International
SMART Journal
International SOCIAL MENTALITY AND RESEARCHER THINKERS Journal
e-ISSN: 2630-631X

RESEARCH ARTICLE
SOCIAL SCIENCES

The Impact of Web-Based Platforms on Developing EFL Speaking Skills: A Comparison of AI-Supported Apps in Turkish EFL Context *

Web Tabanlı Platformların Yabancı Dil Olarak İngilizce Konuşma Becerilerinin Geliştirilmesine Etkisi: Türkiye'deki Yapay Zekâ Destekli Uygulamaların Karşılaştırması

Savaş Okyay [1]
Ramazan Zengin [2]

ABSTRACT

The development of effective English as a Foreign Language (EFL) speaking skills remains a significant challenge for many students. Traditional pedagogical approaches often fail to provide sufficient opportunities for authentic practice and individualized feedback, leading to limited communicative competence. As artificial intelligence (AI) advances in educational settings, web-based platforms have emerged as potential solutions to these long-standing issues. This study examines the efficacy of two AI-powered Web-based platforms, ELSA Speak and Praktika, in enhancing high school students' speaking abilities. A quantitative analysis of pre- and post-intervention assessments, supplemented by qualitative insights into student experiences, evaluates these tools' impact on overall speaking skill. The study also analyzes their impact on sub-skills including dimensions like fluency and accuracy. In addition, the study utilizes Foreign Language Enjoyment Scale (FLES) to reveal the apps' effect on affective factors such as language anxiety and motivation. By providing empirical evidence on AI-driven language learning, this research contributes to the discourse on innovative strategies for improving EFL instruction.

Keywords: EFL, Speaking Skills, Web-Based Learning, ELSA Speak, Praktika, Foreign Language Anxiety, AI-Driven Platforms, FLES, AI-based Language Learning

ÖZET

Yabancı Dil Olarak İngilizce (EFL) etkili konuşma becerilerinin geliştirilmesi, birçok öğrenci için hâlâ belirgin bir zorluk teşkil etmektedir. Geleneksel pedagojik yaklaşımlar genellikle yeterli düzeyde gerçek dünya pratiği ve bireyselleştirilmiş geri bildirim imkânı sunamadığından, iletişimsel yeterlik sınırlı kalmaktadır. Eğitim ortamlarında Yapay zekâ (YZ) kullanımı arttıkça, web tabanlı platformlar uzun süredir devam eden bu sorunlara potansiyel çözümler olarak öne çıkmıştır. Bu çalışma, iki YZ destekli web tabanlı platformun, ELSA Speak ve Praktika, lise öğrencilerinin konuşma becerilerini geliştirmedeki etkililiğini incelemektedir. Ön test ve son test verilerinin nicel analizi ile öğrenci deneyimlerine dair nitel içgörüler, bu araçların genel konuşma becerileri üzerindeki etkisini ortaya koymaktadır. Çalışma ayrıca akıcılık ve doğruluk gibi alt beceriler bakımından da etki analizi sunmaktadır. Buna ek olarak, Yabancı Dil Keyif Ölçeği (FLES) kullanılarak uygulamaların dil kaygısı ve motivasyon gibi duygusal faktörler üzerindeki etkisi değerlendirilmektedir. Yapay zekâ destekli dil öğrenimine ilişkin ampirik kanıtlar sunan bu araştırma, Yabancı Dil olarak İngilizce öğretimini iyileştirmeye yönelik yenilikçi stratejiler tartışmasına katkı sağlamaktadır.

Anahtar Kelimeler: Yabancı Dil Olarak İngilizce, Konuşma Becerileri, Web Tabanlı Öğrenme, ELSA Speak, Praktika, Yabancı Dil Kaygısı, YZ Destekli Platformlar, Yabancı Dil Keyif Ölçeği, YZ Destekli Öğrenme

How to Cite This Article
Okyay, S. & Zengin, R. (2025). "The Impact of Web-Based Platforms on Developing EFL Speaking Skills: A Comparison of AI-Supported Apps in Turkish EFL Context", International Social Mentality and Researcher Thinkers Journal, (Issn:2630-631X) 11(4): 566-579. DOI: https://doi.org/10.5281/zenodo.16421705

Arrival: 30 May 2025
Published: 26 July 2025

INTRODUCTION

In an increasingly globalized world, proficiency in spoken English is a critical skill for academic and professional success as it serves as lingua franca across various fields, from commerce to international affairs (Richards, 2008). Under these circumstances, for English language learners all around the world, including Turkish EFL learners, fluency in spoken English is particularly valuable.

However, as a result of traditional practices in formal education (Kirkgoz, 2007), many Turkish EFL learners struggle to acquire practical speaking skills. When it is combined with cultural anxieties about making mistakes, this environment often inhibits oral confidence, fluency, and accuracy (Al-Sohbi & Preece, 2018; Çelik & Aytin, 2014).

To address these challenges, technology-enhanced learning has emerged as a promising solution. In 2023 alone, language learning apps were downloaded 231 million times worldwide, reflecting the growing

* This article is derived from the author's MA thesis

[1] EFL Teacher, Gaziosmanpaşa MBA Schools, İstanbul, Türkiye
[2] Dr. Öğr. Üyesi, İstanbul Aydın Üniversitesi, Eğitim Fakültesi İstanbul, Türkiye

academic interest in digital language tools (businessofapps.com, 2024). ELSA Speak and Praktika are two of these platforms and are widely adopted, with a combined 25 million downloads. These two applications provide valuable and different perspective opportunities to language learners.

On the other hand, Shadiev and Yang (2020) point out that despite growing interest in AI-assisted language learning, empirical research on the effectiveness of such tools remains limited. The situation is the same for ELSA Speak and Praktika. Despite their popularity mentioned above, they have received relatively little scholarly attention regarding their estimated impact on speaking proficiency. Studies on Praktika are especially scarce in comparison to ELSA speak and other market-leading platforms. Studies that compare two differently-oriented speaking-enhancing platforms are even scarcer.

This study seeks to bridge this gap by evaluating the comparative effectiveness of ELSA Speak and Praktika in enhancing speaking proficiency among Turkish high school students.

Specifically, it examines:

1. Whether there is a significant difference in speaking proficiency improvement between learners using ELSA Speak and those using Praktika.
2. How students' perceptions of foreign language enjoyment (FLE) change after using these platforms.
3. The comparative strengths and limitations of each platform in supporting speaking skill development.

