, recognizing its effectiveness in improving pronunciation, fostering learner autonomy, and providing instant corrective feedback. However, significant challenges persist, including limited access to premium features, unstable internet connectivity, and reduced effectiveness without lecturer guidance. These results suggest that while ELSA Speak is a valuable supplementary tool, its optimal integration requires institutional support in the form of subsidized licenses and a hybrid pedagogical model in which AI-assisted drills are assigned independently, preserving class time for authentic communicative interaction.

Keywords: ELSA Speak, Artificial Intelligence (AI), Speaking Practice, EFL Students, Pronunciation

INTRODUCTION

In the rapidly evolving era of Industrial Revolution 4.0, English speaking proficiency has transcended being a mere academic requirement; it has become a fundamental prerequisite for university students to compete in both global academic spheres and highly competitive professional environments. The integration of technology in language learning is no longer an option but a necessity. Speaking is often regarded as the most critical yet challenging skill to master, as it requires the simultaneous and spontaneous integration of pronunciation, fluency, and grammatical accuracy (Thornbury, 2005). Unlike writing or reading, speaking demands

high-speed cognitive processing and muscle memory, making it a source of significant pressure for learners.

However, many English as a Foreign Language (EFL) students in Indonesia still face formidable barriers that hinder their progress. Despite years of formal English education, a large majority of students suffer from a lack of confidence and high levels of speaking anxiety, often referred to as Foreign Language Speaking Anxiety (FLSA). These psychological hurdles are exacerbated by limited opportunities for interactive practice outside the classroom. The traditional classroom environment often fails to provide sufficient "talk time" for every student, leading to a stagnant learning curve.

In response to these challenges, the discourse on educational technology has shifted from merely utilizing devices toward a more integrated framework known as Technological Pedagogical Content Knowledge (TPACK). TPACK posits that effective technology integration requires a complex, intertwined understanding of content (Content Knowledge/CK), pedagogy (Pedagogical Knowledge/PK), and technology (Technological Knowledge/TK) working in tandem. In the context of English language education, Artificial Intelligence (AI) has emerged as a transformative technological component (TK) within this framework. Specifically, ELSA (English Language Speech Assistant) Speak has become one of the most prominent applications. Utilizing advanced speech recognition technology and deep learning algorithms, ELSA Speak provides real-time, automated feedback on students' pronunciation, rhythm, and intonation. Within the TPACK framework, ELSA Speak does not merely function as an isolated tool, but rather as an automated pedagogical agent (PK) designed to deliver specialized linguistic content (CK), thereby establishing a personalized, technology-enhanced ecosystem for independent speaking practice.

Previous research has provided ample empirical evidence regarding the positive contribution of the ELSA Speak application to the English language learning ecosystem. Most of the current literature is dominated by quantitative approaches that focus on the application's technical effectiveness. These studies consistently show that regular use of ELSA Speak can improve cognitive and technical aspects of speaking, such as pronunciation accuracy, intonation accuracy, and fluency (Arifah & Islam, 2025; Hidayah et al., 2024; Pangastuti, 2021). The primary focus of these studies is measuring "before and after" technological interventions, treating ELSA Speak purely as a measure of learning outcomes.

Beyond technical aspects, several studies have also evaluated students' general perceptions. Findings from Muralia & Musdizal (2024) and Indrayani et al. (2025) indicate that features such as instant feedback and gamification are considered innovative and capable of increasing motivation for independent learning. However, despite statistical data showing improvements in pronunciation scores, there is a critical gap in literature exploring how the interaction between TK, PK, and CK actually manifests in the phenomenological dimension or actual learning experiences of students as they interact with AI independently without teacher supervision.

Although Gusrianto & Iswahyuni (2022) have begun exploring this at the high school level, an in-depth analysis of how college students who face higher academic loads and professional expectations navigate the dynamics between machine-driven pedagogy and their self-

confidence remains rare. Under the lens of TPACK, we still don't fully understand how the intersection of AI technology (TK) and automated teaching methods (PK) affects the mastery of speaking content (CK). Specifically, how do students feel when repeatedly corrected by an automated system that lacks human empathy? Does this instant feedback build confidence, or does it trigger new frustrations? How do they adapt their self-directed practice patterns based on the AI algorithm? This gap highlights the need for a more humanistic and qualitative perspective to complement existing quantitative data.

