IREELL
Indonesian Review of English Education, Linguistics, and Literature
P-ISSN: 3025-2369 I E-ISSN: 3021-8101
https://jurnalfaktarbiyah.iainkediri.ac.id/index.php/ireel/index

The Effectiveness of ELSA Speak in Facilitating Students' Pronunciation in Storytelling

Hesa Nizar Arosy
MI Muhammadiyah 1 Pare

Nizararosy@gmail.com

ABSTRACT

Pronunciation is an essential component of oral communication, especially in English as Foreign Language (EFL) storytelling. However, many junior high school students face difficulties with clarity, intonation, and fluency. This study aimed to investigate the effectiveness of the ELSA Speak application in improving students' English pronunciation in storytelling. A quantitative quasi-experimental design was applied, involving 71 eighth-grade students of SMPN 1 Kunjang, divided into two groups: Class 8F as the experimental group using ELSA Speak and Class 8G as the control group taught through conventional methods. Both pre-test and post-test were administered to measure pronunciation performance, and the data were analyzed using Analysis of Covariance (ANCOVA). The results revealed that students in the experimental group achieved significantly better pronunciation improvement compared to the control group. The ANCOVA analysis confirmed a significant effect of ELSA Speak on students' pronunciation ($p < .001$). The application's real-time and individualized feedback mechanism enabled students to detect and correct pronunciation errors instantly, which fostered autonomous learning and consistent practice. Unlike delayed teacher correction, ELSA Speak provided immediate reinforcement that accelerated accuracy and fluency development. These findings highlight the effectiveness of incorporating AI-based applications into English language learning, particularly for enhancing pronunciation in storytelling contexts.

Keywords: artificial intelligence, ELSA Speak Application, English language learning, pronunciation, storytelling

Copyright: © 2025 by Authors.
This is an open access publication under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International (CC BY SA) license (https://creativecommons.org/licenses/by-sa/4.0/).

Introduction

a. Background and Objective(s)

Mastering a language is a complex process involving multiple components, among which pronunciation plays a central role. Producing correct sounds is not only about articulation but also includes suprasegmentally features such as intonation, stress, and rhythm to convey meaning (Gilakjani, 2016; Pennington & Rogerson-Revell, 2019). Pronunciation strongly affects storytelling performance because fluency, clarity, and prosody determine listeners' understanding of the message and emotions expressed. However, as English is a foreign language in Indonesia, many students struggle to use these features effectively in storytelling tasks. Even grammatically correct sentences may become hard to understand when mispronounced (Nugraha et al., 2022).

Similar pronunciation difficulties among Indonesian EFL learners have also been reported in other recent studies. For instance, Hibatulloh et al. (2024) found that secondary school students often faced problems in pronunciation and fluency despite task-based speaking activities. Likewise, Wulansari and Fakhrurriana (2024) showed that using the Online Oxford Dictionary significantly helped improve learners' pronunciation and vocabulary mastery in EFL contexts. These findings highlight that technology-mediated pronunciation practice is increasingly recognized as a promising pedagogical approach in Indonesia.

Despite its importance, pronunciation often receives limited attention in formal education. It is typically taught as part of speaking activities rather than as a standalone skill. Conventional techniques such as drills, minimal pairs, and tongue twisters can be helpful but are often teacher-centered and lack personalized feedback(Basuki, 2018; Dandee & Pornwiriyakit, 2022). Moreover, due to limited classroom focus on pronunciation and the influence of local accents, teachers may unintentionally shape students' pronunciation according to their own accents (Anes et al., 2023). As a result, many junior high school students struggle with pronunciation in narrative texts and show low oral confidence.

Recent technological advancements, especially Artificial Intelligence (AI), have opened new opportunities for language learning. One example is the AI-powered ELSA Speak application, which provides learners with individualized, immediate, and interactive feedback on pronunciation, stress, and intonation. Previous studies (Adawiah & Muliati, 2024; Kholis,

2021; Lestari et al., 2024) have shown that ELSA Speak can improve pronunciation skills at various educational levels. However, most of these studies have focused on the technical aspects of pronunciation rather than its use in communicative tasks such as storytelling.