By addressing these research questions, the study not only evaluates the efficacy of these AI-based tools but also provides insights into their potential integration into EFL curricula.

LITERATURE REVIEW

This section provides an overview of the historical development of technology-assisted language learning, the role and challenges of speaking skills in EFL contexts, and the advantages and limitations of AI-driven web-based platforms. It also examines the core features of ELSA Speak and Praktika, alongside relevant findings from existing research, highlighting key gaps in the literature.

CALL, MALL and AI Era in EFL Learning

Initially, the integration of technology into language education has evolved over six decades, beginning with Computer-Assisted Language Learning (CALL) in the 1960s. Initially, CALL followed a behaviorist approach, relying on repetitive drills and grammar exercises (Warschauer & Healey, 1998). While these early applications reinforced linguistic structures, they lacked interactive and communicative components. In following decades, academics sought improvements and implemented new versions of this phenomenon; communicative CALL first and integrated or integrative CALL later.

Furthermore, with the rise of Mobile-Assisted Language Learning (MALL) in the 2000s, language learning became more flexible and interactive. Mobile applications introduced gamification, real-time feedback, and personalized learning paths, increasing learner engagement (Shahbaz & Khan, 2017). Social media platforms and online messaging further extended opportunities for spontaneous communication in English.

Over the past 15 years, AI-enhanced web-based platforms have transformed language education profoundly (Zheng & Young, 2024). These platforms utilize speech recognition, adaptive learning algorithms, and real-time conversational practice to create immersive learning environments. This transition from CALL to MALL, and now AI-driven applications, has significantly expanded access to interactive speaking practice.

Challenges in Developing Speaking Skills in Turkish EFL Contexts

Building on this historical overview, speaking is a fundamental skill in language learning, yet it remains one of the most challenging aspects of EFL learning. Oral proficiency requires managing grammatical accuracy, fluency, and pronunciation simultaneously, which increases cognitive load and anxiety (Richards, 2008). In Turkish EFL classrooms, several key challenges hinder speaking development:

- Limited Exposure to Authentic English: Students lack access to natural speech patterns, phonological variations, and interactive communication outside the classroom (Kara et al., 2017).
- Traditional Teacher-Centered Instruction: Classroom environments often prioritize written language over oral skills, with an emphasis on grammar and memorization rather than interactive communication (Kirkgoz, 2007).
- Foreign Language Anxiety (FLA): Fear of making mistakes or negative peer evaluation leads to avoidance behaviors, reducing students' willingness to engage in speaking activities (Asmali, 2016).

Given these constraints, there is a growing demand for innovative learning environments that provide interactive, low-anxiety speaking practice opportunities, which AI-supported platforms aim to address.

AI-Supported Platforms for Language Learning

Web-based platforms have revolutionized language education by offering flexible, interactive, and multimodal learning experiences. These platforms provide real-world language exposure, access to varied accents and pronunciation models, and adaptive feedback mechanisms (Ngoc, 2024). They also promote learner autonomy, encouraging students to practice outside the classroom and improve their confidence in oral production.

Key advantages of AI-driven platforms include:

- ✓ Real-time feedback on speaking performance (Chapelle, 2001; Burston, 2014, Jiang, 2022).
- ✓ Gamification and interactive tasks that maintain learner engagement (Jeong, 2016; Godwin-Jones, 2017).
- ✓ AI-powered adaptive learning, tailoring lesson difficulty based on learner progress (Alshumaimeri & Alshememry, 2023).
- ✓ Simulated conversational practice, allowing students to engage in contextualized dialogues without fear of judgment (Lin & Mubarok, 2021).
- ✓ Despite these benefits, some challenges persist:
- ✓ Access limitations due to device and internet requirements (Irzawati, 2021).
- ✓ Overreliance on automated feedback, which may lack the depth of human interaction (Folgieri et al., 2024).
- ✓ Cultural and contextual adaptation, as AI platforms may not fully align with learners' local linguistic needs.

Overview of ELSA Speak and Praktika: AI-Enhanced Learning Tools

Expanding upon these insights, ELSA Speak and Praktika are two AI-driven platforms designed to support EFL learners in developing speaking proficiency. While both provide real-time AI feedback, they differ in their focus and approach.

ELSA Speak: AI-Powered Pronunciation Coach

On the one hand, ELSA Speak is an AI-powered tool that enhances pronunciation and phonetic accuracy–subskills Levis (2008) identifies as crucial–and fluency using advanced speech recognition technology. Key features include:

- ✓ Personalized pronunciation feedback, aligned with CEFR standards.
- ✓ Speech analyzer, which detects phonetic inaccuracies.
- ✓ Gamified challenges to improve engagement.
- ✓ Adaptive learning paths, adjusting content based on user performance.

Praktika: AI-Supported Conversational Fluency Development

On the other hand, Praktika focuses on conversational fluency and real-world dialogue practice through AI-driven simulations. Key features include:

- ✓ AI avatars that simulate conversational partners.
- ✓ Structured dialogue practice, covering various real-life situations.
- ✓ Real-time corrective feedback during interactions.
- ✓ Progress tracking for continuous improvement.

Table 1: A Comparison of The Features of ELSA Speak and Praktika

Aspect	ELSA Speak	Praktika
Core Purpose	Pronunciation accuracy and fluency refinement.	Conversational fluency and adaptability.
Strengths	Advanced phonetic feedback with detailed correction	Immersive simulations that mimic real-world conversational contexts
	Adaptive learning paths that tailor lessons to individual progress	A broad range of conversational topics (151+)
	Includes tools such as the Speech Analyzer and Course Finder	Real-time corrections during dialogue simulations

Drawbacks	Limited focus on broader conversational skills and interaction	Does not provide detailed phonetic-level feedback or structured guidance for pronunciation issues
	May feel repetitive for advanced learners once major phonetic challenges have been resolved.	Limited evidence from research about its long-term effectiveness in improving conversational fluency or addressing cultural nuances in communication.

Source: Generated by the author

Hence, while ELSA Speak is ideal for learners seeking structured pronunciation improvement, Praktika is more suitable for those looking to develop spontaneous speaking skills. A hybrid approach combining both tools could offer a comprehensive solution for EFL instruction. Table 1 displays a comparison between the apps.

