

The Impact of ELSA Speak on Enhancing English Pronunciation and Speaking Skills

Ashwag A. Almethen

Department of Curriculum and Instruction - College of Education - King Faisal University - Al-ahsa - Saudi Arabia

(Received: 05-04-2026; Accepted: 07-06-2026)

Abstract: This study explored the impact of ELSA Speak as a supplementary instructional tool for enhancing English pronunciation and speaking skills among secondary-level students. A quantitative quasiexperimental pretest-posttest design was used, with the participation of 44 students divided into two groups: the experimental and control groups. Within a 4-week period, ELSA Speak was used in the experimental group as a supplementary tool to traditional instruction. However, only the traditional approach was used for the control group. Oral assessment and a Likert-scale teacher perception questionnaire were used to collect data.

Based on the findings, the experimental group achieved a significantly higher mean performance (9.31) compared to that of the control group (2.12). The results showed the following:

- These findings were confirmed with an independent t-test, indicating a higher performance for the experimental group.
- The evaluators also reported higher levels of engagement among participants in the experimental group, noting that ELSA Speak had an important role in providing individualized practice.
- The findings of the study indicated that AI-powered tools such as ELSA Speak can meaningfully complement traditional instruction in language teaching.

Keywords: Artificial Intelligence - MALL - Communicative Language Teaching

أثر برنامج ELSA Speak على تحسين مهارات النطق والتحدث باللغة الإنجليزية

أشواق أحمد المذن

قسم المناهج وطرق التدريس - كلية التربية - جامعة الملك فيصل - الأحساء - المملكة العربية السعودية

(تاريخ الاستلام: 2026-04-05؛ تاريخ القبول: 2026-06-07)

مستخلص البحث: استهدفت الدراسة بيان فعالية برنامج (تحدث ألسا) كأداة مساعدة لتحسين مهارات النطق والتحدث باللغة الإنجليزية لدى طلاب المرحلة الثانوية. ولتحقيق أهداف الدراسة استخدمت الباحثة تصميمًا كميًا وفق المنهج شبه التجريبي الذي يعتمد على المجموعتين (التجريبية والضابطة)، وعلى الاختبار القبلي والبعدي. وقد بلغ عدد المشاركين في الدراسة (44) طالبًا، قُسموا إلى مجموعتين: الضابطة وعددها (22) طالبًا، والتجريبية وعددها (22) طالبًا. وتم تطبيق الدراسة على مدار أربعة أسابيع، استخدم برنامج (تحدث ألسا) مع المجموعة التجريبية كأداة مساعدة للتعليم التقليدي، في حين استخدم المنهج التقليدي مع المجموعة الضابطة. وجمعت البيانات من خلال التقييم الشفهي للطلاب، واستبانة تقييم المعلمين وفق مقياس ليكرت وأظهرت النتائج ما يلي:

- حققت المجموعة التجريبية التي استخدمت برنامج (تحدث ألسا) متوسطًا أعلى بشكل ملحوظ حيث بلغ (9.31)، مقارنةً بمتوسط أداء المجموعة الضابطة الذي بلغ (2.12)، وقد تم التأكد من هذه النتائج باختبار (ت) للاستقلال، مما يشير إلى تفوق المجموعة التجريبية.
- أفاد القيمين بتحقيق مستوى أعلى من التفاعل بين الطلاب في المجموعة التجريبية، مشيرين إلى أن برنامج (تحدث ألسا) يؤدي دورًا هامًا في توفير تدريب فردي.
- إن الأدوات المدعومة بالذكاء الاصطناعي، مثل: (تحدث ألسا) قادرة على إثراء أساليب التدريس التقليدية في تعليم اللغات بشكل فعال.

الكلمات مفتاحية: الذكاء الاصطناعي - التعلم عبر الأجهزة المحمولة - تدريس اللغة التواصلية.



DOI: 10.12816/0062615

(*) Corresponding Author:

Ashwag A. Almethen
Department of Curriculum and Instruction - College of Education
- King Faisal University - Al-ahsa - Saudi Arabia
Email: asmethen@kfu.edu.sa

(*) للمراسلة:

أشواق أحمد المذن
قسم المناهج وطرق التدريس - كلية التربية
- جامعة الملك فيصل - الأحساء - المملكة
العربية السعودية
البريد الإلكتروني:
asmethen@kfu.edu.sa

1. Introduction

Due to factors such as globalization, the ability to communicate effectively in spoken English has become a primary goal among second and foreign language learners in education. Aside from grammatical accuracy and good lexical knowledge, pronunciation and speaking intelligibly are key to successful communication. According to Pennington (2021), individuals in leadership positions who struggle with pronunciation could experience breakdowns in communication, diminished confidence, and limited participation in spoken engagements. This could even occur in cases where their language proficiency is above average. This indicates the central value of oral communication skills within communicative approaches to teaching the English language.

