IJSS: International Journal of Social Science
Vol. 2 No. 2, August 2026, pp: 366-379
ISSN: 3090-0646
DOI: https://doi.org/10.51805/ijss.v2i2.563

The Effectiveness Of Using Elsa Speak Application In Teaching English Pronunciation Accuracy At 10th Grade Student Of SMK Islam Terpadu Marifa Tasikmalaya

Leni Nurlela[1], Wahyudin Ahmadi[2], Rini Septa Anggraini[3]
Panca Sakti University Of Bekasi
Email: ellenmy09@gmail.com

ABSTRACT

Pronunciation plays a crucial role in effective English oral communication. However, vocational high school students in Islamic boarding school (pesantren) settings often struggle with pronunciation accuracy due to native language interference, limited classroom practice time, and high speaking anxiety. Therefore, this study aimed to investigate the effectiveness of using the AI-based ELSA Speak application in improving the English pronunciation accuracy of 10th grade students at SMK Islam Terpadu Marifa Tasikmalaya. A quantitative approach with a pre-experimental one-group pretest-posttest design was employed in this study. The research involved 21 10th grade Computer and Network Engineering (TKJ) students selected through saturated sampling (N = 21). Data were collected using oral pronunciation tests administered before and after a five-meeting intervention, assessing both segmental features (vowels and consonants) and suprasegmental features (word stress and intonation). Content validity of the instrument was verified through Aiken's V, and reliability of the instrument was verified through Cronbach Alpha and Cohen Kappa (Weighted Kappa). The data were analyzed using descriptive statistics, the Shapiro-Wilk normality test, a Paired Sample T-test, and Normalized Gain (N-Gain) analysis. The research results revealed a statistically significant increase in the mean pronunciation accuracy score from 57 on the pretest to 73 on the post-test ($p < 0,001$). The paired sample t-test demonstrated $t(20) = -12,573$ $p < 0,001$, leading to rejection of the null hypothesis (H_0) and acceptance of the alternative hypothesis (H_a). Furthermore, the N-Gain score averaged 0,4047 (40,47%), indicating a moderate level of instructional effectiveness. These findings confirm that integrating ELSA Speak into structured classroom activities provides an effective AI-assisted learning environment that enhances phonetic awareness, reduces native language interference, and lowers foreign language anxiety.

Keywords: *ELSA Speak, Pronunciation Accuracy, Artificial Intelligence (AI), Mobile-Assisted Language Learning (MALL), Vocational High School.*

INTRODUCTION

As a system of auditory symbols, language enables human beings to communicate, engage in social interaction, and express cognitive as well as effective states. Halliday (1978) defines language can be understood as a social semiotic system, not simply a collection of grammatical rules, but rather a resource that humans employ to construct meaning within social contexts. According to Kurniawan (2024) language functions as the main means of communication, social identity, and social interaction tools, whether in oral, written, or sign form, which are studied in a cultural group. English is among the languages that have gained global

prominence, functioning today as an international medium of communication. Recognized as one of the most influential languages worldwide, it is applied across diverse domains such as education, commerce, science, technology, and diplomacy

In this globalized world, the role of English is becoming increasingly important in various aspects of life. Therefore, mastering English is no longer considered an option, but rather a necessity. However, learning English is not only related to understanding grammar and vocabulary, but also includes aspects of Pronunciation that play an important role in effective communication.

Pronunciation refers to the vowel and consonant sounds produced through the articulation of the mouth, tongue, and lips, emphasis on specific syllables, the tone and tone of speech, and the way of combining sounds between words to make them sound more fluid and less rigid. According to Kumar & Sathyanathan (2025) Good pronunciation of words or sentences while speaking will make it easier for the listener to understand the speaker's intentions, whereas incorrect pronunciation can lead to misunderstandings of meaning. In addition, good and correct pronunciation can increase confidence when speaking. Many English learners are hesitant or afraid to speak for fear of making mistakes in pronunciation. Pronunciation also plays a role in listening skills. When someone is used to pronouncing words correctly, it becomes easier for them to understand what others are saying. This is because speaking and listening skills are interrelated. The better a person is at Pronunciation, the better their ability to understand spoken language in interacting in real-life situations.