Research Gaps and Future Directions

Although AI-powered platforms present a promising avenue for improving speaking proficiency, existing research remains limited, in comparison to the growth in number of AI-using applications, particularly regarding Turkish EFL learners. Table 2 summarizes current literature and gap in the field:

Table 2: Research Gaps İn AI-Driven EFL Tools

Platform/Technology	Research Availability	Research Gaps
ELSA Speak	Limited studies (e.g., Darsih et al., 2020; Sholekhah and Fakhrurriana, 2023)	Long-term effectiveness in fluency and pronunciation
Praktika	No specific studies available	Empirical evidence on speaking skill improvement, cultural adaptability
General Conversational AI in EFL	General studies available on conversational AI but not Praktika-specific	Limited focus on cultural nuances, contextual learning

Source: Generated by the author

Prospectively, to optimize AI's role in EFL instruction, future research should explore the long-term effects of AI-based platforms, their impact on learner motivation, and their adaptability to different linguistic and cultural contexts. Integrating AI-driven tools into formal EFL curricula also requires systematic evaluation of their pedagogical effectiveness.

In summary, AI-enhanced web-based platforms have the potential to revolutionize EFL speaking instruction by providing personalized, real-time feedback and immersive language practice. However, existing research remains fragmented, necessitating further empirical investigation into the comparative effectiveness of platforms like ELSA Speak and Praktika. By addressing these research gaps, this study aims to contribute to the growing body of knowledge on AI-supported language learning and its implications for Turkish EFL learners.

METHODOLOGY

This section outlines the methodological framework adopted in this study. It details the research design, participant selection, data collection instruments, intervention process, and data analysis procedures, ensuring transparency in the steps taken to generate valid and reliable findings.

Research Design

This study employs a quasi-experimental research design using quantitative methods to evaluate the impact of AI-supported web-based platforms, ELSA Speak and Praktika, on Turkish EFL students' speaking performance. A quasi-experimental approach was selected to allow controlled intervention within a real educational setting while maintaining ecological validity.

The study consists of two experimental groups: Experimental Group 1 (ELSA Speak Group), whose students used ELSA Speak for six weeks, and Experimental Group 2 (Praktika Group), whose students used Praktika for six weeks. This structure facilitates a comparative analysis of how AI-supported platforms influence speaking skills. The quasi-experimental nature ensures that interventions are implemented within an authentic classroom environment while maintaining the ability to detect causal relationships.

A pre-test and post-test design was used to assess both speaking proficiency and affective variables (motivation and anxiety) before and after the intervention. Statistical analyses were conducted using SPSS V.27, allowing for rigorous quantitative evaluation.

In addition to the quantitative measures mentioned above, the study also utilizes semi-structured interview questions to provide qualitative insights and learn about the impact of the intervention on the students. Six volunteer students, 3 students from each group, answered 7 questions after the 6-week intervention and a

thematic analysis was applied on the obtained data to reveal key concepts out. All processes were administered by the teacher-researcher, ensuring consistency throughout the study.

Participants

The study was conducted with 40 private high school students enrolled in the 10th grade, aged between 14 and 15. All of them study B1-B2 (CEFR) level course books on an eight-hour-per-week schedule as part of the institutional curriculum. The research took place at a private high school in Istanbul.

A purposive sampling strategy was used to ensure participants had a proficiency level that made measuring the intervention's outcomes feasible. 10th graders in Turkish high school curricula are mostly on the verge of B1-B2 levels and it is a threshold for being competent in oral production or not. The teachers and the course books they study are the same to reduce any extra variables that may impact on the results. Also, convenience sampling strategy was used to make sure the possible participants are motivated to improve speaking skills, had access to the necessary technology (smartphone, tablet, or computer) and were willing and able to use AI-based learning applications.

Participants were from two different 10th-grade classes. After including them using the methods mentioned above, each class was assigned to one of the applications. Table 3 presents the participants' demographic information.

Table 3: Participants' Demographic Information

Category	Distribution
Application	ELSA Speak: 21, Praktika: 19
Gender	Female: 21, Male: 19
Self-Rated Proficiency Level	Beginner: 2, Low-Intermediate: 9, Intermediate: 22, High-Intermediate: 4, Advanced: 3
Self-Rated Performance	Far Below Average: 2, Below Average: 9, Average: 19, Above Average: 6, Far Above Average: 4

Source: Generated by the author

Data Collection Instruments

Multiple data collection tools were used to assess speaking performance and affective variables (motivation, enjoyment, and anxiety).

Standardized Speaking Test

A CEFR-aligned speaking test was developed by an experienced English teacher who is also the researcher. The speaking test consists of 24 topic-oriented questions.

To validate the test, the topics that the questions aim to assess are chosen in line with the topics from the program's coursebooks and those available in the applications. Its evaluation is based on four main criteria: task achievement (relevance of responses, task fulfillment), delivery of speech (fluency, hesitation, etc.), grammar & vocabulary (accuracy and frequency of structures, lexical variety), pronunciation (phonetic clarity, intelligibility, etc.). Each component was rated on a five-point scale (1–5) to enable quantitative pre-test and post-test comparisons. Student performance was assessed using a rubric that is appropriate for their estimated proficiency level and it is adapted from the Cambridge English Language Assessment Guidelines (cambridgeenglish.org).

To ensure reliability, students' performances were graded by two assessors; the main course teacher, who is also the researcher, and a second experienced English teacher. Inter-rater reliability processes are taken into consideration before the calculation of final scores. It was ensured that the assessed scores were within a twenty-percent interval.

Foreign Language Enjoyment Scale (FLES)

The Foreign Language Enjoyment Scale (FLES) (Dewaele & MacIntyre, 2014) was administered before and after the intervention to measure affective responses to speaking practice. It assessed: enjoyment of speaking English, confidence in language use, reduction in speaking-related anxiety.

Responses were collected via a 10-point online survey (Google Forms), with higher scores indicating greater enjoyment and reduced anxiety.

Semi-structured interview Questions

Volunteered students answered seven semi-structured interview questions to provide deeper insights into their experience with using the platforms. The teacher/researcher prepared these questions with the purpose mentioned above. Table 4 shows these questions:

Table 4: Semi-Structured Interview Questions

Can you describe your experience using [ELSA Speak/Praktika]?
What features of the platform did you find most useful for improving your speaking skills, and why?
Were there any challenges or frustrations you faced while using the platform? If so, how did you deal with them?
How did using the platform affect your confidence in speaking English?
Do you think the platform has helped you improve your speaking skills? If yes, in what way? If no, why not?
If you could change one thing about the platform to make it more effective for learning English, what would it be?
How does practicing on the platform compare to traditional classroom speaking activities?