Addressing this gap, the present study investigates the effectiveness of the ELSA Speak application in enhancing junior high school students' English pronunciation in storytelling. It aims to contribute not only to pronunciation improvement but also to students' oral performance in narrative contexts, supporting more engaging and effective English language learning.

b. Theoretical Framework/Literature Review

Pronunciation is one of the key elements of oral communication in English as a Foreign Language (EFL). It encompasses both segmental features (individual vowel and consonant sounds) and suprasegmental features such as intonation, stress, rhythm, and phrasing, all of which shape the intelligibility and naturalness of speech (Harmer, 2007). Inaccuracies in these aspects can lead to misunderstanding, even when grammar and vocabulary are correct. The importance of pronunciation is especially evident in storytelling, where clarity and prosody determine whether the message and emotions of a narrative are effectively conveyed.

Theoretical perspectives also underline its importance. Flege's Speech Learning Model (SLM) (1995) posits that learners' second language (L2) pronunciation is influenced by their first language (L1) phonetic system, but with sufficient exposure, new sounds can still be acquired. Meanwhile, Communicative Language Teaching (CLT) emphasizes intelligibility over native-like accuracy (Richards & Rodgers, 1986), highlighting that pronunciation instruction should aim at effective communication rather than imitation of native accents. Integrating pronunciation into communicative tasks such as storytelling aligns with this objective.

Despite its significance, pronunciation often receives limited attention in classroom practice. Traditional approaches such as drills, minimal pairs, and tongue twisters help raise awareness of sound distinctions but rely heavily on teacher input and rarely offer personalized feedback (Basuki, 2018; Dandee & Pornwiriyakit, 2022). Teachers themselves also face challenges due to limited training and the influence of local accents (Karakas, 2019). Consequently, pronunciation tends to be overshadowed by grammar and vocabulary instruction, leaving many students with low oral fluency and confidence.

Recent advancements in Artificial Intelligence (AI) have created new opportunities for pronunciation learning. AI-powered applications, such as ELSA Speak, use Automatic Speech Recognition (ASR) to provide instant, individualized feedback on pronunciation, stress, and intonation. These tools enhance accuracy, foster learner autonomy, and increase motivation through gamification and interactive features (Gusti Bagus et al., 2024). Empirical studies confirm their benefits across various educational levels:) reported significant improvements in fluency and accuracy among high school students, while Al-Shallakh (2024) found notable gains among university learners. (Nurlita et al., 2024) noted that 87% of junior high school students perceived ELSA Speak as motivating and effective. Similarly (Mubarok & Aziez, 2024)showed AI-based tools provide more personalized feedback than non-AI alternatives like Google Translate. Other ASR applications, such as Google Read Along, have also demonstrated positive effects (Abimanto & Sumarsono, 2024)

While prior research highlights the technical effectiveness of AI-based pronunciation tools, few studies have examined their use in communicative or performance-based contexts such as storytelling. Most focus narrowly on segmental accuracy rather than fluency, expressiveness, and oral performance. Furthermore, there is limited research at the junior high school level, where students are at a critical stage in developing pronunciation habits.

Addressing this gap, the present study combines pronunciation and communicative language teaching theories with the practical affordances of AI. It examines how ELSA Speak can enhance both segmental and suprasegmental aspects of pronunciation within the meaningful context of storytelling. By doing so, this research extends prior findings and provides pedagogical insights into integrating technology for more effective EFL instruction.

c. Research Problem

Based on the background and the research gap identified in the literature review, this study focuses on improving students' English pronunciation through the use of ELSA Speak application. Pronunciation learning in many classrooms still relies heavily on conventional teaching methods that often lack immediate and personalized feedback for students. It offers an innovative alternative by providing interactive exercises and automatic evaluations of pronunciation errors. Therefore, this study seeks to determine whether the use of ELSA Speak

application can significantly enhance students' English pronunciation in storytelling compared to traditional teaching methods.

Methods

This study employed a quantitative quasi-experimental design to investigate the effectiveness of the ELSA Speak application in facilitating students' English pronunciation in storytelling. The design involved a pre-test and post-test with two groups: an experimental group that received treatment using ELSA Speak and a control group that was taught through conventional methods. The use of this design enabled the researcher to compare the effectiveness of AI-based pronunciation training against traditional classroom instruction while controlling for initial differences in students' abilities.