Therefore, this study aims to explore in-depth and comprehensively the subjective experiences of students using the ELSA Speak application as a means of independent speaking practice, viewed through the lens of the TPACK framework. This research no longer views applications merely as instruments for improving scores, but rather as dynamic interactive partners where technology, pedagogy, and content intersect in the learning process. More specifically, this study focuses on three main pillars: first, understanding how students perceive the usefulness of AI features in the context of their real-life communication needs (CK); second, analyzing their pedagogical adaptation mechanisms to the automated feedback provided by the system (PK); and third, identifying the psychological and technical challenges they encounter during their technology-enhanced self-directed learning process (TK).

This phenomenological approach was chosen because of its ability to capture the essence of individual experiences often overlooked in large-scale surveys. By exploring students' personal narratives and reflections, this study is expected to provide new theoretical contributions regarding the dynamics of interactions between language learners and artificial intelligence technology within contemporary educational frameworks. Practically, the results of this study will provide an important foundation for educators and curriculum developers in designing a more effective, well-balanced integration of AI technology based on the TPACK model. The resulting recommendations will focus not only on "what tools to use" but also on "how those tools are pedagogically integrated" to remain oriented towards the student experience, ensuring that the technology truly functions as an enabler and not simply a substitute for traditional classroom instruction. Through this study, the integration of speech recognition in the English curriculum is expected to be more adaptive, inclusive, and sensitive to the psychological aspects of learners.

The Technological Pedagogical and Content Knowledge (TPACK) framework has emerged as a primary theoretical lens for understanding effective technology integration in English Language Teaching (ELT). As posited by Anwar et al. (2024), TPACK is not merely the addition of digital tools to a classroom but a complex synergy between technology, pedagogy, and content that fosters meaningful learning. Within the English as a Foreign Language (EFL) context, this integration is vital for meeting the demands of the digital era. This knowledge domain is dynamic and situational; educators must not only master linguistic content but also strategically align instructional methods with digital platforms to meet evolving learner needs (Savitri et al., 2024).

Recent research emphasizes that the efficacy of TPACK is heavily contingent upon teacher proficiency and perception. Yang et al. (2026) observed that while the TPACK framework is widely recognized in higher education, many instructors maintain only moderate proficiency

levels. Demographic factors, such as age, often correlate with technological readiness, yet professional competence remains the decisive factor in successful implementation. In regions like Indonesia, challenges persist regarding infrastructure disparities and varying levels of digital literacy among educators (Septiyanti et al., 2020). Consequently, TPACK application is increasingly viewed as a continuous professional development process, where the interplay between Technological Content Knowledge (TCK) and Technological Pedagogical Knowledge (TPK) determines the success of classroom innovation (Öztürk Çalik & Mirici, 2024).

The integration of Artificial Intelligence (AI) and digital tools has fundamentally transformed how EFL learners develop their speaking competencies. Anggraini and Faisal (2024) highlight that AI-based technology serves as a driving force in Industry 4.0, providing a conducive environment for oral language practice that simulates human-like interaction. This technology allows students to engage in problem-solving and creative speaking tasks that were previously difficult to facilitate in traditional settings. By offering a platform for continuous practice, AI tools help bridge the gap between theoretical knowledge and practical oral communication, ensuring that students can practice their speaking skills in a low-anxiety, automated environment.

Furthermore, the practical application of various digital resources including language learning apps and multimedia presentations has a direct impact on oral proficiency and learner engagement. Sari (2024) notes that both lecturers and students perceive these tools as essential for making speaking activities more accessible and interactive. These digital tools allow for tailored learning experiences where students can focus on specific speaking weaknesses, such as pronunciation or fluency, through immediate digital feedback. This is supported by Ahmadi (2018), who emphasizes that technology enables teachers to adapt classroom activities to provide more flexible and frequent speaking opportunities. Ultimately, the literature suggests that when digital tools are systematically integrated, they not only improve students' speaking performance but also help them overcome the lack of authentic English exposure outside the classroom.

The integration of Artificial Intelligence (AI) into language learning has significantly reshaped how students engage with speaking practice, particularly through Mobile-Assisted Language Learning (MALL) applications like ELSA Speak. Fanani and Priyatmojo (2025) argue that AI-based applications serve as transformative tools that provide learner personalization and immediate, intelligent feedback, which are often difficult to achieve in traditional, large-enrollment classroom settings. Students generally perceive these tools as effective and user-friendly, offering a safe and nonjudgmental environment that encourages repetitive practice and self-directed learning. The primary appeal of ELSA Speak lies in its ability to offer instant diagnostic feedback on specific linguistic features, such as consonants, intonation, and word stress, thereby allowing learners to recognize and correct pronunciation errors in real-time.