Some of the factors that affect students' pronunciation difficulties are the difference in sound system between English and Indonesian, which is the mother tongue. Then, the lack of English practice in the classroom causes students to not be used to using English orally. And also learning methods that still focus on grammar and vocabulary aspects. In addition, students' low self-confidence and fear of making mistakes are the most significant obstacles in the process of practicing English. Aryani et al. (2025) explained that vocational school students often experience psychological barriers in speaking English for fear of being ridiculed for their accents. These problems result in students' communication skills in English being less effective and often having difficulty communicating their intentions clearly that they are difficult for the interlocutor to understand. This shows that mastering Pronunciation skills is a very important aspect and needs to be considered in the learning process.

Although mastery of grammar and vocabulary is often the primary focus in English language learning, pronunciation remains one of the aspects that poses the greatest difficulty for vocational high school students in Indonesia. Various studies show that many students in Indonesia still make mistakes in pronouncing vowel and consonant sounds, placing word stress, and using intonation, making their speech often difficult for listeners to understand. Based on several studies in Indonesia, students' pronunciation skills remain a challenge in English language learning, including at the vocational high school (SMK) level. An improvement in the pronunciation skills of 11th grade students at SMKN 2 Purwokerto was observed by Hoeriyah (2024) following the use of the ELSA Speak Application, outperforming conventional instruction. Furthermore, research by Nurul Eka Musnianti et al. (2025)

demonstrates that incorporating the ELSA Speak application effectively enhances students' ability to pronounce vowels at SMK PGRI 1 Giri Banyuwangi. These findings suggest that enhancing vocational students' pronunciation performance requires integrating interactive technology equipped with real time feedback mechanisms.

This situation was also confirmed by the results of a preliminary study conducted by researcher at SMK Islam Terpadu Marifa Tasikmalaya through initial observations and discussions with English Teachers. The observation results showed that the majority of 10th grade students still struggled to pronounce English vocabulary correctly. Many students tend to pronounce words according to Indonesian spelling, pay little attention to word stress and intonation, and lack confidence when asked to say words or sentences in English. Pronunciation instruction is also still limited, so students have few opportunities to receive feedback on their pronunciation errors. Therefore, learning materials that provide pronunciation practice are needed.

Along with the rapid development of digital technology in the world, various innovations in language learning are beginning to develop. One of them is through the use of artificial intelligence-based applications. One of the currently trending apps that can be used for language learning and can help improve Pronunciation skills is the ELSA Speak App.

According to the official website of ELSA Speak, ELSA Speak (English Language Speech Assistant) is an artificial intelligence (AI)-based English Pronunciation and speaking trainer app that helps users improve their speaking skills like a native speaker by providing instant, accurate, and personalized feedback in real-time. With the speech recognition feature, the app is able to detect pronunciation errors and provide specific corrections to each word or sound spoken.

ELSA Speak is equipped with various features specifically designed to improve pronunciation skills. One of its key features is artificial intelligence (AI) based speech recognition, which automatically analyses a user's pronunciation and provides immediate feedback on pronunciation errors for each word, sound, word stress, and intonation. In addition, ELSA Speak also provides phonemic transcription using International Phonetic Alphabet (IPA) symbols. This feature helps users understand how to pronounce a word based on the correct phonetic symbols, so that students not only imitate the sounds of native speakers but also understand the sound representation of each word. Guided by these dual capabilities, learners are able to pinpoint their phonetic mistakes and fix them autonomously through iterative practice.

The use of ELSA Speak in Pronunciation learning is expected to be an alternative solution to overcome the limitations of practice time in the classroom and help students practice independently outside of class hours. According to Kumalasari & Fitri Wulandari (2025), incorporating technological tools like ELSA Speak enriches student learning by delivering tailored feedback that creates a dynamic, impactful educational setting. Furthermore, the interactive nature and private evaluation features of the application boost learners' drive and self-assurance by reducing the anxiety associated with making public mistakes. Further exploration into the utility of the ELSA Speak application for teaching pronunciation in

vocational education is thus necessary. It is hoped that this research will yield effective interventions to refine students' phonological accuracy, enabling clearer and more confident real-world communication.