Source: Generated by the author

Intervention Process

The study was conducted over six weeks, with structured engagement in ELSA Speak and Praktika. The teacher-researcher used both applications for two weeks in a row and documented effective platform usage strategies before the intervention. In Week 1, students completed the pre-speaking test and FLES questionnaire and then received detailed instructions on their respective platforms by the teacher-researcher according to the notes taken before. During Weeks 2–6, the ELSA Speak Group engaged in daily pronunciation and fluency exercises while the Praktika Group participated in interactive AI-based conversations, with the teacher monitoring platform usage, providing guidance on how to utilize specific features, and addressing technical issues. Because the researcher-teacher met the groups four days a week, it was possible to manually track platform usage. In Week 6, students completed the post-speaking test and FLES questionnaire. Following the final week of the intervention, related students answered the interview questions.

Data Analysis

Quantitative statistical methods were employed to analyze the impact of ELSA Speak and Praktika on speaking proficiency. Descriptive statistics—mean, standard deviation, and range—were calculated for pre-test and post-test scores. Normality tests (Kolmogorov–Smirnov and Shapiro–Wilk) assessed whether data followed a normal distribution. Inferential analyses included paired-sample t-tests to compare pre-test and post-test scores within each group, Wilcoxon Signed-Rank tests to verify t-test results in cases of non-normality, and Pearson/Spearman correlation analyses to examine relationships between speaking test scores and FLES results. All statistical analyses were conducted using SPSS V.27 to ensure rigorous data interpretation.

Content analysis was also applied to process the data derived from the qualitative phase of the study. Thematic analysis was conducted using related features of ATLAS.ti and GPTs.

Ethical Considerations

Informed consent was obtained from students and their parents, data confidentiality was strictly maintained with all responses anonymized, and ethical approval was granted by the relevant institutional review board, with official permission obtained from the Turkish Ministry of Education.

RESULTS & FINDINGS

This section presents the quantitative and qualitative findings obtained in the study. First, statistical analyses of students' speaking test results and affective factors (enjoyment, motivation, and anxiety) are reported. The results are summarized using tables and graphical representations. Content analysis of the answers to interview questions displays the findings and provides insights into the students' experiences.

Pre-Test and Post-Speaking Tests Results

Both groups (ELSA Speak, Praktika) were assessed using:

- ✓ Two speaking tests, as a response to RQ1 and RQ3, featuring four constructs; task achievement, delivery of speech, grammar & vocabulary, pronunciation.

Reliability, Normality, Correlation Analyses

The speaking tests showed strong internal consistency, according to the reliability analysis, with Cronbach's Alpha values for both groups above 0.84. Further confirming the substantial correlations between test items were the inter-item correlation coefficients ($r > 0.75$). Both groups' pre-test scores, according to normality

tests, had a normal distribution, even if the total data was only marginally normal. There was a slight variation in the ELSA Speak group and a substantial divergence in the aggregated data, but the post-test scores for the Praktika group satisfied the normality assumptions. While pointing out small normality deviations in the post-test scores, these results confirm the tests' reliability. Table 5 presents related analyses results.

Table 5: Reliability, Normality and Correlation Results of Speaking Tests

Measure	Praktika Group	ELSA Speak Group	Aggregated Data	Interpretation
Cronbach's Alpha	0.889	0.849	0.865	High internal consistency
Standardized Alpha	0.896	0.860	—	Strong internal consistency
Pearson Correlation	0.811	0.754	0.763	Significant positive correlation between pre- and post-test scores
Shapiro-Wilk (Pre-Test, p)	> 0.05	> 0.05	0.047	Normal (for groups), borderline normal (aggregated)
Shapiro-Wilk (Post-Test, p)	> 0.05	0.088	0.001	Normal (Praktika), slight deviation (ELSA Speak), non-normal (aggregated)

Source: Generated by the author

A Comparison of Mean Scores

Table 6 shows the means of speaking tests. Each group displayed improvement in overall speaking performance.

Table 6: Mean Scores Of Speaking Tests

Group	Pre-Speaking Test (M)	Post-Speaking Test (M)
ELSA Speak (n=21)	59.74	72.89
Praktika (n=19)	63.10	67.57

Source: Generated by the author

- ✓ ELSA Speak Group: Significant in overall speaking performance (+13.16).
- ✓ Praktika Group: Notable gains in overall speaking proficiency (+4.48).

Paired T-Test and Wilcoxon Signed-Rank Test Results

To assess the differences between experimental groups, a paired t-test was conducted first.

Table 7: Paired T-Test Results (Praktika And Elsa Speak Groups)

Group	Mean Difference	t-value	df	p-value	Significance
Praktika	-4.48	-1.156	20	0.261	Not Significant
ELSA Speak	-13.16	-3.339	18	0.004	Significant

Source: Generated by the author

- ✓ Elsa speak users showed a significant difference while praktika users showed an improvement in their overall score but it was not significant.

To verify paired t-test results a wilcoxon signed-rank test was also conducted:

Table 8: Wilcoxon Signed-Rank Test Results (Praktika And Elsa Speak Groups)

Group	Z-value	p-value	Significance
Praktika	-1.062	0.288	Not Significant
ELSA Speak	-2.707	0.007	Significant

Source: Generated by the author

Wilcoxon Signed-Rank Test results also support the paired t-test results in terms of significance of enhancements.

Dimensional Comparison of Speaking Test Results

To examine the dimensional impact of ELSA Speak and Praktika and compare the results with each other, another Wilcoxon-Rank Test was conducted and it yielded satisfactory results. Figure 1 displays the mean scores by dimension for separate groups.