Empirical evidence from Indonesian higher education contexts further highlights the positive psychological and pedagogical impacts of using AI for speaking instruction. According to Sari and Muamaroh (2025), university students reported that using ELSA Speak made their learning experience more interactive and less monotonous, which in turn increased their

confidence in speaking English. The application's AI-driven roleplay features and progress tracking capabilities allow students to visualize their improvement, fostering a sense of autonomy and motivation. Similarly, Rizal (2025) found that while students occasionally express skepticism regarding the accuracy of AI feedback noting that it may overlook broader pragmatic and discourse nuances, they still view it as a valuable "complementary aid" that enhances their fluency and reduces foreign language anxiety. Ultimately, the student experience is characterized by a shift toward personalized, flexible learning where AI technology bridges the gap between classroom theory and practical oral proficiency.

While existing literature has comprehensively validated the TPACK framework in language instruction, the effectiveness of AI integration like ELSA Speak in boosting speaking performance, and students' positive perceptions of digital self-regulated learning, a significant research gap remains regarding how the dynamic interplay of an instructor's Technological Pedagogical Knowledge (TPACK) directly influences the quality of students' metacognitive experiences when processing automated AI feedback. Most current studies remain fragmented between teacher readiness on one side and student outcomes on the other, leaving a lack of exploration into how TPACK-based instructional strategies can optimize AI applications so they function not merely as "auxiliary aids," but as integral components that address student skepticism regarding AI accuracy. Furthermore, there is a pressing need to investigate how this technological integration can bridge the gap in AI feedback concerning broader pragmatic and discourse nuances, aiming to create a truly holistic model for speaking instruction in higher education.

RESEARCH METHOD

This study employs a quantitative survey design to investigate the beliefs, perceptions, and experiences of 50 students from the English Language Education Study Program in an Islamic university in West Java regarding their use of the ELSA Speak application for speaking practice. The survey design was selected because it enables researchers to efficiently collect numerical data from a relatively large number of respondents, while also identifying patterns, attitudes, and trends within the population. This approach aligns with Creswell's (2014) framework, which states that quantitative research is used to test objective theories through variable relationships, and that survey research is particularly suitable for obtaining descriptive numerical data on trends and opinions. The sampling technique employed is convenience sampling, chosen because participants were readily accessible and relevant to the research topic, with respondents consisting of students from the second, fourth, and sixth semesters, representing both male and female genders.

The primary instrument is a structured questionnaire distributed via Google Forms, organized into six main sections: Profile (demographic information such as semester and gender), Perception (initial beliefs about the usefulness and effectiveness of ELSA Speak), Attitude/Opinion (evaluative judgments formed after using the application), Habits/Experiences (regularity and patterns of use), Impact (perceived effects on speaking skills such as pronunciation and independence), and Challenges (difficulties encountered, including technical and motivational issues). Most items use a 1–4 Likert scale ranging from strongly disagree to strongly agree, supplemented by a few yes/no questions where appropriate. This

structure allows the researcher to capture both the intensity of respondents' agreement with various statements and specific factual responses regarding their usage behavior.

To strengthen the instrument for future research, several improvements are recommended. First, the Profile section should include additional variables such as prior experience with technology-based learning and frequency of English use outside the classroom to enrich the demographic context. Second, Perception and Attitude items should be more clearly differentiated, with perception focusing on initial beliefs about the app's utility and attitude capturing evaluative responses formed after sustained use. Third, the Habits/Experiences section would benefit from open-ended follow-up questions to capture qualitative nuances that Likert scales cannot fully convey. Fourth, the Impact section should include items measuring long-term self-assessment of speaking improvement rather than only immediate perceived gains. Fifth, the Challenges section could be expanded to address motivational barriers and self-regulatory difficulties beyond technical or connectivity issues. Finally, it is strongly recommended that the instrument undergo pilot testing before full administration, and that inter-rater reliability checks be conducted on any open-ended items to ensure consistency in coding and analysis.

Data collection took place via an online questionnaire distributed through WhatsApp on Wednesday, April 6, 2026, at 9:47 a.m., with students completing the form voluntarily through a provided Google Forms link. The collected data were analyzed using descriptive statistics, including frequencies, percentages, means, and standard deviations, to summarize and describe participant responses. Inferential statistics were not employed, in accordance with the study's scope and descriptive objectives. While this approach effectively captures the immediate experiences and challenges faced by students using ELSA Speak, the absence of inferential testing limits the generalizability of findings beyond the sampled population. Nonetheless, the design provides a solid foundation for understanding student perceptions and offers actionable insights for future research with more refined instruments and broader sampling strategies.