Multiple investigations have validated the benefits of implementing ELSA Speak in language learning. Specifically, research by Rismawati et al. (2022) indicates a boost in self-directed learning, which aligns with finding from Rinaepi et al. (2022) showing substantial progress in learner pronunciation following application-based interventions. Similarly, the findings Anggraini (2022) state that AI technology in ELSA Speak helps overcome phonetic constraints in English learners. However, most of this research is conducted at the public high school or university level that focuses on purely academic competence. There is still a scarcity of studies that specifically examine the effectiveness of ELSA Speak at the vocational level which has the characteristics of a vocational curriculum and a dominant practice hour load. In addition, students of Marifa Integrated Islamic Vocational School have a background in the pesantren environment that requires a flexible but controlled approach to independent learning media.

While the impact of ELSA Speak has been widely explored across various academic levels in Indonesia, the majority of existing literature centers in public junior and senior high schools (SMP, SMA, and SMK), leaving a gap regarding contexts with distinct student demographics. To this day, research examining the use ELSA Speak in Integrated Islamic Vocational High Schools (SMK Islam Terpadu), particularly in Tasikmalaya Regency, remains very limited. Furthermore, previous studies generally focused only on improvements in pronunciation skills following the use of ELSA speak, without considering the characteristics of the learning environment for students in pesantren-based schools, which have relatively limited time allocated for English language instruction. Consequently, the originality of this research stems from its target population: 10th grade students at SMK Islam Terpadu Marifa Tasikmalaya, whose unique educational background and characteristics distinguish them from subject in prior studies. This study also provides insight into the effectiveness of ELSA Speak in improving pronunciation accuracy within the context of integrated Islamic vocational schools, an area that has not been extensively studied. Through this research, the existing body of knowledge conserning AI-based media in pronunciation pedagogy is expected to be expanded, specifically within Islamic Vocational high schools in Indonesia.

Based on the description above, the researcher is interested in conducting experimental research entitled: "The Effectiveness of Using Elsa Speak Application in Teaching English Pronunciation Accuracy At 10th Grade Student of SMK Islam Terpadu Marifa Tasikmalaya".

METHOD

This study concludes that the use of the ELSA Speak application has the potential to improve the English pronunciation accuracy of tenth-grade students at SMK Islam Terpadu Marifa Tasikmalaya. The research involved 21 students from class X TKJ during the second semester of the 2025/2026 academic year. Using a quantitative pre-experimental one-group pretest-posttest design, the study

examined students' pronunciation performance before and after receiving learning activities supported by the ELSA Speak application.

The pronunciation assessment focused on four important indicators: vowel accuracy, consonant accuracy, word stress, and intonation. These indicators were selected to provide a comprehensive measurement of students' pronunciation abilities. During the treatment, students used ELSA Speak through their smartphones and headsets to practice pronunciation. The application provided speech analysis and immediate feedback, allowing students to identify pronunciation problems and repeat their practice until their performance improved. This interactive process encouraged students to become more actively involved in their pronunciation learning.

The research procedure consisted of several stages, beginning with a pretest to identify students' initial pronunciation abilities. The students then participated in learning activities focusing on vowels, consonants, word stress, and intonation, followed by a posttest to measure their pronunciation performance after the treatment. The collected data were analyzed using descriptive statistics, gain scores, the Shapiro-Wilk normality test, and a paired sample t-test. N-Gain analysis was also used to determine the level of improvement achieved by the students.

Based on the research design and analysis procedures, ELSA Speak can be considered a promising technology-based learning tool for pronunciation instruction. Its ability to provide immediate feedback and opportunities for repeated practice can support students in developing greater awareness and accuracy in producing English sounds. Furthermore, the application can make pronunciation practice more interactive and independent, reducing students' dependence solely on teacher correction.