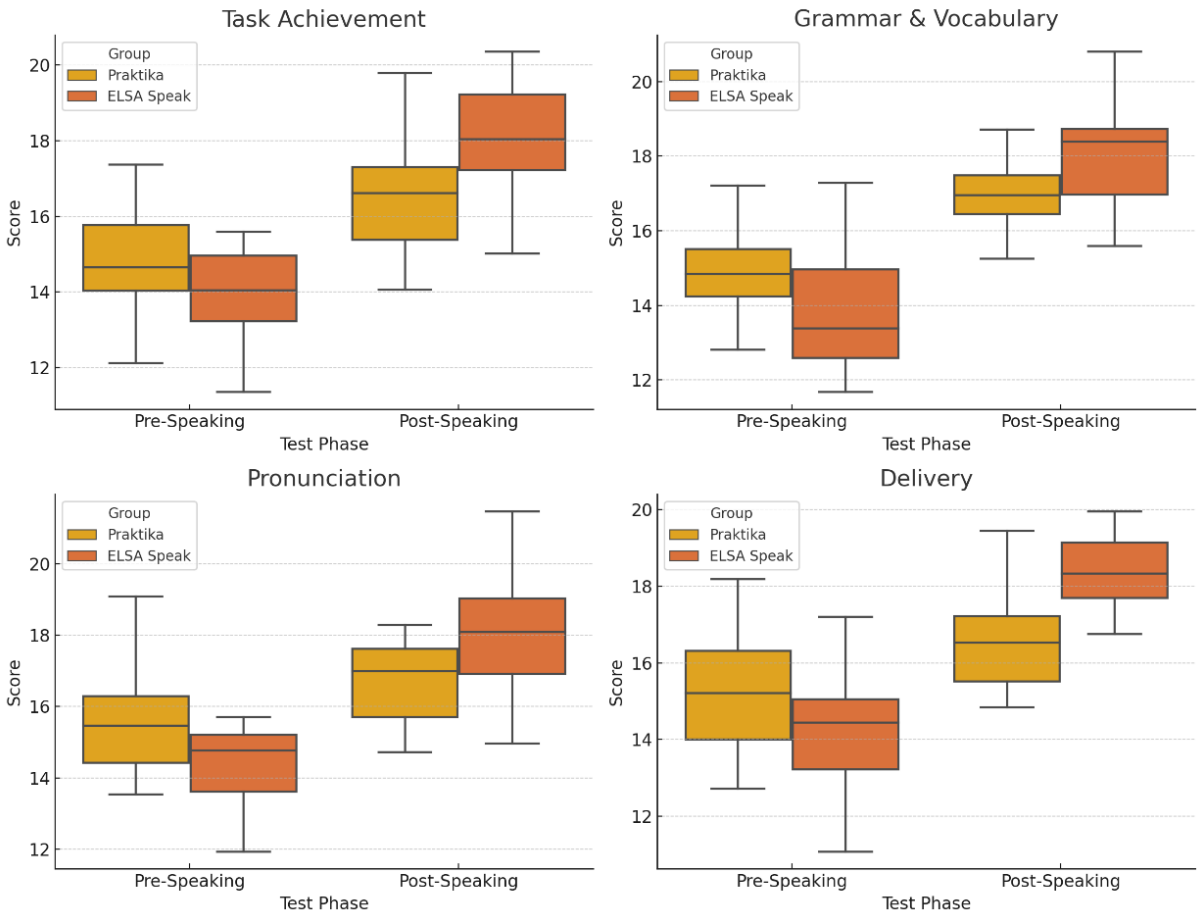

Figure 1: Pre- and Post-Speaking Test Score Distribution by Dimension
Source: Generated by the author

Table 9 presents the results of Wilcoxon Signed-Rank Test for each dimension. P-value of each aspect in rubric helps us understand which results are significant.

Table 9: Dimensions' Wilcoxon Signed-Rank Test Results

Group	Dimension	Z-value	p-value
Praktika	Task Achievement	-0.838	0.402
	Grammar & Vocabulary	-0.115	0.908
	Pronunciation	-0.774	0.439
	Delivery of Speech	-1.469	0.142
ELSA Speak	Task Achievement	-1.047	0.295
	Grammar & Vocabulary	-2.428	0.015
	Pronunciation	-2.563	0.010
	Delivery of Speech	-2.843	0.004

Source: Generated by the author

- ✓ Praktika: There are improvements in all dimensions but they are not statistically significant.
- ✓ ELSA Speak: There are statistically significant improvements in all dimensions except for 'Task Achievement'.

In conclusion, ELSA Speak seems to have a major influence on three important areas in the context of EFL performance: grammar and vocabulary, pronunciation, and speech delivery. Praktika did not result in statistically significant improvements across the tested dimensions.

Pre- and Post-FLES Results

The study utilizes FLES in response to RQ2. It inquiries into shifts in students' perceptions of FLE, as the Methodology part of the study notes.

Reliability, Normality Correlation Results of Pre-and Post-FLES

The reliability analysis confirmed that the FLES assessment items consistently measured the intended constructs, with Cronbach's Alpha values of 0.786 (pre-FLES) and 0.752 (post-FLES). Normality analysis,

based on a non-significant Shapiro-Wilk test, indicated that the pre-FLES dataset largely met normality assumptions, allowing for the use of a paired t-test, cross-verified with a Wilcoxon signed-rank test. Correlational analysis revealed strengthened relationships between key factors such as confidence and anxiety, enjoyment and mistake worry, and perceptions of a supportive learning environment, suggesting that the intervention positively influenced learners' attitudes and beliefs. Table 10 shows the results of the analyses mentioned above.

Table 10: Reliability, Normality Correlation Results of Pre-and Post-Fles

Name of Analysis	Result of the Analysis	Findings
Reliability	Cronbach's Alpha: - Pre-FLES: 0.786 - Post-FLES: 0.752	The scale demonstrated internal consistency, confirming that the assessment items reliably measured the intended constructs.
Normality	Shapiro-Wilk test	Most items in the Praktika and ELSA Speak groups remained normally distributed, with minor variations in skewness and kurtosis.
Correlational	Pearson and Spearman correlations	Several significant correlations indicated a positive shift in learners' attitudes:

Source: Generated by the author

Paired T-Test and Wilcoxon Signed-Rank Test Results

The analyses of the pre- and post-FLES results for the Praktika and ELSA Speak groups produced mixed outcomes. The paired t-tests showed no significant changes for the item "Even if I am well prepared for class, I feel anxious about it" in both groups (Praktika: $p = 0.602$; ELSA Speak: $p = 0.574$).

Given the small sample sizes and minor normality deviations, the Wilcoxon Signed-Rank Test was also performed for cross-verification. While the aggregated data largely mirrored the paired t-test findings, the Wilcoxon Test identified statistically significant changes in a few specific items:

- ✓ "I learned to express myself better in the English language" ($p = 0.008$),
- ✓ "Making errors is part of the learning process" ($p = 0.050$, marginally significant),
- ✓ "There is a good atmosphere" ($p = 0.011$).

These significant results were confirmed through cross-verification between the paired t-test and Wilcoxon test (See Table 11).

In conclusion, although most items showed no significant change, the intervention led to nuanced improvements in self-expression, error perception, and classroom atmosphere, which warrant further discussion in relation to RQ2.