The population of the study consisted of eighth-grade students of SMPN 1 Kunjang in the 2024/2025 academic year. From this population, two intact classes were selected as the sample through purposive sampling. Class 8F, comprising 36 students, was assigned as the experimental group, while Class 8G, comprising 35 students, served as the control group. Thus, the total number of participants was 71 students. The two classes were considered comparable in terms of age, language proficiency, and prior exposure to English instruction, making them suitable for the quasi-experimental design. Due to the fixed class schedules and administrative regulations at the research site, random assignment of students was not feasible. Therefore, purposive sampling was used to select two intact classes with similar proficiency levels and demographic characteristics. This approach ensured comparable groups while maintaining ecological validity. Moreover, statistical control (ANCOVA) was applied to adjust for any pre-existing differences between groups.

Instruments

The instrument used in this study was a pronunciation test administered in the form of storytelling tasks. The test required students to narrate a short story orally, and their performance was recorded and evaluated using a scoring rubric focusing on fluency, accuracy, intonation, and stress (Mustadi, as cited in Cahyani, 2021). Both pre-test and post-test were conducted for the experimental and control groups to measure improvements in pronunciation. To ensure reliability, the scoring was validated through inter-rater agreement involving two English teachers.

The data obtained from the pre-test and post-test were analyzed using Analysis of Covariance (ANCOVA) in SPSS 26.0. This statistical technique was applied to determine the significant differences in students' pronunciation performance between the experimental and control groups, while controlling for their pre-test scores. The ANCOVA results provided evidence of whether the treatment using ELSA Speak had a significant effect on students' pronunciation in storytelling. Descriptive statistics, including mean scores and standard deviations, were also presented to illustrate the magnitude of improvement in both groups.

Result and Discussion

Research Findings

The researcher explains the data gathered throughout the research in this chapter. The data provides an answer to the research question. Data was collected from two distinct classes' pre- and post-tests. On April 14th, 2025, the experimental and control groups took part in a pre-test. On May 26th, 2025, the experimental and control groups were given a post-test.

This section presents the pre-test and post-test results of both experimental and control groups, comparing students taught using the ELSA Speak application with those using conventional methods. To determine the effectiveness of the treatment, the researcher used ANCOVA in SPSS version 25 for Windows to analyze the data obtained from both the pre-test and post-test.

1. Description of the Data

The data were taken from students of SMPN 1 Kunjang, samples were taken from 71 students. The sample consisted of two groups, 35 students as the experimental class from VIII F and 36 students as the control class from VIII G.

Each test, both the pre-test and post-test, required students to perform a narrative text reading task. Their pronunciation performance was then assessed using the ELSA Speak application. The following are the results or scores of the pre-test and post-test for both the experimental and control classes.

a. The Result of Pre-test

To determine the students' pronunciation ability in reading narrative texts before receiving treatment, the researcher administered a pre-test to both the experimental and

control classes. The students' performances were assessed using the ELSA Speak application. The table below presents the overall pre-test scores of both groups.

Table 1.1

The Result of Pre-test

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation
CONTROL	36	40	15	55	28.81	10.214
EXPERIMENTAL	35	50	15	65	40.37	14.347
Valid N (listwise)	35					

Table 1.1 presents the data indicating that the number of students in both the experimental and control groups was equal, with a total number of 71 participants. For both the experimental group and the control group, 15 was the lowest pre-test score. The experimental group's highest score was 65, whereas the control group's was 55. The experimental group's average score was 40,37 whereas the control group's was 28,81.

The experimental group's standard deviation was thus 14,347, while the control group's was 10,214 Based on the descriptive statistics, the experimental group had a higher average pre-test score (40.37) compared to the control group (28.81), indicating a notable initial difference. Therefore, ANCOVA was used in the further analysis to control for this pre-existing difference and to measure the true effect of the treatment. It can be concluded that the students in the experimental and control groups possessed comparable English pronunciation abilities in storytelling prior to the use of ELSA Speak in the experimental group.

b. The Result of Post -Test

The post-test was administered to both the experimental and control groups after the treatment was completed. To assess students' improvement in pronunciation, the researcher calculated and compared the scores from both groups. The post-test used the same narrative text as the pre-test A summary of the post-test results for both the experimental and control groups is presented in the following table.