Despite its emphasis, pronunciation instruction remains among the most challenging elements of language teaching to effectively implement in formal classroom settings. Barriers such as limited instruction time, large classroom sizes, and competing priorities in the curriculum tend to restrict opportunities for important aspects such as individualized speaking practice and detailed feedback that can positively contribute to English learning. In an attempt to mitigate these limitations, mobile-assisted language learning (MALL) applications continue to gain prominence as supplementary tools that can further extend pronunciation practice beyond the classroom setting (Xuhong et al., 2025). Among these, ELSA Speak (English Language Speech Assistant) has especially gained significant attention due to its use of artificial intelligence in education to provide detailed, real-time pronunciation feedback (Sholekhah & Fakhurriana, 2023). Nevertheless, the extent to which this tool produces measurable improvements in pronunciation and speaking performance, particularly when compared to traditional approaches, is insufficiently explored. This study aimed to fill this gap and contribute to the existing limited literature.

2. Background and Context

Communicative competence has become a defining concept in contemporary English language teaching. Rather than solely focusing on grammatical knowledge, this competence emphasizes the ability to employ language effectively and appropriately in real-life scenarios (Elmahdi et al., 2025). Within this

framework, pronunciation has a key role, given that intelligible speech is a prerequisite for successful oral communication. In cases where pronunciation errors affect a listener's comprehension, communication is likely to fail or be delayed despite the speaker achieving grammatical accuracy. This highlights the significance of proper pronunciation and speaking skills.

Despite the importance of the two, studies have consistently found that pronunciation instruction tends to be marginalized in classroom practice. This could result from various factors, including insufficient training in pronunciation pedagogy, with pronunciation often being addressed implicitly rather than through systematic instruction (Pennington, 2021). According to studies conducted in diverse educational settings, learners tend to receive limited corrective feedback on key areas such as pronunciation, which enables persistent errors that may become more permanent over time (Hoque et al., 2024; Moyi & Galadima, 2020). These issues are especially pronounced in contexts where English is taught as a foreign language and opportunities for authentic spoken interaction are limited.

Advances in technology have increased interest in MALL as a means of addressing existing instructional gaps. Consequently, MALL evolved from earlier computer-assisted language learning approaches with an emphasis on portability, learner autonomy, and increased exposure to the target language (Yaman & Ekmekçi, 2016). Today, increasing numbers of pronunciation applications are employing speech recognition and artificial intelligence in education to provide more immediate and individualized feedback that cannot be easily replicated in traditional classrooms.

ELSA Speak, one of the most recognizable applications, is specifically designed to support pronunciation development among learners. It can analyze learners' speech at multiple levels, including areas such as the individual sound, sentence-level pronunciation, and emphasis, and consequently provide detailed feedback that can help improve accuracy and intelligibility (Nushi & Sadeghi, 2021). Given that a learner can practice independently, at their own pace, and receive instant feedback, the application has been promoted as valuable to classroom instruction but not as a replacement for it. Nevertheless, it is important to

evaluate the educational value of this technology through empirical research, exploring the actual learning outcomes and not merely the learner's perceptions.

2.1 Problem Statement

Although MALL applications are increasingly finding their place in language education, especially with the advent of and widespread use of artificial intelligence, limited empirical evidence indicates the effectiveness of the application as a supplementary teaching tool in traditional classroom settings where formal instruction is employed. Where studies report positive learner attitudes towards these applications, the findings tend to be based on self-reported data or short-term interventions. This raises concerns regarding the reliability of reported outcomes.

Some of the existing meta-analytical and review studies on MALL identify persistent methodological concerns that directly complicate how the existing research findings are interpreted. For instance, Mihaylova et al. (2022) note that, while their study reports positive effects of MALL applications, their findings "should be treated with caution due to risks of high bias and low quality of evidence" (p. 265). Due to such limitations, including others such as inconsistent research designs, variable measures of assessment, and insufficient control groups, it becomes difficult to determine whether observed and reported improvements in pronunciation or speaking skills can be reliably attributed to the use of mobile applications. Claims about the effectiveness of pronunciation applications such as ELSA Speak must therefore be interpreted carefully and supported by more rigorous research.

Several studies have suggested that ELSA Speak can improve learners' pronunciation accuracy and speaking confidence (Sa'diyah et al., 2025). However, much of the existing research lacks rigorous comparative designs that evaluate the tool against traditional classroom instruction. Moreover, few quantitative studies conducted in secondary and tertiary educational settings systematically measure pronunciation and speaking outcomes. This has led to a need for research examining whether the tool provides notable benefits when used as a supplementary teaching tool in the secondary context.

2.2 Research Questions

1. To what extent does the use of ELSA Speak improve participants' English pronunciation and speaking skills?
2. How does the performance of participants using ELSA Speak compare with that of those receiving traditional classroom instruction only?

2.3 Research Objectives

The main objectives of this study are as follows:

- To evaluate the effectiveness of ELSA Speak in improving learners' English pronunciation and speaking performance.
- To compare pronunciation and speaking outcomes between learners using ELSA Speak as a supplementary tool and those relying solely on traditional classroom instruction.

2.4 Significance of the Study

The study findings contribute empirical evidence to ongoing discussions and literature on the pedagogical value of AI-driven pronunciation tools in language education. With a focus on measurable pronunciation and speaking outcomes rather than perceptions alone, the study aimed to clarify whether ELSA Speak offers actual benefits in formal instructional contexts. The findings could enable educators, curriculum designers, and institutions to make informed decisions about integrating mobile-assisted pronunciation tools into English language programs. By addressing identified gaps in the literature, the study could form a good foundation for future research on technology-supported pronunciation instruction.