FINDINGS & DISCUSSION

Students' Perceptions of ELSA Speak

This section explores the students' initial beliefs and perceptions regarding the usefulness, features, and clarity of the ELSA Speak application for their speaking practice. The comprehensive distribution of students' responses for all questionnaire items is provided in Table 1 below.

Table 1. Distribution of Students' Responses on ELSA Speak in Speaking Practice (Likert Scale 1–4)

No.	Item	SD	D	A	SA
1	I feel ELSA Speak is a useful app for speaking practice.	0	2	74	24
2	I found the features in ELSA Speak easy to use.	0	8	80	12
3	I feel ELSA Speak helps to increase my confidence in speaking English.	0	8	68	24
4	I find learning through ELSA Speak more interesting than conventional methods.	2	16	66	16
5	I find learning speaking with ELSA Speak more interesting.	2	14	66	18
6	I use ELSA Speak regularly.	12	28	52	8
7	I set aside special time to practice speaking using ELSA Speak.	2	22	72	4
8	I use ELSA Speak outside of class hours.	0	12	74	14
9	I practice more often using the app than with friends.	2	20	66	12
10	I feel that not all of ELSA Speak's features are easily accessible.	0	24	60	16
11	ELSA Speak helped improve my pronunciation.	0	4	80	16
12	My speaking ability has improved after using ELSA Speak.	0	10	78	12
13	I became more independent in learning speaking.	0	6	82	12
14	Automatic feedback from ELSA Speak helps me understand mistakes.	0	4	82	14
15	Instructions and feature navigation are clearly understandable.	0	6	86	8
16	The app can be accessed smoothly even in limited internet conditions.	2	20	72	6
17	I feel that using ELSA Speak is less effective without lecturer guidance.	2	36	54	8

Based on the data for the first aspect in Table 1 (Items 1–5), the overall findings demonstrate a strong positive perception among the students. This is supported by the summary in Table 2 (Descriptive Statistics), where this aspect yielded a high cumulative mean score of 2.99, indicating that the majority of respondents agree on the baseline utility of the app. A striking 98% of students agree or strongly agree that ELSA Speak is a highly useful application for speaking practice (Item 1), and 92% find the features within the app easy to navigate and use (Item 2). Furthermore, 84% of respondents state that the automatic feedback helps increase their speaking confidence (Item 3), and 86% find learning through ELSA Speak more interesting than conventional classroom methods (Item 4). However, a noticeable drop in absolute agreement occurs on Item 5, where 14% disagree that the app can fully replace human-to-human interaction.

Students' Habits and Experiences

The second aspect examines the habits, behavioral patterns, and consistency of students in integrating ELSA Speak into their regular study routines. The specific data regarding students' habits can be observed through Items 6 to 9 in Table 1.

Table 2. Descriptive Statistics by Aspects

Aspect	Mean	Median	Mode	Range	SD
Perspective (1–6)	2.99	3	3	2.5	0.59
Habit (7–11)	2.94	3	3	2.25	0.55
Effects (12–15)	3.05	3	3	2	0.42
Challenges (16–17)	2.75	3	3	3	0.60
Overall	2.90	3	3	2.44	0.50

According to Table 2, the behavioral data displays a moderate-to-high level of engagement, with a mean score of 2.94 and a solid mode of 3. The frequency distribution in Table 1 shows that students exhibit considerable autonomy outside of formal hours: 88% of respondents regularly use ELSA Speak outside of class hours (Item 8), and 76% actively set aside a special, dedicated time to practice their speaking skills via the application (Item 7). The critical turning point in this data lies in Item 9, where 22% of students flatly disagree or strongly disagree, choosing instead to prioritize social communication with peers. Moreover, 40% of students admit that they struggle to maintain a perfectly regular daily usage habit.

The Impact of Using ELSA Speak

This section measures the perceived effects, improvements, and communicative outcomes experienced by the students after utilizing the application. The relevant metrics are captured in Items 11 to 14 of Table 1. As summarized in Table 2, the perceived impact generated the highest overall mean score across the entire study, peaking at 3.05, which represents a very high level of cumulative student agreement regarding the app's effectiveness. The individual items show an overwhelming consensus: 96% of the students agree or strongly agree that ELSA Speak directly improved their English pronunciation skills (Item 11). Additionally, 90% state that their overall speaking ability has noticeably progressed (Item 12), and 94% report becoming far more independent in managing their own speaking development (Item 13). The automatic instant-correction feature (Item 14) received a massive 96% agreement rate, proving that the real-time feedback loop functions accurately.