Table 1.2

The Result of Post-Test

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation
Control	36	48	17	65	40.31	12.368
experimental	35	45	40	85	63.09	12.521
Valid N (listwise)	35					

Table 1.2 displays the descriptive statistics of the post-test results for both the control and experimental groups. The control group, consisting of 36 students, had post-test scores ranging from 17 to 65, with a mean score of 40.31 and a standard deviation of 12.368. This indicates a moderate level of improvement compared to the pre-test, where their average score was 28.81.

The experimental group, which included 35 students, achieved post-test scores between 40 and 85, with a mean score of 63.09 and a standard deviation of 12.521. This represents a notable improvement from their pre-test mean of 40.37. These results suggest that students in the experimental group, who used the ELSA Speak application, showed a greater increase in their English pronunciation scores compared to those in the control group, who did not receive the same treatment.

Thus, it can be inferred that the use of ELSA Speak had a positive impact on improving students' pronunciation abilities in storytelling, as reflected by the higher mean score and greater range of improvement in the experimental group.

2. Data analysis

At this point, the researcher explains the results of the pre-test and posttest calculations, which are then entered into the Analysis of Covariance (ANCOVA). The purpose of this calculation is to determine the significant effect the ELSA Speak application on students' English pronunciation in storytelling. Due to several requirements for computing the data, there are several assumptions that must be met before performing data analysis using ANCOVA. The first requirements, such as; the distribution of the data should be called normal, the variance between the experimental and control groups should be called homogeneous, there should be no interaction between the pre-test and the group; and also,

the relationship between the pre-test and post-test should be the assumption test, are presented as follows.

a) The Result of Normal Distribution

Researchers conducted a normality test to meet the requirements of ANCOVA. A normality test is used to see whether the scores of each variable are normally distributed or not. Tests for normality can be done with Shapiro Wilk.

Table 2.1

The Result of Normal Distribution

Tests of Normality

	Class	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		statistic	df	Sig.	Statistic	df	Sig.
Score	Pre-test G Control class	.136	36	.089	.940	36	.051
	Post-test G Control class	.093	36	.200*	.974	36	.550
	Pre-test F experimental	.155	35	.033	.955	35	.166
	Post-test F experimental	.169	35	.013	.943	35	.070

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

To ensure the assumption of normality was met before conducting the ANCOVA, the researcher used the Shapiro-Wilk test, which is more appropriate for small sample sizes (fewer than 100 participants). As shown in Table 4.3, the significance values of the Shapiro-Wilk test for each group are as follows: 0.051 for the pre-test of the control group, 0.550 for the post-test of the control group, 0.166 for the pre-test of the experimental group, and 0.070 for the post-test of the experimental group. Since all significance values are greater than 0.05, it can be concluded that the data from both the experimental and control groups are normally distributed. Therefore, the assumption of normality required for conducting the ANCOVA was satisfied.

b) The Result of Homogeneity Test

The next assumption to fulfill the ANCOVA requirements is the variance between groups, which is homogeneous. The Levene test was used to determine the variance of

homogeneity between groups. If the Levene test result is greater than 0.05, it means that the data variance between groups is homogeneous.

Table 2.2

The Result of Homogeneity of Variance

Test of Homogeneity of Variance					
		Levene Statistic	df1	df2	Sig.
score	Based on Mean	2.039	3	138	.111
	Based on Median	1.873	3	138	.137
	Based on Median and with adjusted df	1.873	3	134.683	.137
	Based on trimmed mean	2.091	3	138	.104

To fulfill one of the assumptions required in conducting ANCOVA, the researcher carried out the Levene's Test of Homogeneity of Variance. This test aims to determine whether the variance between the control and experimental groups is equal. Based on the output shown in Table 2.2, the significance value based on the mean is 0.111, based on the median is 0.137, based on the median with adjusted degrees of freedom is 0.137, and based on the trimmed mean is 0.104. Since all significance values are greater than 0.05, it can be concluded that there is no significant difference in variance between the groups. Therefore, the assumption of homogeneity of variance is fulfilled, and the data are suitable for further analysis using ANCOVA.

c) The Result of Homogeneity Regression

The regression homogeneity test is the third assumption that must be met before testing the hypothesis using ANCOVA. The regression homogeneity test aims to determine whether there is a relationship between covariates and independent variables. This assumption will be fulfilled if it is found that there is no relationship between the covariate and the independent variable on the condition that the sig value of the covariate*independent variable is greater than 0,05.