2.5 Literature Review

2.5.1 Theoretical Framework

- Communicative Language Teaching (CLT)

CLT is an approach that largely focuses on meaningful interaction as not only a goal but also a means of learning language. Pronunciation, in this case, is viewed as the capacity to produce intelligible and comprehensive language (Pennington, 2021). The instruction employed to support intelligibility is in alignment with CLT principles, enabling learners to participate in communicative activities more effectively. Tools such as ELSA Speak can support CLT by enabling the learner to address

their individual pronunciation challenges outside the classroom setting, enhancing their readiness for communicative interaction.

- Technology Acceptance Model (TAM)

The TAM is a framework that demonstrates how and why users adopt technology. As Dziak (2024) explains, perceived usefulness and perceived ease of use are among the most important factors in this case. In language learning contexts, students, as well as teachers, are likely to be drawn to a tool when they believe it will enhance their performance and when the tool is accessible and user-friendly. According to research where the framework has been applied to educational technologies, positive perceptions have been shown to enhance engagement and the frequency of practice. However, this is not necessarily guaranteed regarding learning effectiveness. Thus, it is important to conduct an empirical evaluation of the learning outcomes.

- MALL

MALL evolved from computer-assisted language learning (CALL), with an emphasis on mobility, flexibility, and learner autonomy (Yaman & Ekmekçi, 2016). Unlike CALL, MALL's strength lies in its enabling the learner to practice their language beyond the classroom setting. It thus enhances exposure by increasing opportunities for repetition. By providing immediate feedback and enabling individualized practice, MALL creates an excellent environment for students to develop their pronunciation skills. Although many studies have found the approach to produce positive outcomes, the majority of these studies, as Abduh (2019) states, focused on "reading skills, writing, grammar, and vocabulary learning. Very few studies have examined the effect of MALL on pronunciation" (p. 30).

- Empirical Studies on AI-Driven Pronunciation Tools and ELSA Speak

Research on AI-driven pronunciation tools indicates potential benefits for pronunciation accuracy and speaking performance. Studies examining applications such as Speechace report improvements in areas such as learners' pronunciation competency, especially in cases where feedback tends to be detailed and immediate (Jingjing & Andi, 2025; Nguyen et al., 2025). However, variability in effectiveness has been

reported depending on the learner's proficiency levels and instructional contexts.

Regarding ELSA Speak, studies indicate a positive impact on the learner's pronunciation accuracy as well as their speaking skills. Specifically, classroom-based research and case studies have noted an improvement in the learner's speaking performance as well as an increase in confidence when the tool has a supplementary role (Sa'diyah et al., 2025). However, critical evaluations of the application have identified limitations such as potential inaccuracies in speech recognition as well as an emphasis on native-like pronunciation norms, which do not always align with intelligibility-focused pedagogy (Nushi & Sadeghi, 2021).

Motivation has been identified as a key factor influencing the development of pronunciation. Based on studies on pronunciation learning consistency, findings have reported a relationship between learner motivation and achievement in pronunciation, indicating that tools that help enhance engagement and practice frequency can indirectly enhance pronunciation (Stoughton & Kang, 2024). However, motivation alone is insufficient without effective instructional design and reliable feedback mechanisms.

Systematic reviews of mobile-assisted pronunciation research generally highlight the need for more rigorous quantitative studies that specifically examine production outcomes rather than only perceptions (Stoughton & Kang, 2024). These reviews consistently identify a lack of controlled comparative research, especially in formal education contexts.

3. Methodology

3.1 Research Design

For this study, a quantitative quasiexperimental pretest-posttest design was employed to investigate the impact of AI-integrated tools on proficiency. The approach was ideal for capturing the professional observations of educators regarding student progress. Specifically, the study compared two pedagogical environments, one where the evaluators employed ELSA Speak as a supplementary tool and the other where the evaluators only used the traditional instructional methods. The first was the experimental condition, whereas the second was the control.

3.2 Participants and Sampling

The study participants included 44 secondary students enrolled in English language classes. Convenience sampling was used to select these participants for accessibility and instructional considerations. The group was divided into two smaller groups: the experimental ($n = 22$) and control ($n = 22$) groups.

ELSA Speak was used by those in the experimental group as a supplementary pronunciation tool alongside traditional classroom instruction, while those in the control group only received traditional classroom instruction.

Baseline equivalence between the two groups was ensured by having them complete a pronunciation and speaking pretest before intervention. The test results indicated comparable proficiency levels before the intervention period.

Aside from the student participants, English language evaluators also participated in the study. The evaluators had two key roles: supervising classroom implementation and completing perception questionnaires about engagement among participants and the usefulness of ELSA Speak as an instructional tool. The evaluators were therefore not regarded as the primary analysis group for the study.

3.3 Research Setting

This study was conducted within the secondary-level English language classroom, particularly where ELSA Speak was employed. The intervention was planned well to fit into the curriculum over 4 to 6 weeks. During this period, evaluators in the experimental group facilitated the use of ELSA Speak as a supplementary tool.

3.4 Data Collection Instruments

To satisfy the pretest and posttest requirements, participants completed speaking assessments before and after the intervention. Key areas evaluated during the assessment included fluency in language, accuracy in pronunciation, intelligibility, and overall speaking performance.