Therefore, the implementation of ELSA Speak may serve as an alternative strategy for English teachers who aim to strengthen students' pronunciation skills through technology-assisted learning. Future researchers are encouraged to conduct similar studies with larger samples, longer treatment periods, and more controlled research designs to obtain broader evidence regarding the effectiveness of ELSA Speak in English pronunciation learning.

RESULTS

Analysis Requirement Test

The normality test was conducted to determine whether the pretest and post-test data were normally distributed, a key prerequisite for selecting appropriate statistical analysis. Because the sample size was small (N=21), the Shapiro-Wilk test was applied via IBM SPSS Statistics.

Table 1. Test of Normality used Shapiro Wilk

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pretest Pronunciation Accuracy	.166	21	.132	.965	21	.625
Posttest Pronunciation Accuracy	.093	21	.200*	.977	21	.876

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

a. Lilliefors Significance Correction

As shown in the table below, the Shapiro Wilk test yielded a statistic of $W = 0.965$ with a significance value of $p = 0.625$ ($p > 0.05$) for the pretest pronunciation accuracy. Similarly, the post-test pronunciation accuracy obtained a statistic of $W = 0.977$ with a significance value of $p = 0.876$ ($p > 0.05$). since the significance values for both pretest and post-test exceed the standard alpha level of 0.05, it can be concluded that the data from both testing stages are normally distributed. Consequently, parametric statistical analyses with Paired Sample t-test can be used for further hypothesis testing.

Hypothesis testing Results

Because the pretest and post-test data met the assumption of normal distribution, a paired samples t-test was conducted to determine whether a statistically significant difference existed in students' pronunciation accuracy before and after the intervention.

Table 2. Paired Sample Correlations

Paired Samples Correlations

		N	Correlation	Significance One-Sided p	Significance Two-Sided p
Pair 1	Pretest Pronunciation Accuracy & Posttest Pronunciation Accuracy	21	.825	<.001	<.001

Table 3. Paired Sample Test

Paired Samples Test

		Paired Differences					t	df	Significance One-Sided p	Significance Two-Sided p
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper				
Pair 1	Pretest Pronunciation Accuracy - Posttest Pronunciation Accuracy	-16.714	6.092	1.329	-19.487	-13.941	-12.573	20	<.001	<.001

Furthermore, the Paired Samples Correlations revealed a strong positive relationship between the pretest and post-test scores (r = 0.825, p < 0.001). The results of the Paired Sample t-test demonstrated a statistically significant difference between the pretest and post-test scores, t (20) = -12.573, p < 0.001 (two tailed). Because the p-value is less than the significance alpha level of 0.05 (p < 0.05), the null hypothesis (H_0) is rejected and the alternative hypothesis (H_a) is accepted. This confirms that the intervention had a significant effect on improving students' pronunciation accuracy.

Effectiveness Testing Results

To determine the level of effectiveness of using the ELSA Speak application in improving students' pronunciation accuracy, a Normalized Gain (N-Gain) score analysis was calculated based on Hakes's (1998) formula. The N-Gain score measures the improvement in student learning outcomes relative to the maximum possible improvement.

Table 4. Descriptive Statistics of N-Gain Score

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
NGain_Score	21	.11	.77	.4047	.17504
NGain_Persent	21	10.87	76.92	40.4719	17.50373
Valid N (listwise)	21				

As shown in the table above, the mean N-Gain score obtained from 21 students was 0.4047 (40.47 %), with a minimum score of 0.11 (10.8%) and maximum score of 0.77 (76.92%). Based on Hakes's, (1998) classification criteria, an N-Gain score between 0.30 and 0.70 ($0.30 \leq g \leq 0.70$) falls into the moderate category. In terms of percentage, a score ranging from 40% to 55% is interpreted as moderately effective.