Table 11: Cross-Verification of Paired T-Test and Wilcoxon Results

Item	Paired T-Test p-value	Wilcoxon p-value	Cross-Verification
Learned to express better	0.310	0.008	Yes
Errors part of learning process	0.372	0.050	Yes
Good atmosphere	0.523	0.011	Yes

Source: Generated by the author

Interview Questions Findings

While our quantitative results demonstrated significant improvements in pronunciation accuracy (ELSA Speak: $p < .01$; Praktika: non-significant) and modest gains in fluency (Section 4.2), students' own reflections help explain how and why these changes occurred. A thematic analysis of semi-structured interviews with six volunteers (three per group) revealed five key themes:

Features

Students drew attention to the platforms' targeted feedback and dialogue simulations. One of the ELSA Speak users noted, "ELSA Speak helped me improve my skills because it showed my pronunciation mistakes more accurately than many other apps," and this highlights the platform's strength in pinpointing phonetic errors.

Challenges

Repetitive content and trial-period limits emerged as drawbacks. As Student 4 noted, "After a while, the same things started happening," indicates the need for greater variety and sustained access.

Confidence

Interactive, low-pressure practice boosted some learners' self-belief, though not uniformly. Student 5 remarked, "It made me feel like I was talking to someone and increased my self-confidence and knowledge," whereas Student 1 admitted, "I didn't increase my confidence much, even though it made me feel like someone was in front of me".

Speaking Skills

Real-time correction and contextualized dialogue fostered fluency and accuracy. As Student 5 reflected on Praktika, "It increased my speaking and listening skills and helped me become more informed" underscoring the platforms' role in skill development.

Comparison

Many students contrasted AI practice with classroom speaking, often praising the reduced social pressure. "If you are a person that usually gets nervous around so many people, this app would help you because there is no one else," noted a Praktika user, suggesting that a private, judgment-free environment can alleviate anxiety.

Altogether, these student testimonials illuminate the mechanisms behind our statistical findings, the platforms' interactive feedback loops and low-anxiety environment, while also pointing toward content variety and extended, full-feature access as areas for better enhancement.

DISCUSSION

This section interprets the study's outcomes by connecting them to the research questions and extant literature. It appraises students' advancement, satisfaction levels, and the relative efficacy of ELSA Speak and Praktika, concluding with pragmatic implications and avenues for future inquiry.

Effectiveness of AI-Supported Platforms on Speaking Skills

RQ1: Is there a significant difference in speaking skill improvement between Turkish EFL learners using ELSA Speak and those using Praktika?

The findings demonstrate that ELSA Speak outperformed Praktika in enhancing speaking proficiency, with a mean increase of +13.16 versus +4.48. Statistical tests confirmed these results via the Wilcoxon Signed Rank Test ($p = 0.007$) and the paired t test ($p = 0.004$).

These outcomes correspond with contemporary research. Nuraini et al. (2024) documented a significant enhancement in pronunciation scores among vocational high school learners employing ELSA Speak. Febri (2023), Sholekhah and Fakhrurriana (2023) likewise reported tangible improvements in pronunciation accuracy and grammar using the same application. Such studies imply that ELSA's real time corrective feedback provides focused scaffolding analogous to instructor-led support.

From a theoretical perspective, Krashen's (1982) Affective Filter Hypothesis elucidates these results: ELSA's interventions likely diminished learners' anxiety—an imperative consideration in Turkish EFL environments, where exam-oriented pedagogies frequently exacerbate speaking-related stress according to Iskender and Savaşçı (2023).

While Praktika yielded moderate but non-significant enhancements in speech delivery and task completion, this pattern aligns with Ngo et al. (2023) and Elsani et al. (2023), who observed that conversational AI platforms often necessitate prolonged engagement to manifest measurable fluency gains. Godwin Jones (2023) further asserts that conversational AI demands more profound incorporation into learners' routines to achieve substantial improvement.

Moreover, the "Features" theme from learner interviews corroborated these quantitative findings: participants emphasized ELSA Speak's precise pronunciation feedback as a pivotal element in their demonstrable progress. This aligns with Bandura's (1997) self-efficacy framework, whereby successful, scaffolded performance fortifies learners' confidence in their linguistic capabilities.

Learners' Perceptions of Foreign Language Enjoyment (FLE)

RQ2: Is there a significant difference in learners' perceptions of foreign language enjoyment before and after using these platforms?

Although a few positive correlations emerged between enjoyment and reduced anxiety ($r = 0.664$, $p < 0.001$) and between confidence and anxiety reduction ($r = 0.676$, $p < 0.001$), aggregate FLE scores did not exhibit a statistically significant uptick. This diverges from Chang et al. (2022), Sayed et al. (2024), who documented pronounced psychological advantages of AI enhanced language tools. The absence of a notable shift here may

stem from the intervention's six-week span and contextual factors. Dincer et al. (2019) demonstrated that, within Turkey's EFL milieu, perceptions of instructor autonomy support were vital for engagement and performance, underscoring learners' dependency on classroom interactions. Similarly, Öztürk and Ok (2014) revealed that teacher immediacy and motivational conduct are perceived by students as the primary stimuli for their language learning engagement.

Interviews revealed a "Confidence" theme: most of the students recounted amplified self-assurance following the intervention. This concurs with MacIntyre and Gregersen's (2012) insights into how affirmative emotional experiences bolster language acquisition, though increased confidence alone may not instantaneously influence overarching enjoyment in exam-driven contexts.

Comparative Strengths and Weaknesses of ELSA Speak and Praktika

RQ3: What are the strengths and weaknesses of each platform in developing speaking skills among Turkish EFL learners?

ELSA Speak exhibited significant enhancements across pronunciation, grammar, and fluency sub skills. Dai and Wu (2023), Hapsari and Wu (2022) similarly determined that immediate, tailored corrective feedback substantially augments pronunciation precision, affirming our findings.

Praktika's conversational architecture aligns with Swain's (1995) Output Hypothesis, which emphasizes language production as a mechanism for fluency development. Despite Praktika's modest, non-significant quantitative gains, qualitative data illustrated its capacity to alleviate anxiety and encourage spontaneous speech. Zhang, Meng, and Ma (2024) observed that AI-mediated dialogue can elevate learners' willingness to communicate and diminish anxiety—echoing our interview feedback.

However, Koç and Savaş (2024) noted that AI tools frequently struggle to emulate genuinely unpredictable conversational dynamics. While Praktika may mitigate affective barriers, enduring fluency advancements likely hinge on complementing AI engagement with instructor facilitation and peer collaboration, particularly in Turkish EFL settings where blended pedagogical approaches are culturally consonant.