The participants' responses were scored using a structured analytic rubric adapted from communicative pronunciation assessment criteria typically employed in English language instruction. Scores were assigned on a 5-point scale, ranging from 1 (poor) to 5 (excellent) across four key

categories of assessment. These included fluency, accuracy in pronunciation, intelligibility, and overall spoken performance. To score 5 on intelligibility, for instance, the speaker's speech was required to be easy to understand, even if the participant did not have a native accent. The speaking performances were evaluated independently by two English language evaluators to maintain consistency in scoring. Additionally, a Cohen's kappa coefficient was calculated to establish interrater reliability.

Finally, a structured consensus discussion was used to resolve minor differences in the scores. In a scenario where the evaluators differed by a single point on the rubric scale, they reexamined the audio recording in question together against the descriptors in the rubric until they reached a unanimous final score. Where discrepancies exceeded a single point, an average of the two scores was used to maintain consistency in evaluation.

A review by English language instructors familiar with the assessment practices was conducted to establish the content validity of the assessment instrument.

3.5 Procedure

The study was conducted in four main phases, including the baseline evaluation intervention, exit evaluation, and perception survey. During the baseline evaluation, the evaluators in the two groups administered and scored the speaking pretest for the participants. This helped establish a baseline of the current performances. During the intervention, which took 4–6 weeks, the evaluators in the experimental group employed ELSA Speak in the learning routines of their participants, while the control group continued to receive the conventional instructions. In the exit evaluation phase, posttests of the participants' speaking skills were conducted to document any shifts in performance. Finally, following the posttest, during the perception survey, the experimental group completed the perception questionnaire. This provided their account of the tool's impact on them.

3.6 Intervention Procedure

For the experimental group, the intervention phase lasted 4 weeks, with participants using ELSA Speak as a supplementary pronunciation tool alongside the traditional classroom instruction. During this period, the participants engaged with

the tool an average of three times per week. Each session lasted 20–30 minutes, depending on the day's activities.

Key activities completed using ELSA Speak included word pronunciation exercises, stress and intonation exercises, sentence repetition, and guided speaking tasks, with the tool providing automated feedback. Specifically, ELSA Speak provided immediate corrective feedback in the areas of accuracy in pronunciation and intelligibility.

During the intervention phase, the evaluators played a vital role in supervising the implementation process, monitoring participation among the participants, and ensuring all the participants completed their assigned tasks. However, the application was largely responsible for providing feedback during pronunciation tasks.

Unlike those in the experimental group, students in the control group only received regular instruction. Thus, they did not use ELSA Speak. The main activities for this group included textbook-based speaking tasks, oral exercises in the classroom, and teacher-led pronunciation correction.

3.7 Data Analysis

The data collected were analyzed using statistical measures. This helped determine the impact of the intervention. Descriptive statistics were used to calculate the test scores and evaluate the perception survey. However, a paired sample t-test was used to measure growth within the student groups of the two groups of evaluators. An independent sample t-test was used to compare the performance gains between the experimental and control groups. This helped determine whether the evaluators observed a statistically significant improvement in participants using ELSA Speak compared to those who did not.

3.8 Instruments

3.8.1 Data Analysis Tools

- The quantitative data were analyzed using R (Base Software 4-series). Descriptive statistics were calculated to summarize the pre and posttest scores as well as responses to the teacher questionnaires.
- Paired sample t-tests were used to compare within-group improvements, while

independent sample t-tests were used to compare performances between the two groups.

- Statistical significance was set at $p < .05$.
- The software (R) was used to analyze survey responses.

3.8.2 Data Storage

1. All collected data were arranged in MS Excel before being exported for statistical analysis.
2. Participants were assigned unique identifiers (e.g., C01, C02, E01, E02...). This helped ensure anonymity.
3. The digital files were stored in a password-protected PC, encrypted if necessary, backed up on a secure external drive, and only accessible to the principal researcher.

3.8.3 Data Retention and Disposal

The data will be stored for a year in case they need to be retrieved for important reasons. After this period, if they are no longer required, they will be permanently deleted and any printed copies shredded.

3.8.4 Ethical Standards

The data collection process and analysis procedures complied with the institutional research ethics and international standards regarding privacy, confidentiality, and data handling.

3.8.5 Ethical Considerations

The study strictly adhered to ethical standards. All the participants provided informed consent, and they were assured that their classroom data and personal identities were not collected. Participation was voluntary, and evaluators could withdraw their data from the study at any time.