Discussion of Research Findings

The statistical analysis presented in the findings section demonstrates a noticeable improvement in the students' English pronunciation accuracy following the integration of the ELSA Speak application. The empirical evidence is supported by a statistically significant increase in mean scores from the pretest (M = 57) to the post-test (M = 73), alongside a paired sample t-test outcome of $t(20) = -12.573$, $p < 0.001$. Because the p -value falls well below the standard threshold of $\alpha = 0.05$, the null hypothesis (H_0) is rejected in favor of the alternative hypothesis (H_a). This confirms that targeted intervention using an AI-driven Mobile-assisted Language Learning (MALL) tool produces a statistically significant positive effect on the pronunciation accuracy of 10th grade students at SMK Islam Terpadu Marifa Tasikmalaya.

To further contextualize the magnitude of this improvement, the Normalized Gain (N-Gain) analysis yielded a mean score of 0.4047 (40.47%). Based on Hake's (1998) effectiveness criteria, an N-Gain falling within the range of 0.30 to 0.70 represents a moderate level of instructional effectiveness. This result is particularly meaningful considering the structural constraints within the research settings: a short five meeting intervention window, a relatively small sample size (N = 21), and the strictly regulated mobile device policy characteristic of a *pesantren* -based vocational school.

The baseline pretest scores highlighted substantial initial challenges among participants, with 80.95% of the class falling into the "Poor" or "Very Poor categories (M = 57). This baseline performance aligns with earlier findings by Rothgerber (2022), who observed that EFL learners frequently experience first language interference. In this study, the students' native phonological systems primarily Sundanese and Indonesian exerted strong influence over their speech, causing them to substitute unfamiliar English phonemes with familiar native equivalents.

The structural divergence between the vowel systems of Indonesian and Received pronunciation English presented a major hurdle. Students initially struggled distinguish short vowels from long vowels (e.g., confusing /ɪ/ in sit with /i:/ in seat) and frequently monophthongized complex English diphthongs. Similar patterns occurred with consonant production, particularly voiced/voiceless dental fricatives (/θ/, /ð/) and postalveolar affricates (/t/ , /d/), which are absent in the students' native languages.

The immediate visual and auditory feedback mechanisms inherent to ELSA Speak's Automatic Speech Recognition (ASR) engine played a crucial role in addressing these segmental errors. When practicing target vocabulary, the application provided real-time, color-coded phonetic transcriptions alongside International Phonetic Alphabet (IPA) representations. This explicit feedback loop provided high-quality, comprehensible input (i + 1) while encouraging controlled output generation, aligning with Krashen's Input Hypothesis and Swain's Output Hypothesis. By viewing exact sound breakdowns on screen, students moved beyond passive repetition toward conscious phonetic monitoring, enabling them to self-correct articulatory positioning for problematic English vowels and consonants.

While segmental accuracy improved consistently across the cohort, progress on suprasegmental elements, specifically word stress and intonation, exhibited a distinct dynamic. During initial sessions, participants displayed a marked tendency toward monotone delivery and flat stress distribution. This behavior was driven by both transfer from Indonesian and speaking anxiety arising from fear of making mistakes.

Although ELSA Speak provides effective automated assessments for syllable stress placement, its algorithmic evaluation of overall sentence intonation, pitch variation, and natural speech rhythm remains somewhat rigid. As noted in the theoretical review, ASR tools can occasionally struggle to accurately evaluate broad contours of human tone. To address this technical limitation, the intervention integrated Vygotsky's (1978) concept of instructional scaffolding. The teacher acted as a human mediator, complementing the application's automated drill cycles with targeted mini lessons on falling pitch (in WH-questions and statements) and rising pitch (in Yes/No questions). Combining AI-driven individual drills with teacher guided communicative feedback ensured that students improved not only their broader expensive intonation.

A unique dimension of this study lies in its institutional context: an integrated Islamic vocational high school (SMK Islam Terpadu) operating within a pesantren (Islamic boarding school) environment. Under standard conditions, school regulations strictly restrict routine student smartphone access during operational hours. Consequently, introducing Mobile-Assisted Language Learning required carefully managed classroom integration rather than unrestricted off campus homework assignments.