Table 2.3

The Result of Homogeneity Regression

Tests of Between-Subjects Effects

Dependent Variable: Post-test

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	12087.037[a]	2	6043.519	52.642	.000
Intercept	8991.568	1	8991.568	78.322	.000
Pretest	2877.757	1	2877.757	25.067	.000
teaching strategy	4080.165	1	4080.165	35.540	.000
Error	7806.625	68	114.803		
Total	208461.000	71			
Corrected Total	19893.662	70			

a. R Squared = .608 (Adjusted R Squared = .596)

Based on the output of the ANCOVA analysis presented in the table 2.3, the significance value for the pretest variable is 0.000, which is less than 0.05. This indicates that the pretest has a significant effect on the posttest scores, fulfilling the assumption of linearity and justifying the use of the pretest as a covariate in the model. Furthermore, the significance value for the teaching strategy variable is also 0.000, showing that there is a statistically significant difference between the control and experimental groups in terms of posttest scores after controlling for the pretest. The F values for both variables—25.067 for pretest and 35.540 for teaching strategy also indicate strong effects.

Additionally, the R Squared value of 0.608 means that approximately 60.8% of the variance in the posttest scores can be explained by the combination of pretest scores and teaching strategy. These findings suggest that the ANCOVA model is valid, and that the use of the ELSA Speak application as a teaching strategy had a significant effect on students' English pronunciation in storytelling after controlling for their initial pretest performance.

d) The Result of Test of Linear Relationship Covariate and Dependent Variable

To account for variations in the independent variables, covariates were included in the analysis. Covariate testing is primarily used to assess the connection between covariates and dependent variables while accounting for the dependent variable (for each particular group). There must be a significant correlation between the covariate and the dependent variable with a significance level lower than 0.05 in order to continue with the

ANCOVA analysis. The Result of Linear Relationship Covariate and Dependent Variable The group significance value is 0,000, which is less than 0,05 (0,000)

Table 2.4

The Result of Linear Relationship Covariate and Dependent Variable

Tests of Between-Subjects Effects Dependent Variable: Post-Test					
Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	12087.037[a]	2	6043.519	52.642	.000
Intercept	8991.568	1	8991.568	78.322	.000
Pretest	2877.757	1	2877.757	25.067	.000
teaching strategy	4080.165	1	4080.165	35.540	.000
Error	7806.625	68	114.803		
Total	208461.000	71			
Corrected Total	19893.662	70			

a. R Squared = .608 (Adjusted R Squared = .596)

The significance value for the group is 0.000, which is less than 0.05 (0.000 < 0.05), as shown in the table above. This indicates a significant relationship between the covariate and the dependent variable. Therefore, it can be concluded that there is a linear association between the covariate and the dependent variable.

e) The Result of ANCOVA

The fulfillment of the assumptions required for ANCOVA analysis has been shown in the previous sections. The final step is the ANCOVA test itself. ANCOVA serves as a hypothesis testing method that enhances the accuracy of research findings by statistically controlling for the influence of other variables, such as pre-test scores, IQ, and similar factors (Mackey, 2005).

Table 2.5

The Result of ANCOVA

Tests of Between-Subjects Effects Dependent Variable: Post-Test					
Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	12087.037[a]	2	6043.519	52.642	.000
Intercept	8991.568	1	8991.568	78.322	.000
Pretest	2877.757	1	2877.757	25.067	.000

Teaching Strategy	4080.165	1	4080.165	35.540	.000
Error	7806.625	68	114.803		
Total	208461.000	71			
Corrected Total	19893.662	70			

a. R Squared = .608 (Adjusted R Squared = .596)

Based on the table above, the group significance value is 0.000, which is lower than the threshold of 0.05 (0.000 < 0.05). This indicates that the researcher has sufficient evidence to reject the null hypothesis. In other words, the null hypothesis can be rejected, and the alternative hypothesis is accepted. In conclusion, students who were taught using the ELSA Speak application showed significantly better results in English pronunciation for storytelling than those who were taught through conventional methods. This suggests that the use of ELSA Speak is effective in facilitating students' pronunciation skills in the context of storytelling.