4. Results

4.1 Descriptive Statistics (Under Traditional Instructional Conditions)

Table 1

Group	Pretest	posttest
Control	59.50	61.62
Experimental	59.95	69.26

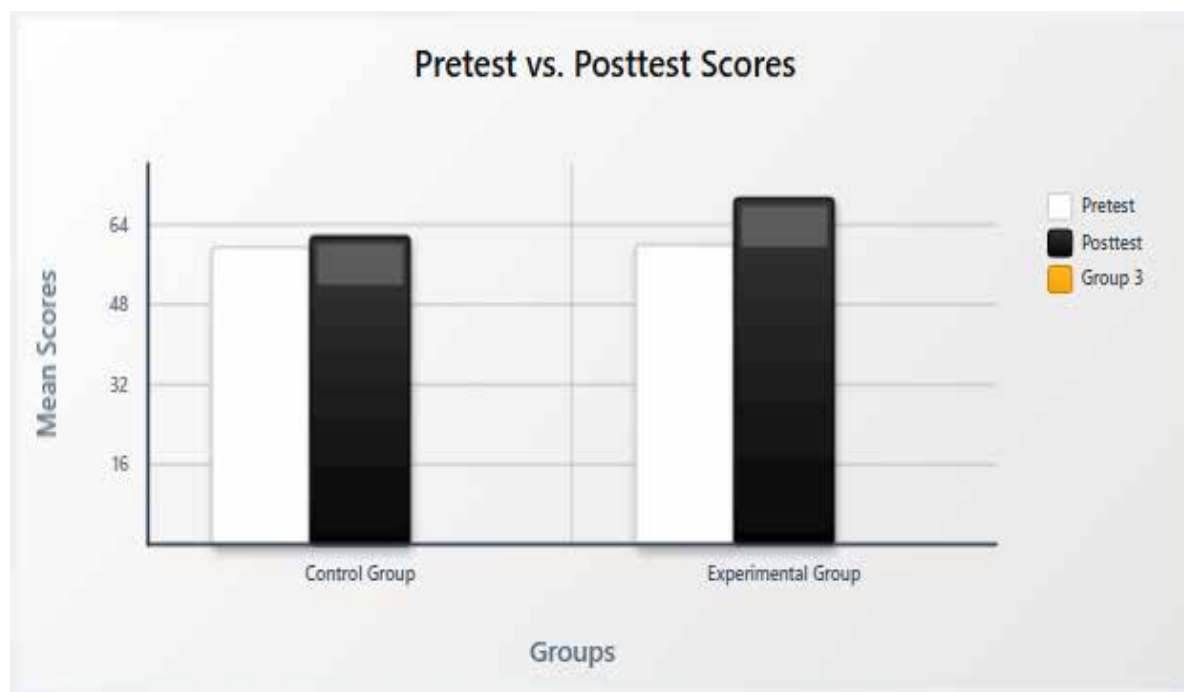


Figure 1

The pretest scores obtained for the experimental and control groups were comparable before the intervention. This was a good indicator of relative equivalence in baseline pronunciation and proficiency in speaking. Based on teacher evaluations, the participants in the control group ($n = 22$) recorded a pretest mean score of 59.50 (SD: 4.89) and a posttest mean score of 61.62 (SD: 4.37). The mean gain was 2.12 points (SD: 5.80). The results demonstrated marginal improvement in pronunciation and speaking performance among these participants under traditional instruction.

4.2 Descriptive Statistics (ELSA Speak Supplementation)

For the experimental group ($n = 22$), the teacher-assigned scores yielded a mean score of 59.95 (SD: 4.47) for the pretest and 69.26 (SD: 6.01) for the posttest, resulting in a significantly higher mean gain of 9.31 points (SD: 7.61). As reported by the evaluators involved in the implementation and use of ELSA Speak, this indicated a significant improvement in pronunciation and speaking performance among this group of participants.

4.2.1 The baseline scores were largely comparable for the two groups.

Table 2: Teacher Questionnaire

Questionnaire Item	Mean	SD
ELSA Speak enhanced student engagement	4.45	0.59
participants practiced pronunciation more frequently	4.61	0.50
The application was easy to integrate	4.18	0.73
participants demonstrated greater speaking confidence	4.39	0.66

The evaluators generally reported positive perceptions regarding the implementation of ELSA Speak. Among the items, the best rated was an increase in pronunciation practice opportunities. This strongly indicated that the evaluators viewed the application as an effective tool in extending speaking practice beyond the limitations of the classroom.

4.3 Comparison Within the Groups

A paired sample t-test was conducted to determine whether the teacher-reported evaluations of the pronunciation and speaking performance improved among participants within each of the two groups.

For the control group, the difference between the pretest and posttest scores [$t(21) = 1.72$, $p = .101$] was not statistically significant. Conversely, the paired sample t-test for the experimental group revealed significant improvement between the two tests [$t(21) = 5.74$, $p < .001$]. Whereas evaluators observed substantial gains among participants when using ELSA Speak, the difference between the pre and posttest scores was not significant when using only traditional instruction.

4.3.1 Comparison Between the Two Groups

An independent sample t-test was used to examine teacher-reported performance differences between the two groups. The test revealed a statistically significant difference between the

experimental ($M = 9.31$, $SD = 7.61$) and control ($M = 2.12$, $SD = 5.80$) groups, $t(39.24) = 3.53$, $p = .001$, Cohen's $d = -1.06$. This indicated a large effect size. The 95% confidence interval for the mean difference ranged from 3.07 to 11.31.

The findings indicate that students using the application demonstrated significantly greater improvement in speaking performance and pronunciation compared to those who only received traditional instruction.

4.3.2 Evaluators' Questionnaire Results

Evaluators in the experimental group also completed a Likert-scale questionnaire. This provided contextual insights into the implementation process in the classrooms. The evaluators generally reported positive perceptions regarding student engagement as well as the integration of ELSA Speak into the classroom routines. Additionally, they reported an increase in opportunities for individualized pronunciation practice as well as a notable increase in student engagement in using spoken English outside the classroom setting.