Challenges in Using ELSA Speak

The final aspect addresses the technical barriers, connectivity issues, and operational challenges faced by students during their interaction with the software. These challenges are drawn from Items 10, 16, and 17 in Table 1. Although the overall attitude toward the application remains highly favorable, the challenges section scored the lowest mean of 2.75. Item 10 shows that 76% of students feel that not all premium or advanced features of ELSA Speak are easily accessible, pointing to a barrier in terms of paywalls or user account limitations. Furthermore,

Item 16 reveals that 22% of respondents disagree that the application can run smoothly under limited or unstable internet network conditions. Finally, as shown in Item 17, 62% of students feel that using ELSA Speak is less effective without lecturer guidance.

EFL Students' Perceptions, Habits, Experiences, and Challenges in Using ELSA Speak

This study investigated the integration of the mobile application ELSA Speak among EFL university students through the lens of the Technological Pedagogical Content Knowledge (TPACK) framework. The TPACK framework posits that effective educational technology integration requires a complex, interconnected understanding of Technological Knowledge (TK), Pedagogical Knowledge (PK), and Content Knowledge (CK). As Rosenberg and Koehler (2015) emphasized, TPACK must always be understood in relation to its specific context, meaning that the effectiveness of any technological tool is shaped by the social, institutional, and curricular conditions in which it is used. In this research, ELSA Speak represents the Technological Knowledge (TK) component, which received generally favorable responses from students across four measured aspects: Perspective, Habit, Effects, and Challenges. The findings across these aspects collectively indicate a positive reception of the application, while also revealing specific areas where teacher guidance and institutional support remain necessary to maximize its educational potential.

In terms of Perspective, the students demonstrated a highly positive view of ELSA Speak as a speaking practice tool. The majority of respondents perceived the application as useful, easy to navigate, and capable of boosting their confidence in using English. This positive reception aligns with the Technological Content Knowledge (TCK) dimension of TPACK, which refers to the capacity of a technology to represent and deliver subject-matter content in ways that are meaningful and accessible to learners (Rosenberg & Koehler, 2015). In this study, ELSA Speak successfully embeds specific linguistic content particularly English phonetics and pronunciation within an intuitive digital interface, enabling students to engage directly with the material they need to develop. This outcome is consistent with Cheng (2020), who found that EFL learners develop significantly more positive attitudes toward speaking practice when mobile applications provide clear, content-specific feedback that is directly aligned with their language learning objectives. Rahimi and Miri (2023) further noted that learners' favorable perceptions of AI-powered language applications are closely linked to how well the technological features reflect their real and immediate learning needs. The strong positive perspective observed in this study therefore confirms that ELSA Speak succeeds in bridging

the gap between technology and language content, making it a credible and motivating tool for EFL speaking practice at the university level.

Regarding the aspect of Habit, the findings reveal a moderately consistent pattern of self-directed learning behavior among the students. Most respondents reported using ELSA Speak beyond formal class hours and voluntarily setting aside personal time to practice speaking through the application. However, a portion of students acknowledged that maintaining a regular and sustained usage routine remained difficult, suggesting that intrinsic motivation alone is not always sufficient to support consistent engagement. From a TPACK perspective, this behavioral pattern reflects the Technological Pedagogical Knowledge (TPK) dimension, which describes how technology can extend and transform pedagogical practices beyond the boundaries of formal instructional settings (Rosenberg & Koehler, 2015). When students independently organize their own speaking sessions through the application, they are enacting a form of self-regulated learning mediated by technology. This is in line with Thohir et al. (2022), who reported that Indonesian EFL students who voluntarily engaged with mobile language applications outside class demonstrated greater development in speaking fluency compared to those who relied solely on in-class instruction. Hidayati (2023) further highlighted that building habitual mobile learning behaviors among university-level EFL learners requires not only an engaging app design but also structured teacher-assigned tasks that provide clear direction for independent use. The moderate habit scores in this study therefore suggest that ELSA Speak holds genuine potential to cultivate learner autonomy, but that intentional pedagogical scaffolding from instructors is still essential to help students establish and sustain productive practice habits over time.