Despite these constraints, structured 60 minutes sessions equipped with headsets allowed students to maximize their focus during target practice. The private feedback provided by the application created a low anxiety "safe space" for oral practice. Students who were previously hesitant to speak in front of their peers out of fear of public mispronunciation were able to practice repeatedly without fear on ridicule. This increase in learner autonomy and psychological comfort directly contributed to the positive shift in post-test performance (33.3% of students achieved "Good" scores and 28.5 reached "Very Good")

The findings of this research strongly align with and extend existing CALL and MALL literature. The significant post treatment score gain mirrors findings by Apriyawanto et al (2025) and Mubarok et al (2025), who recorded comparable enhancements in syllable stress and phoneme accuracy among secondary and

vocational students. Furthermore, the positive impact on learner engagement echoes the observations of Nuraeni et al (2023) and Ngan et al. (2024), who highlighted how real-time ASR feedback fosters confidence in speaking.

Importantly, this study fills a crucial gap in the literature by providing that AI-based pronunciation tools can be effectively implemented within specialized Islamic vocational contexts. While previous studies predominantly focused on public general high schools or university environments with open smartphone access, these empirical results demonstrate that with teacher supervision and structured sessions design, ELSA Speak remains highly effective even under restricted mobile usage policies.

In summary, the integration of the ELSA Speak application served as an effective technological bridge for 10th grade students at SMK Islam Terpadu Marifa Tasikmalaya. By combining automated ASR feedback with targeted teacher mediation, the intervention successfully reduced mother tongue phonological interference, improved pronunciation accuracy, and postered a supportive, low anxiety environment for English language acquisition.

CONCLUSION

The findings of this study indicate that the use of the ELSA Speak application contributes significantly to improving the English pronunciation accuracy of 10th grade students at SMK Islam Terpadu Marifa Tasikmalaya. The study involved 21 students from class X TKJ and employed a quantitative pre-experimental design using a one-group pretest-posttest approach. Students' pronunciation was assessed based on four key aspects: vowel accuracy, consonant accuracy, word stress, and intonation. These aspects were considered important because accurate pronunciation is closely related to the clarity and effectiveness of oral communication in English.

The results show a substantial improvement in students' pronunciation performance following the intervention. The mean pronunciation score increased from 57 in the pretest to 73 in the post-test. This improvement was statistically significant, as indicated by the paired sample t-test, which produced a value of $t(20) = -12.573$ with $p < 0.001$. Because the p-value was below the 0.05 significance level, the null hypothesis (H0) was rejected and the alternative hypothesis (Ha) was accepted. These results demonstrate that the use of ELSA Speak had a statistically significant positive effect on students' pronunciation accuracy.

The magnitude of the improvement was further examined using the Normalized Gain (N-Gain) analysis. The mean N-Gain score was 0.4047, or 40.47%, which falls within the moderate effectiveness category. The individual N-Gain scores ranged from 0.11 (10.8%) to 0.77 (76.92%), indicating that the degree of improvement differed among students. Thus, while the intervention produced a meaningful improvement in pronunciation performance, its overall instructional effectiveness can be characterized as moderate.

One factor that may explain these findings is the feedback system provided by ELSA Speak. The application uses artificial intelligence and Automatic Speech Recognition (ASR) technology to provide immediate feedback on students' pronunciation. This feature allows learners to identify errors in individual sounds and pronunciation patterns and to repeat the exercises based on the feedback

received. The availability of phonetic transcription and International Phonetic Alphabet (IPA) representations also supports students in developing greater awareness of English sound production. Through repeated practice and self-monitoring, students were given greater opportunities to identify and correct their own pronunciation difficulties.

Nevertheless, the findings suggest that ELSA Speak is most appropriately used as a complementary instructional tool rather than as a replacement for teacher-led instruction. Although the application provides useful support for segmental features such as vowels and consonants, its ability to assess broader aspects of suprasegmental pronunciation, particularly sentence-level intonation and natural speech rhythm, remains limited. Teacher intervention therefore remains important for providing explanations, communicative practice, and instructional scaffolding, particularly in areas where automated feedback may not fully capture the complexity of natural spoken English.