CONCLUSION

The findings confirm that ELSA Speak is particularly effective in improving pronunciation, grammar, and speech delivery, while Praktika enhances conversational fluency but does not significantly impact technical accuracy. Consequently, ELSA Speak appears to be the more effective tool for structured speaking improvement, offering targeted corrective feedback that refines pronunciation and grammatical precision. On the other hand, Praktika's emphasis on contextual dialogue practice provides an engaging yet less structured approach, making it more beneficial for students seeking confidence in spontaneous speech rather than accuracy.

Furthermore, thematic analysis of student interviews showed that low-or-no-pressure conversational practice (Praktika) and accurate, real-time feedback (ELSA Speak) were listed as important elements impacting learners' progress, placing the numerical gains in context and also pointing out areas that need improvement, like content variety and continuous access.

On the other hand, despite these improvements, the limited impact on foreign language enjoyment and anxiety underscores the importance of further research into individual and contextual factors that influence the psychological benefits of AI-assisted tools. While some students demonstrated increased confidence and reduced anxiety, the lack of a statistically significant overall increase in enjoyment suggests that AI-driven language learning experiences may not universally enhance motivation and engagement. This finding is particularly relevant for Turkish EFL learners, as it highlights the role of cultural attitudes toward language learning, classroom dynamics, and learner preferences in determining the effectiveness of AI-enhanced instruction.

Practical Implications

The study's findings have important implications for language educators, curriculum designers, and policymakers. Given the demonstrated benefits of AI-driven platforms, educators should integrate structured AI applications like ELSA Speak into language learning curricula to reinforce technical speaking skills. However, since conversational fluency also plays a crucial role in language proficiency, a blended learning model that incorporates both structured and conversational AI tools may provide the most comprehensive speaking development experience.

Additionally, teacher training programs should emphasize the pedagogical integration of AI-driven tools more which will ensure that instructors can guide students in maximizing the benefits of personalized feedback and adaptive learning pathways which are vital points as Cahyono and Rosita stated (2023) in their study. Educators must also recognize that technology alone is not a complete solution. Human interaction, feedback, and motivation still remain integral to effective language learning.

However, in many Turkish public schools, limited technology availability, large class sizes, and uneven internet connectivity can impede straightforward AI integration into the curricula.

We could list the guidelines in a few steps:

Phase 1: Needs & Infrastructure Audit. Partner with the Ministry of Education (MEB) to survey existing technology labs, internet bandwidth, and teacher digital-literacy levels; secure at least one dedicated AI-lab session per week (20–25 students per lab).

Phase 2: Curriculum Mapping & Teacher Training. Embed ELSA Speak like pronunciation modules into the MEB's "Speaking & Listening" units (e.g., Weeks 3–4 of Term 1) and conduct a two-day regional workshop for instructors on using AI dashboards, interpreting analytics, and troubleshooting.

Phase 3: Blended Implementation. Establish a weekly cycle combining: one AI-pronunciation lab session including ELSA Speak-like platforms for phonetic drills; one AI-conversation session including Praktika-like platforms for guided dialogues; and one in-class, teacher-led communicative-strategy workshop to increase authenticity.

By slotting platforms-based pronunciation drills directly into end-of-chapter speaking tasks within the coursebook syllabus and aligning it conversational scenarios with speaking outcomes, schools can ensure that AI practice targets both curricular goals and national-exam requirements.

Teacher training should include how to read AI reports, set individualized learning paths, and scaffold feedback, delivered through a mix of in-service days and online "train-the-trainer" follow-ups, with ongoing support via a dedicated help-desk or peer network moderated by EFL specialists.

Finally, pilot schools should collect pre- and post-implementation data on speaking scores and Foreign Language Enjoyment Scale (FLES) ratings each semester, then share anonymized reports with the MEB curriculum office to inform broader adoption.

By following these steps, grounded in audit, curriculum alignment, teacher empowerment, and rigorous evaluation, Turkish EFL programs can harness both structured and conversational AI tools. This blended approach maximizes gains in pronunciation, grammar, and fluency while maintaining the human interaction and motivational support essential to effective language learning.

Limitations & Future Research Directions

While this study contributes to the growing body of research on AI in EFL education, several limitations should be acknowledged. The small sample size limits the generalizability of findings, emphasizing the need for future studies with larger, more diverse participant groups. Additionally, the short six-week intervention period may not have been sufficient to capture long-term effects on fluency, motivation, and speaking confidence. Future research should adopt longitudinal designs to explore the sustained impact of AI-driven learning.

Moreover, the study primarily relied on quantitative methods, which, while valuable, do not provide in-depth insights into students' personal experiences, perceptions, and attitudes toward AI-based language learning. Future studies should incorporate qualitative approaches, such as interviews, focus groups, and reflective journals, to gain a richer understanding of learner experiences.

To further refine AI-assisted language learning, future research should explore:

1. Long-Term Impact of AI in EFL – Investigating the lasting effects of AI-driven speaking tools on fluency, accuracy, and learner motivation over an extended period.
2. Comparative Studies with Other AI Tools – Evaluating how other AI-enhanced platforms compare to ELSA Speak and Praktika in terms of effectiveness and learner engagement.
3. Cultural and Psychological Influences – Assessing how cultural attitudes toward AI, foreign language anxiety, and learner self-perception influence engagement and learning outcomes.

4. Teacher Training & AI Integration – Examining how educators can best leverage AI-driven tools in diverse classroom settings to enhance speaking instruction.

By addressing these research gaps, future studies can provide a deeper understanding of AI-driven learning and optimize its implementation in EFL education. Ultimately, the effective integration of structured feedback tools like ELSA Speak and conversational AI models like Praktika has the potential to transform speaking instruction, making language learning more engaging, accessible, and effective for diverse learner populations.

REFERENCES

Al-Sobhi, B. M. S., & Preece, A. S. (2018). Teaching English speaking skills to the Arab students in the Saudi school in Kuala Lumpur: problems and solutions, International Journal of Education and Literacy Studies, 6(1), 1-11.

Alshumaimeri, Y. A., & Alshememry, A. K. (2023). The extent of AI applications in EFL learning and teaching, IEEE Transactions on Learning Technologies, 17, 653–663.

Asmali, M. (2016). Willingness to communicate of foreign language learners in Turkish context, Procedia-Social and Behavioral Sciences, 232, 188–195.

Bandura, A. (1997). Self-efficacy: The Exercise of Control. New York, W. H. Freeman and Company.