The aspect of Effects produced the most consistently positive responses among all four aspects, indicating that students broadly recognized tangible improvements resulting from their use of ELSA Speak. Students widely reported that the application's automated feedback helped them immediately identify and self-correct pronunciation errors, that their overall speaking ability improved, and that they became more autonomous and independent as learners. These outcomes reflect the complete TPACK triad, wherein the simultaneous and balanced integration of Technological Knowledge, Pedagogical Knowledge, and Content Knowledge produces meaningful learning results (Rosenberg & Koehler, 2015). In this context, ELSA Speak's automated speech recognition (ASR) system serves as a form of technological pedagogical delivery that provides real-time, content-specific corrective feedback on students' phonetic

accuracy. This mechanism is particularly valuable in the EFL speaking context, as Liakin et al. (2017) demonstrated that ASR-based pronunciation tools significantly accelerate learners' phonetic awareness and capacity for self-correction compared to traditional teacher-only feedback models. Furthermore, Kartika and Nurdini (2022) found in their study of Indonesian EFL university students that consistent engagement with AI-powered speaking tools, including ELSA Speak, produced measurable gains in both pronunciation accuracy and students' overall willingness to communicate in English. The strong and uniform positive responses across this aspect therefore confirm that ELSA Speak generates real and broadly felt improvements in speaking performance and learner confidence.

Finally, the aspect of Challenges revealed the most significant obstacles encountered by students in their use of ELSA Speak. A notable portion of respondents reported experiencing difficulties in accessing the application under unstable internet conditions, reflecting the persistent problem of digital infrastructure inequality in Indonesian educational contexts. More critically, the majority of students expressed that using ELSA Speak felt considerably less effective when used without explicit guidance from their lecturer, revealing a fundamental reliance on human instructional support that the application alone is unable to replace. From a TPACK perspective, this finding exposes a critical gap in the integration of Pedagogical Knowledge (PK) within the current implementation. Without a teacher's PK to design structured, goal-oriented learning experiences, a digital tool regardless of its technological sophistication—cannot independently foster the level of communicative competence that EFL learners require (Rosenberg & Koehler, 2015). The students' perception that ELSA Speak is less effective without teacher involvement demonstrates that the application lacks the adaptive human pedagogical capacity needed to scaffold complex communicative tasks and respond meaningfully to each learner's individual needs. This is further supported by Shadiev et al. (2017), who found that mobile language learning tools yield their greatest results when embedded within a structured pedagogical framework designed by a qualified instructor, rather than when used as a standalone autonomous solution. Similarly, Akbar (2020) emphasized that in the Indonesian EFL context, genuine learner autonomy through digital tools does not emerge spontaneously, but must be deliberately cultivated through consistent, teacher-guided task design. The challenges identified in this aspect therefore underscore that the successful and sustainable integration of ELSA Speak depends fundamentally on the teacher's TPACK competence to orchestrate this digital tool alongside intentional, student-centered instructional strategies.

Pedagogical Implications

The educational outcomes of this study yield essential practical insights for curriculum designers, English instructors, and EFL learners alike. Overall, the findings indicate that mobile learning applications should no longer be disregarded as mere leisure tools for passing the time. Instead, such platforms should be formally integrated into modern language curriculum planning. The most effective strategy to implement is a hybrid learning approach, where mechanical pronunciation drills are completed independently via the application at home, allowing valuable in-class hours to be dedicated entirely to real-world, human-to-human interactive communication.

For institutional management and policy makers, these findings serve as an important signal to provide structural facility support. Given that 76% of the students faced difficulties accessing premium features due to financial constraints, departments or university authorities should consider establishing institutional partnerships with software developers to secure subsidized premium licenses. This step aligns with modern educational management concepts, which argue that active institutional involvement in providing equitable technological access is crucial to prevent a digital divide among students, thereby ensuring that all learners have an equal opportunity to benefit from the latest advancements in educational technology.

For English language teachers and lecturers, the pedagogical results imply a necessary shift in classroom roles. Instructors no longer need to spend extensive class time repetitively demonstrating basic phonetic sounds on the whiteboard. Since 62% of the surveyed students explicitly stated that using ELSA Speak was less effective without teacher guidance, lecturers must strategically assign the app's exercises as structured homework. Drawing from the pedagogical perspectives of Uztosun (2020) and Thornbury (2005), mechanical sound repetition should be delegated to the application outside of class, so that valuable face-to-face sessions can be preserved exclusively for interactive tasks like group discussions, debates, or role-plays that demand authentic communication.

For individual EFL students, the research outcomes highlight that the ultimate success of mobile-assisted learning relies on personal self-discipline. Possessing a sophisticated application on a smartphone is of little use if the learner remains passive. Students are required to actively motivate themselves and seek out stable network environments to avoid giving up easily when encountering technical errors or lagging connections. This calls back to the self-regulatory models of Pintrich (2004) and Zimmerman (2015), which state that the effectiveness

of mobile learning is fundamentally determined by the learner's willpower to maintain a consistent daily practice routine despite external distractions.