Discussion

The results of this research align with the theory of pronunciation learning through technology-assisted applications, as supported by previous studies. According to Wahyuni (2018), mobile-assisted pronunciation tools can support learners by offering immediate feedback and promoting learner autonomy. The ELSA Speak application, which provides real-time correction and interactive pronunciation training, is in line with this concept.

The data analysis in this study shows that students who were taught using the ELSA Speak application performed better in pronunciation tasks compared to those taught using traditional methods. This can be seen from the difference in post-test mean scores between the experimental and control groups. The experimental group scored an average of 63.09, while the control group only achieved 40.31. This significant difference proves that the ELSA Speak application had a positive impact on students' English pronunciation, especially in the context of storytelling.

Furthermore, the results of the ANCOVA analysis also support the effectiveness of ELSA Speak. The significance value for the teaching strategy variable (i.e., the type of instruction used: ELSA Speak vs. conventional method) is also 0.000, which is smaller than the significance level of 0.05. This indicates that there was a statistically significant difference between the two groups after controlling for pre-test scores. The R-squared value of 0.608

shows that 60.8% of the variance in post-test scores was influenced by the pre-test and the teaching strategy used.

Before conducting ANCOVA, several assumptions were tested and all were met. The data were normally distributed, as seen from the Shapiro-Wilk values greater than 0.05 in both groups. The homogeneity of variance was also confirmed with a significance value above 0.05 in Levene's test. There was no interaction between the covariate (pre-test) and the independent variable (teaching strategy), meaning the assumption of homogeneity of regression slopes was fulfilled. Additionally, there was a linear relationship between the covariate and the dependent variable, with a significance value of 0.000 < 0.05.

These findings are consistent with previous research conducted by Adawiah and Muliati (2024), who reported that the use of the ELSA Speak application significantly improved students' pronunciation accuracy and confidence in speaking activities at the vocational high school level. Similarly, in the current study, the ELSA Speak application allowed students to practice independently and receive real-time feedback, which contributed to the improvement of their pronunciation performance in storytelling tasks.

Other studies in the Indonesian context have similarly emphasized the value of integrating digital tools into English instruction. For instance, Darussalam and Ningrum (2023) demonstrated that mind-mapping techniques effectively enhanced students' speaking performance, while Hibatulloh et al. (2024) noted that task-based instruction improved oral accuracy and motivation. Together with Wulansari and Fakhrurriana's (2024) evidence on digital dictionary use, these studies collectively suggest that technology-assisted learning environments can meaningfully support pronunciation and speaking development.

In conclusion, based on the findings and statistical analysis, it can be determined that the ELSA Speak application is effective in improving students' English pronunciation. This application not only provides practice opportunities but also enhances students' confidence in speaking, particularly when performing storytelling tasks. The results of this research reinforce the importance of integrating digital tools into the English learning process to support pronunciation development more effectively. However, this study has several limitations that should be acknowledged. One of the limitations is the relatively short treatment duration, which consisted of only three sessions.

A longer period of exposure to the ELSA Speak application might have resulted in even more significant improvements. Additionally, factors such as students' individual motivation, prior familiarity with mobile learning technology, and differing levels of interest in storytelling may have influenced the outcomes. These variables were not directly controlled in this study but may have contributed to the students' pronunciation performance.

Conclusion

Based on the findings, this study confirms that the ELSA Speak application effectively enhances junior high school students' English pronunciation in storytelling. Students who practiced with ELSA Speak achieved significantly greater gains in accuracy and fluency than those taught with conventional techniques ($p < 0.05$). These improvements indicate that real-time, individualized feedback helps learners identify and correct errors more effectively than teacher-led drills alone.

Beyond the statistical results, the findings have broader pedagogical implications. Teachers can integrate ELSA Speak as a supplementary tool for storytelling activities, allowing students to practice both in class and independently. The app's instant feedback can help address individual pronunciation needs without taking additional class time, while its gamified features sustain learner motivation. This aligns with communicative language teaching principles that emphasize intelligibility and fluency over native-like perfection. Overall, the study suggests that incorporating AI-based tools like ELSA Speak can make pronunciation instruction more interactive, personalized, and contextually meaningful, offering valuable insights for EFL pedagogy and curriculum design.