Table 3

Feature	Control Group (n=22)	Experimental group (n=22)
Instructional method(s)	Regular classroom instruction only (textbook exercises, oral tasks, teacher-led correction)	Regular classroom instruction paired with ELAS Speak application
Intervention	4 weeks of standard classroom instruction	4 weeks of practicing three times a week for 20 to 30 minutes per session.
Pretest mean score	59.50 (SD: 4.89)	59.95 (SD: 4.47)
Posttest mean score	61.62 (SD: 4.37)	69.26 (SD: 6.01)
Mean gain in performance	2.12 points (SD: 5.80)	9.31 points (SD: 7.61)
Within-group improvement Paired t-test	$t(21) = 1.72$, $p = .101$ Not statistically significant	$t(21) = 5.74$, $p < .001$ Very statistically significant
Between-group comparison Independent t-test	---	$t(39.24) = 3.53$, $p = .001$ Statistically significant superiority
Effective size (Cohen's d)	---	1.06 (Large effect size)
95 percent confidence interval	---	3.07-11.31
Evaluator observations	Marginal improvement in pronunciation and speaking skill.	Significant performance gains, higher engagement among students, increased individualized practice.

5. Discussion

The first guiding question of the study focused on the extent to which ELSA Speak improves student English pronunciation as well as speaking skills, as evaluated by the evaluators. Following analysis of the teacher-assigned tests, the results show a statistically significant improvement in the group using the tool. Specifically, the evaluators observed meaningful gains where participants engaged with ELSA Speak. However, evaluators in the control group did not report significant gains in these areas. The results indicate that AI-powered tools are vital in enhancing pronunciation and speaking performances and should be considered if educators wish to boost performances in these areas.

The CLT lens can be used to meaningfully interpret these findings. The framework emphasizes intelligibility as well as effective communication over native-like accuracy. In this case, pronunciation is viewed as a functional component of communicative competence (Pennington, 2021). The gains observed among participants in the experimental group indicate that the AI-powered tool was successful in enhancing the development of intelligible spoken output by enabling them to repeatedly practice targeted pronunciation features as they received timely individualized feedback. These opportunities closely align with the principles of CLT, especially when addressing difficulties in pronunciation that can be easily overlooked or insufficiently addressed in communicative class settings due to limitations such as time constraints. Therefore, ELSA Speak appears to have a complementary role in communication instruction. This interpretation is especially important within the context of Saudi English foreign language (EFL), where instruction in pronunciation has been shown to be challenging due to limitations such as classroom time constraints, large class sizes, and inconsistent teacher preparation in pronunciation pedagogy. According to a study by Algethami and Al Kamli (2025), practices in teaching pronunciation in Saudi classrooms tend to face practical constraints that can limit individualized corrective feedback. Applications such as ELSA Speak can mitigate the issue by serving as supplementary tools, enabling learners to effectively and freely engage in repeated, individualized pronunciation practice outside the classroom. This is particularly beneficial for continued learning at the learners' convenience. This

further reinforces the role of such tools in enhancing communicative readiness.

From the perspective of MALL, the study's results support the concept that mobile technologies can help enhance and extend learning beyond the classroom environment. They thus increase the students' exposure to the target language (Yaman & Ekmekçi, 2016). Evaluators' reports of significant improvement in the experimental group indicated that aspects such as flexibility and immediacy afforded by mobile-based pronunciation practice can positively contribute to enhanced skills in speaking and communicating. Although much MALL research has focused on learners' attitudes and perceptions, pronunciation has not been sufficiently explored (Abduh, 2019). This study addresses this gap by demonstrating measurable improvements in pronunciation and speaking performance and thus evidences the effectiveness of MALL in this area.

The findings also align with research conducted in the Saudi Arabian EFL context, which highlights the growing role of MALL in elevating the acquisition of language. For instance, a study by Alnujaidi (2021) reported increased interest among Saudi EFL teachers toward adopting MALL. Despite concerns about implementation and classrooms, this did not appear to significantly impact interest. However, Keezhatta and Omar (2019) found that mobile-assisted learning significantly enhanced language-related performance among secondary school learners in Saudi Arabia. The two studies represent the teachers' perspective and the impact on students. While the teachers had a favorable perspective, students benefited from the technology. This reinforces the view that mobile technologies can positively supplement traditional instruction by creating an opportunity for increased practice and learner engagement outside the classroom.

Regarding the second question, comparing the performance of the two groups, those using teacher-implemented ELSA Speak and those who only received traditional instruction, the independent sample t-test showed that the former achieved significantly more gains compared to the control group. This reinforces the conclusion that ELSA Speak can have a positive impact when used as a supplementary instruction tool. In the context of Saudi Arabian EFL, recent studies indicate the growing value of artificial intelligence in language

learning. Alqaed (2024) found that students in the country generally had a positive perspective on AI-supported language learning, largely due to the immediate feedback provided as well as the opportunity for individualized learning. Additionally, participants noted that the AI applications increased engagement while also supporting the development of language across various skills. This aligns well with the present findings, where evaluators reported significant improvements in areas of speaking and pronunciation where ELSA Speak was used. The findings are consistent with several empirical studies suggesting that AI-driven applications for pronunciation can help improve accuracy and performance in speaking through immediate as well as detailed feedback (Jingjing & Andi, 2025; Nguyen et al., 2025).