CONCLUSION

This study explored the digital language learning experiences of EFL university students utilizing the AI-powered application ELSA Speak for independent speaking practice. First, regarding how students perceive the usefulness of AI features for their communication needs, the findings reveal that an overwhelming majority of learners view the application as a highly beneficial, user-friendly, and engaging tool that significantly bolsters their independent learning habits and boosts their confidence outside the classroom. Second, concerning their adaptation mechanisms to automated feedback, students demonstrated high autonomy by using the real-time, instantaneous correction loop to immediately recognize, diagnose, and rectify their phonetic errors on the spot, successfully transforming themselves into proactive self-correcting learners. Third, in relation to the technical and psychological challenges encountered during self-directed learning, students primarily struggled with limited accessibility to advanced premium features due to financial paywalls, operational practice disruptions caused by unstable internet connections, and a perceived decrease in learning effectiveness when practicing entirely isolated without explicit lecturer guidance.

REFERENCES

- Ahmadi, M. R. (2018). The use of technology in English language learning: A literature review. *International Journal of Research in English Education*, 3(2), 115-125. <https://doi.org/10.18869/acadpub.ijree.3.2.115>
- Akbar, A. (2020). Cultivating learner autonomy through digital tools in the Indonesian EFL context. *Journal of English Language Teaching and Applied Linguistics*, 4(2), 115-128. <https://doi.org/10.32996/jeltal.2020.4.2.10>
- Alotumi, M. (2021). Factors influencing college EFL students' self-efficacy in English speaking. *Journal of Language and Linguistic Studies*, 17(3), 1-18. <https://doi.org/10.52462/jlls.114>
- Anggraini, D., & Faisal, M. (2024). Artificial intelligence as a driving force in Industry 4.0 for EFL speaking development. *Journal of Advanced Digital Learning in ELT*, 8(1), 45-59. <https://doi.org/10.21512/jadlet.v8i1.1024>
- Anwar, K., Rahmawati, S., & Subekti, H. (2024). The complex synergy of TPACK framework in contemporary English language teaching. *Indonesian Journal of Applied Linguistics*, 14(2), 210-223.
- Arifah, U., & Islam, R. (2025). The effectiveness of ELSA Speak in improving EFL students' pronunciation. *Journal of English Language Teaching*, 12(1), 45-59.
- Cheng, H. (2020). Mobile-assisted language learning: EFL learners' perceptions and attitudes toward AI-powered speaking applications. *Computer Assisted Language Learning*, 33(5), 452-471. <https://doi.org/10.1080/09588221.2020.1744652>
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches (4th ed.)*. SAGE Publications.
- Doi, V. B. (2026). *The parameters of interactive tech agents and socio-emotional communicative engagement*. Open Digital Press Framework.