Burston, J. (2014). MALL: The pedagogical challenges", Computer Assisted Language Learning, 27(4), 344–357.

Cahyono, A. & Rosita, R. (2023). The impact of using AI-based language learning platforms on English speaking skills of college students, Transformational Language Literature and Technology Overview in Learning (Transtool), 2(2), 1-8.

Chang, Y., Lee, S., Wong, S. F., & Jeong, S. (2022). AI-powered learning application use and gratification: An integrative model, Information Technology & People, 35(7), 2115–2139.

Chapelle, C. A. (2001). Computer Applications in Second Language Acquisition: Foundations for Teaching, Testing, and Research, Cambridge University Press.

Curry, D. (2024). Language learning app revenue and usage statistics, Business of Apps.

Çelik, S., & Aytin, K. (2014). Teachers' views on digital educational tools in English language learning: Benefits and challenges in the Turkish context, Teaching English as a Second or Foreign Language--TESL-EJ, 18(2), 1-18

Dai, Y., & Wu, Z. (2023). Mobile-assisted pronunciation learning with feedback from peers and/or automatic speech recognition: A mixed-methods study, Computer Assisted Language Learning, 36(5–6), 861–884.

Darsih, E., Wihadi, M., & Hanggara, A. (2020). Using ELSA app in speaking classes: Students' voices", International Journal of Advanced Multidisciplinary Research and Studies, 295-300.

Dincer, A., Yeşilyurt, S., Noels, K. A., & Vargas Lascano, D. I. (2019). Self-determination and classroom engagement of EFL learners: A mixed-methods study of the self-system model of motivational development, SAGE Open, 9(2), 1-16

Elsani, E., Salsabila, R., Putra, M. F. I., Nabila, N. K., & Nahartini, D. (2023). "The effect of using Elsa Speak app for first-semester students' English speaking proficiency", Edukatif Jurnal Ilmu Pendidikan, 5(6), 2644–2655.

Febri, Y. (2023). The effect of using ELSA Speak application to improve students' pronunciation, Jolly Journal of English Education, 1(2), 54-63.

Folgieri, R., Gil, M., Bait, M., & Lucchiari, C. (2024). AI-powered personalised learning platforms for EFL learning: Preliminary results, CSEDU, 2(0), 255–261.

Godwin-Jones, R. (2017). Smartphones and language learning, Language Learning and Technology, 21(2), 3–17.

Godwin-Jones, R. (2023). Emerging spaces for language learning: AI bots, ambient intelligence, and the Metaverse", Language Learning & Technology, 27(2), 6–27.

Hapsari, I. P., & Wu, T. (2022). AI chatbots learning model in English speaking skill: Alleviating speaking anxiety, boosting enjoyment, and fostering critical thinking, Lecture Notes in Computer Science, 444–453.

Irzawati, I. (2021). The utilization of digital platforms in online learning: EFL students' perspectives, Indonesian EFL Journal, 7(2), 131–138.

İskender, Y. C., & Savaşçi, M. (2023). Foreign language teaching anxiety of Turkish EFL teachers, Kastamonu Education Journal, 31(3), 464-483.

Jeong, K. (2016). A study on the integration of Google Docs as a web-based collaborative learning platform in EFL writing instruction, Indian Journal of Science and Technology, 9(39), 1-7.

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, 1-18.

Kara, E., & Ayaz, A. D. (2017). Challenges in EFL speaking classes in Turkish context, European Journal of Language and Literature, 3(2), 50-58.

Kirkgoz, Y. (2007). English language teaching in Turkey: Policy changes and their implementations", RELC Journal, 38(2), 216–228.

Koç, F. Ş., & Savaş, P. (2024). The use of artificially intelligent chatbots in English language learning: A systematic meta-synthesis study of articles published between 2010 and 2024, ReCALL, 37(1), 4–21.

Levis, J. M. (2018). Intelligibility, Oral Communication, and the Teaching of Pronunciation, New York, Cambridge University Press.

Lin, C.-J., & Mubarok, H. (2021). Learning analytics for investigating the mind map-guided AI chatbot approach in an EFL flipped speaking classroom, Educational Technology and Society, 24(4), 16–35.

MacIntyre, P., & Gregersen, T. (2012). Emotions that facilitate language learning: The positive-broadening power of the imagination, Studies in Second Language Learning and Teaching, 2(2), 193–213.

Ngoc, N. K., Thanh, N. T. M., & Thanh, N. T. M. (2024). Tertiary students' perceptions on Elsa Speak application for pronunciation learning, European Journal of Applied Linguistics Studies, 7(1), 1–13.

Nuraini, W. R., Setyawan, M. A., & Pambayun, R. N. (2024). The effect of ELSA Speak on pronunciation of English speaking at vocational high school, English Language Teaching Methodology, 4(3), 499–511.

Öztürk, E., & Ok, S. (2014). Motivational behaviors of teachers in Turkish EFL classes: Perception of students, Journal of Language and Linguistic Studies, 10(2), 120-133.

Richards, J. C. (2008). Teaching Listening and Speaking: From Theory to Practice, New York, Cambridge University Press.

Sayed, B. T., Younes, Z. B. B., Alkhayyat, A., Adhamova, I., & Teferi, H. (2024). To be with artificial intelligence in oral test or not to be: A probe into the traces of success in speaking skill, psychological well-being, autonomy, and academic buoyancy, Language Testing in Asia, 14(49), 1-21.

Shadiev, R., & Yang, M. (2020). Review of studies on technology-enhanced language learning and teaching, Sustainability, 12(2), 524.

Shahbaz, N. M., & Khan, N. R. M. I. (2017). Use of mobile immersion in foreign language teaching to enhance target language vocabulary learning, MIER Journal of Educational Studies Trends & Practices, 7(1), 66–82.

Sholekhah, M. F., & Fakhrurriana, R. (2023). The use of ELSA Speak as a mobile-assisted language learning (MALL) towards EFL students' pronunciation, JELITA Journal of Education Language Innovation and Applied Linguistics, 2(2), 93–100.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), Principle and Practice in Applied Linguistics 125–144. Oxford University Press.

Thi-Nhu Ngo, T., Hao-Jan Chen, H., & Kuo-Wei Lai, K. (2024). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis, ReCALL, 36(1), 4–21.

Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview, Language Teaching, 31(2), 57–71.

Zhang, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate, System, 121, 103259.

Zheng, L., & Yang, Y. (2024). Research perspectives and trends in artificial intelligence-enhanced language education: A review, Heliyon, 1-11.