The study findings are also consistent with classroom-based research on ELSA Speak, which has reported improvements in speaking performance and confidence among learners after using the application alongside formal instruction (Sa'diyah et al., 2025). Ngan et al. (2024) note that "students using ELSA Speak had noticeable improvements in pronunciation and communication confidence" (p. 7592).

The findings obtained through the evaluators' questionnaire align with interpretations drawn from the TAM. Although the model does not directly predict the learning outcomes, positive perceptions of the usefulness of ELSA Speak as well as its ease of use are good indicators of conditions conducive to sustained engagement (Dziak, 2024). Through increased levels of engagement and higher frequencies of practice, the tool could have indirectly contributed to the gains in performance observed in the experimental group. By pairing the evaluators' data with the objective measures of performance, the study strengthens TAM-based interpretations by showing that the favorable perceptions coincided with measurable improvements in the outcomes rather than solely relying on attitudinal data. These interpretations are also supported by more recent research exploring technology acceptance in higher education institutions in the Arab world. As Abu-Tineh et al. (2024) explain, the perceived usefulness and ease of use continue to be important predictors of acceptance within these educational contexts, especially when digital tools contribute to accessibility, learner autonomy, and flexibility. The

generally positive perception reported by evaluators of the present study is therefore a broader reflection of regional trends concerning the integration of educational technologies into language instruction.

Although the results of the study indicate significant advantages for the experimental group, evidenced by the high mean gain and substantial effect size, it is important to interpret the findings alongside the limitations of the study's design. Firstly, using convenience sampling to draw participants from a specific secondary school means that the cohort, which comprised 44 students, did not fully represent diversity in language proficiency profiles. This is especially true when considering the wider Saudi EFL landscape. The lack of true randomization could introduce a risk of selection bias, meaning that the differences in gains between the two groups could have been influenced by baseline variables not measured, including the varying degrees of digital literacy and prior individual motivation.

Additionally, the time constraint of a 4-week period posed a potential constraint on the interpretation of long-term language acquisition. For more permanent habituation, pronunciation and automaticity regarding speaking skills require extended, sustained reinforcement. Therefore, the significant improvement observed in the experimental group could have resulted from the novelty effect, a temporary surge in participant engagement due to the introduction of a novel application. This becomes even more likely when considering the 4-week cycle. Without longitudinal follow-up assessments, it is unclear whether the gains will eventually plateau or regress. Acknowledging these limitations does not aim to diminish the value of the observed large effect size. Instead, it highlights the importance of viewing the quantitative gains as a promising short-term indicator.

Based on the findings, incorporating ELSA Speak as a supplementary tool can significantly benefit pronunciation and speaking instruction in the classroom setting. Although the traditional approach is associated with some positive gains in these areas, the gains are even more significant when using this tool. Using such tools, educators would be better able to support individualized practice, particularly in cases where the size of the class and time constraints are limiting factors. Perhaps more importantly, the findings also suggest that such tools

can enable teachers to ensure that their students are understood without pushing them to sound like native speakers. This prioritizes the learner's need to communicate clearly.

5.1 Limitations and Future Research

A major limitation of the study is the use of convenience sampling. This approach limits generalizability due to the small number of participants. This limitation is further compounded by the intervention period being relatively short. The findings should therefore be interpreted cautiously when generalizing to broader contexts in education. Future studies could consider randomized designs, more diverse raters, and a longer intervention period to mitigate these limitations. Studies could also examine how learners' proficiency levels or use patterns impact pronunciation outcomes.

6. Conclusion

This study aimed to evaluate the effectiveness of AI-driven applications such as ELSA Speak as supplementary tools in traditional English language classrooms. The results showed that the tool had significant instructional value when used alongside the traditional approaches. However, only using the traditional approach resulted in marginal improvements (mean gain 2.12, $p = .101$). This was statistically nonsignificant, especially compared to the results obtained following the integration of ELSA Speak (mean gain 9.31, $p < .001$). These improvements were also well-complemented by the exceptional evaluator-reported student engagement. The findings fill an important gap in MALL literature, demonstrating that when artificial intelligence tools are systematically used to extend the boundaries of learning, they are not merely novelties but result in actual, functional speech improvements.

The findings also contribute to newer discussions regarding the integration of AI and mobile-assisted technologies in English language education in not only Saudi Arabia but also the broader Arab world. Based on more recent regional research, the value of AI-supported and mobile-based learning environments has been emphasized regarding promising engagement among learners, autonomy, and competence in communication. The present study contributes to the body of existing literature by providing empirical evidence that these applications can positively influence measurable speaking and pronunciation outcomes,

especially when incorporated alongside traditional instructional approaches.

7. Recommendations

- For teachers, structured microlearning homework tasks are recommended. Rather than attempting to cram digital practice into already tight classroom schedules, instructors might benefit from leveraging ELSA Speak as a structured homework supplement. This could involve assigning microsessions, about 20-minute sessions, three times per week, to match current textbook topics and themes. This would help bypass constraints resulting from crowded classrooms and limited time.
- Secondly, teachers could use detailed feedback reports generated by the AI tools to focus on errors causing breakdowns in communication. By using approaches such as word stress and key vowel sounds, they might yield better outcomes compared to pressuring learners to sound more like natives, particularly regarding desired outcomes for CLT.
- For policymakers and curriculum designers, blended curriculum integration is recommended. AI-powered pronunciation tools have a valuable role, as the findings indicated, and should be embedded into the official English language curriculum framework. This means moving beyond treating these applications as optional and instead clearly listing them in the curriculum guides as verified supplementary elements. To ensure better outcomes, authorities should also invest in teacher-training workshops so that teachers have the required skills to be fully involved.