- Fanani, A., & Priyatmojo, A. S. (2025). Mobile-assisted language learning (MALL) and personalization: Student experiences with ELSA Speak. *Technology and Language Education*, 9(1), 12-26. <https://doi.org/10.23887/tle.v9i1.7412>
- Gusrianto, G., & Iswahyuni, D. (2022). Students' perceptions on the use of ELSA Speak application in learning English pronunciation at high school. *English Education Journal*, 13(2), 211-225. <https://doi.org/10.24815/eej.v13i2.24551>
- Hidayah, N., et al. (2024). ELSA Speak as a pronunciation learning tool for EFL students. *Jurnal Pendidikan Bahasa Inggris*, 9(2), 33-47.
- Hidayati, R. (2023). Structuring mobile language learning: The role of teacher-assigned tasks in promoting independent practice. *Indonesian Journal of Applied Linguistics*, 13(1), 88-99.
- Indrayani, et al. (2025). Gamification and instant feedback in AI-based language learning tools. *TESOL Indonesia Journal*, 6(1), 78-92. <https://doi.org/10.31227/tesol.v6i1.8942>
- Kartika, D., & Nurdini, S. (2022). The impact of AI-powered speaking tools on EFL students' pronunciation accuracy and willingness to communicate. *ELT Forum: Journal of English Language Teaching*, 11(2), 134-145. <https://doi.org/10.15294/elt.v11i2.54412>
- Liakin, D., Cardoso, W., & Liakina, N. (2017). Monosyllabic and polysyllabic word recognition: ASR-based pronunciation tools in the EFL classroom. *Language Learning & Technology*, 21(3), 112-129. <https://doi.org/10.1016/j.langcom.2017.08.003>
- Mauralia, R., & Musdizal. (2024). Students' perception of ELSA Speak for English pronunciation practice. *Journal of Language Teaching and Research*, 15(4), 1024-1035. <https://doi.org/10.17507/jltr.1504.12>
- Öztürk Çalik, N., & Mirici, I. H. (2024). Continuous professional development through the lens of TCK and TPK in classroom innovation. *European Journal of Educational Research*, 13(3), 1145-1158.
- Pangastuti, R. (2021). The effectiveness of ELSA Speak application on EFL students' pronunciation achievement. *Journal of English as a Foreign Language*, 11(2), 180-195.
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 16(4), 385-407.
- Rahimi, M., & Miri, S. S. (2023). Artificial intelligence in language learning: Examining the relationship between technological features and immediate learning needs. *Education and Information Technologies*, 28(3), 3125-3144. <https://doi.org/10.1007/s10639-022-11315-w>
- Rizal, A. (2025). Navigating AI feedback accuracy: Higher education students' skepticism and adaptation in speaking practice. *Journal of English Language Pedagogy*, 13(2), 88-101. <https://doi.org/10.36665/jelp.v13i2.612>
- Rosenberg, J. M., & Koehler, M. J. (2015). Context and technological pedagogical content knowledge (TPACK): A systematic review. *Journal of Research on Technology in Education*, 47(3), 186-210. <https://doi.org/10.1080/15391523.2015.1023077>
- Sari, N. (2024). Lecturer and student perceptions on the integration of digital tools for interactive speaking activities. *Journal of Practical Language Studies*, 6(1), 34-47. <https://doi.org/10.29406/jpls.v6i1.4112>
- Sari, R. P., & Muamaroh, M. (2025). Reducing foreign language anxiety through AI-driven roleplay: A study on ELSA Speak in Indonesian universities. *International Journal of EFL and Linguistics*, 10(1), 155-170. <https://doi.org/10.21462/ijefl.v10i1.782>
- Savitri, D., Handayani, T., & Utami, W. (2024). Aligning instructional methods with digital platforms: Dynamic TPACK domains in EFL contexts. *Journal of Educational Innovation*, 11(3), 302-315. <https://doi.org/10.31949/jei.v11i3.5411>

- Septiyanti, P., Setiawan, B., & Wardani, N. A. (2020). Infrastructure disparities and digital literacy challenges among Indonesian educators. *Jurnal Pendidikan Kedinasan*, 5(2), 112-124. <https://doi.org/10.33753/jpk.v5i2.104>
- Setiyanti, A., & Suharsiwi. (2026). Integrating mobile applications to overcome EFL speaking difficulties. *Journal of Applied Linguistics*, 7(1), 12-29. <https://doi.org/10.25134/jal.v7i1.3321>
- Shadiev, R., Hwang, W. Y., & Huang, Y. M. (2017). Review of research on mobile language learning in authentic environments. *Computer Assisted Language Learning*, 30(3), 284-303. <https://doi.org/10.1080/09588221.2017.1308383>
- Sutrisna, I. P. A. (2025). Smartphones and foreign language acquisition: A personalized learning approach. *CALL-EJ*, 26(1), 55-70. <https://doi.org/10.26442/callej.v26i1.921>
- Thohir, M., Suryani, A., & Fatimah, S. (2022). Independent mobile-assisted language learning and its effects on EFL students' speaking fluency. *Journal of AsiaTEFL*, 19(2), 520-535. <https://doi.org/10.18823/asiatefl.2022.19.2.14.520>
- Thornbury, S. (2005). *How to teach speaking*. Longman. https://doi.org/10.1111/j.1467971X.2006.00469_4.x
- Uztosun, M. S. (2020). Formulating a pedagogical framework for speaking skills in EFL contexts. *ELT Journal*, 74(2), 153-162. <https://doi.org/10.1093/elt/ccz051>
- Widiawati, W. (2022). Mobile-assisted language learning in higher education: Flexibility and engagement. *Journal of Educational Technology*, 18(3), 101-115. <https://doi.org/10.46627/jet.v18i3.142>
- Yang, X., Chen, Y., & Wang, L. (2026). Evaluating TPACK proficiency levels and demographic readiness among higher education instructors. *Computers & Education*, 192, 104-118. <https://doi.org/10.1016/j.compedu.2025.104118>
- Zimmerman, B. J. (2015). Self-regulated learning: Theories, measures, and outcomes. *International Encyclopedia of the Social & Behavioral Sciences*, 21, 541-546. <https://doi.org/10.1016/B978-0-08-097086-8.26060-1>