References

Abimanto, D., & Sumarsono, W. (2024). Improving English Pronunciation with AI Speech-Recognition Technology. *Acitya: Journal of Teaching and Education*, *6*(1), 146–156. https://doi.org/10.30650/ajte.v6i1.3810

Adawiah, R., & Muliati, A. (2024). The effect of Elsa Speak application on students' English pronunciation development. *Journal of Excellence in English Language Education*, *3*(1), 65–70. https://ojs.unm.ac.id/JoEELE/article/view/59627

Anes, C. L. E., Wachyudi, K., & Pahlevi, M. R. (2023). Persepsi Siswa Terhadap Penggunaan Aplikasi ELSA Speak. *Jurnal Educatio FKIP UNMA*, *9*(2), 1147–1152. https://doi.org/10.31949/educatio.v9i2.3235

Basuki, Y. (2018). the Use of Drilling Method in Teaching Phonetic Transcription and Word Stress of Pronunciation Class. *Lingua Litera: Journal of English Language Teaching*

Learning and Literature, *1*(1), 16.

Dandee, W., & Pornwiriyakit, P. (2022). Improving English Pronunciation Skills by Using English Phonetic Alphabet Drills in EFL Students. *Journal of Educational Issues*, *8*(1), 611. https://doi.org/10.5296/jei.v8i1.19851

Darussalam, G. M., & Ningrum, A. S. B. (2023). Improving students' speaking skill in descriptive text using mind mapping method at seventh grade junior high school. *IREELL: Indonesian Review of English Education, Linguistics, and Literature, 1*(1), 36–48. https://doi.org/10.30762/ireell.v1i1.1100

Gusti Bagus, Ni Wayan Prami Wahyudiantari, & Huang Yan Wang. (2024). The Use of Moving Class Learning Model to Enhance Speaking Skills Viewed from Motivation. *Journal of Language and Literature Studies*, *4*(1), 208–220. https://doi.org/10.36312/jolls.v4i1.1789

Hibatulloh, M. F., Arganata, K. L., Latifah, C. N., & Kurniasari, U. (2024). The implementation of teaching speaking using task-based approach in Indonesian secondary school EFL classes: The state of the art. *IREELL: Indonesian Review of English Education, Linguistics, and Literature, 2*(1), 60–77. https://doi.org/10.30762/ireell.v2i1.2812

Kholis, A. (2021). Elsa Speak App: Automatic Speech Recognition (ASR) for Supplementing English Pronunciation Skills. *Pedagogy: Journal of English Language Teaching*, *9*(1), 01. https://doi.org/10.32332/joelt.v9i1.2723

Lestari, R., Harahap, A. L., Munthe, D., Rusli, M., & Rambe, S. A. (2024). The Improving Of Students English Pronunciation Skill By Using English Pronunciation Digital At Man Labuhanbatu , SUMUT. *Journal on Education*, *06*(02), 14797–14809. https://jonedu.org/index.php/joe/article/view/5352/4258

Mubarok, M. Z., & Aziez, F. (2024). AI and non-AI Tools in Teaching English Pronunciation to EFL Learners. *Canadian Journal of Language and Literature Studies*, *4*(5), 22–33. https://doi.org/10.53103/cjlls.v4i5.178

Nugraha, S. V., Irawan, L. A., Haerazi, & Frank, T. O. (2022). Segmental Aspects of Pronunciation Errors Produced by ELE Students in Classroom Settings. *Journal of Language and Literature Studies*, *2*(2), 88–98. https://doi.org/10.36312/jolls.v2i2.759

Nurlita, A. T., Rezeki, Y. S., & Rahmani, E. F. (2024). Students' Perception Analysis Toward The Use of Elsa Speak Application in Improving Their Pronunciation (A Descriptive Study to The Eighth-Grade Students of SMPN 16 Pontianak in Academic Year 2022/2023). *Jurnal Pendidikan Tambusai*, *8*(3), 46010–46017.

Wulansari, T. W., & Fakhrurriana, R. (2024). The influence of using Online Oxford Dictionary toward pronunciation and vocabulary mastery. *IREELL: Indonesian Review of English Education, Linguistics, and Literature, 2*(1), 113–122. https://doi.org/10.30762/ireell.v2i1.3251