8. References

- Abduh, M. Y. M. (2019). The effect of implementing MALL applications on learning pronunciation of English by EFL learners at Najran University. *International Journal of Linguistics*, 11(6), 29–44.
- Algethami, G., & Al Kamli, H. (2025). Pronunciation teaching in English language classrooms in Saudi Arabia: Teachers' cognition and practice. *SAGE Open*, 15(1), 1-12. <https://doi.org/10.1177/21582440251321299>
- Alnujaidi, S. (2021). Adoption of mobile assisted

- language learning (MALL) in Saudi Arabian EFL classrooms. *Journal of Language Teaching and Research*, 12(2), 312–323. <https://doi.org/10.17507/jltr.1202.13>
- Alqaed, M. A. (2024). AI in English language learning: Saudi learners' perspectives and usage. *Arab World English Journal*, 15(4), 95–110.
- Abu-Tineh, A. M., Alazaizeh, M., Ghamrawi, N., Alshaboul, Y., & Shal, T. (2024). Culture and technology acceptance model at Arab higher education institutions. *Education and Information Technologies*, 1-15.
- Dziak, M. (2024). Technology acceptance model (TAM). EBSCO Research Starters. Retrieved from EBSCO Research Starters
- Elmahdi, O. E. H., AbdAlgane, M., Othman, K., Alhaj, M. K. M., Ibrahim, I. Z. A., Abdelrady, A. H., & Ahmed, O. A. M. (2025). Improving communicative competence and speaking skills in EFL/ESL classrooms: Strategies, techniques, and pedagogical approaches. *Forum for Linguistic Studies*, 7(12), 484–497. <https://doi.org/10.30564/fls.v7i12.12116>
- Hoque, M. S., Oli, L., Thamarai Selvi, M. D., Chowdhury, A. S. M. M. H., & Reza, M. M. (2024). Untangling the tongues: Exploring pronunciation development strategies for learners of English at high schools in Oromia region of Ethiopia. *World Journal of English Language*, 14(5), 257.
- Jingjing, L. Y., & Andi, H. K. (2025). Effectiveness of AI-driven tools in improving English pronunciation training. *Educational Research and Reviews*, 7(1), 56–58.
- Keezhatta, M. S., & Omar, A. (2019). Enhancing reading skills for Saudi secondary school students through mobile assisted language learning (MALL): An experimental study. *International Journal of English Linguistics*, 9(1), 437–447. <https://doi.org/10.5539/ijel.v9n1p437>
- Mihaylova, M., Gorin, S., Reber, T. P., & Rothen, N. (2022). A meta-analysis on mobile-assisted language learning applications: Benefits and risks. *Psychologica Belgica*, 62(1), 252–271.
- Moyi, M. I., & Galadima, S. (2020). Enhancing communicative competence in English through conscious efforts and good reading culture: A study of some universities in Northern Nigeria. *East African Scholars Journal of Education, Humanities and Literature*, 3(10), 525–540.
- Nguyen, N. V., Vo, T. T., & Tran, V. D. T. (2025). AI-driven pronunciation assessment: The impact of SpeechAce on EFL learners' pronunciation competency. *Computer-Assisted Language Learning Electronic Journal*, 26(3), 84–106.
- Nushi, M., & Sadeghi, M. (2021). A critical review of ELSA Speak: A pronunciation app. *Computer-Assisted Language Learning Electronic Journal*, 22(3), 287–302.
- Ngân, P. T., Thủy, L. T., Nguyệt, N. T., Hân, T. P., Huệ, P. T. (2024). The application of Elsa speak software in English teaching at can tho university of technology – a case study. *International Journal of Social Science and Human Research*, 7(10), 7592-7598.
- Pennington, M. C. (2021). Teaching pronunciation: The state of the art 2021. *Relc Journal*, 52(1), 3-21.
- Sa'diyah, L., Bahtiar, Y., Nur, L. C. N., Afidah, N., & Ma'arif, I. B. (2025). The effectiveness of ELSA Speak in improving speaking skills in extracurricular classes at MA Al-Bairuny. *APPLICATION: Applied Science in Learning Research*, 5(1), 27–35.
- Sholekhah, M. F., & Fakhrurriana, R. (2023). The use of ELSA Speak as a mobile-assisted language learning (MALL) tool toward EFL students' pronunciation improvement. *Journal of Education, Language Innovation, and Applied Linguistics*, 2(2), 93–100.
- Stoughton, A. M., & Kang, O. (2024). A systematic review of empirical mobile-assisted pronunciation studies through a perception–production lens. *Languages*, 9(7), 1–15.
- Xuhong, H., Sazalli, N., & Puteh, M. B. (2025). The role of mobile applications in enhancing English speaking proficiency: A systematic review. *International Journal of Academic Research in Business and Social Sciences*, 15(5), 907–922.
- Yaman, İ., & Ekmekçi, E. (2016). A shift from CALL to MALL?. *Participatory Educational Research*, 4(2), 25–32.