The context of the study is also important when interpreting the findings. The research was conducted in an integrated Islamic vocational high school within a pesantren environment, where students' access to smartphones is subject to institutional regulations. Despite these restrictions, the structured use of smartphones and headsets during classroom sessions enabled students to engage in focused pronunciation practice. The private feedback provided by ELSA Speak also created a relatively low-anxiety environment in which students could practice repeatedly without the fear of making mistakes in front of their classmates. This condition may have supported greater confidence and engagement during pronunciation activities.

Overall, the findings confirm that integrating ELSA Speak into structured English pronunciation instruction can improve students' pronunciation accuracy. The increase in the mean score from 57 to 73, the significant paired sample t-test result of $p < 0.001$, and the mean N-Gain of 0.4047 (40.47%) provide empirical evidence of its positive contribution to pronunciation learning. At the same time, the moderate N-Gain result and the limitations of automated assessment indicate that ELSA Speak should be integrated with appropriate teacher guidance. In this context, the application offers a practical form of AI-assisted language learning that can provide repeated practice, immediate feedback, and a more supportive environment for developing English pronunciation skills among vocational high school students.

BIBLIOGRAPHY

Adlakha, Dr. S. (2022). Role of respiratory system in producing speech sounds. *International Journal of Research in English*, *4*(1), 31–33. https://doi.org/10.33545/26648717.2022.v4.i1a.129

Ahsin Fahmi Mubarok, Mochammad Abdullah Salim, & Zhenita Deliany. (2025). AI Based Applications to Foster Pronunciation Accuracy. *Essence: Journal of English Language Teaching, Linguistics, and Literature*, *2*(1). https://doi.org/10.33367/essence.v2i1.7479.

Ahmed, A. A. A., Hassan, I., Pallathadka, H., Keezhatta, M. S., Noorman Haryadi, R., Al Mashhadani, Z. I., ... & Rohi, A. (2022). MALL and EFL learners' speaking: Impacts of Duolingo and WhatsApp applications on speaking accuracy and fluency. *Education Research International*, *2022*(1), 6716474.

Aiken, L. R. (1980). Content Validity and Reliability of Single Items or Questionnaires. *Educational and Psychological Measurement*, *40*(4), 955–959. https://doi.org/10.1177/001316448004000419

Aiken, L. R. (1985). Three Coefficients for Analyzing the Reliability and Validity of Ratings. *Educational and Psychological Measurement*, *45*(1), 131–142. https://doi.org/10.1177/0013164485451012

Aisyah Rizqa Fitri Siregar, Yani Lubis, & Siti Julaika. (2023). The Exploration of English Consonant: A Literature Review. *Jurnal Pendidikan Dan Sastra Inggris*, *3*(2), 42–50. https://doi.org/10.55606/jupensi.v3i2.1982

Albano, R., Curzi, Y., & Fabbri, T. (2020). Qualitative and Quantitative Research. In *System, Actor, and Process: Keywords in Organization Studies* (pp. 80–83). Routledge. https://doi.org/10.4324/9781003057260-26

Ambalegin, A. (2021). Phonological analysis of English vowel pronunciation. *KnE Social Sciences*, 28–45.

Andaristo, A. A., Tatihoran, N., & Firdaus, R. (2025). The Concepts of Population and Sampling in Quantitative Research and the Selection of Participants/Key Informants in Qualitative Research. *Al Hikmah Indonesian Journal of Early Childhood Islamic Education*, *9*(2), 359–373. https://doi.org/10.35896/ijecie.v9i2.910

Anggraini, A. (2022). Improving Students' Pronunciation Skill Using Elsa Speak Application. *Journey: Journal of English Language and Pedagogy*, *5*(1), 135–141. https://doi.org/10.33503/journey.v5i1.1840

Apriyawanto, A., Chandra, A., Rauf, W., & Mulyarti, M. (2025). Improving Syllable Stress Accuracy in EFL Pronunciation: An Experimental Study Using ELSA Speak Application. *Jurnal Edukasi Saintifik*, *5*(2), 77–86. https://doi.org/10.56185/jes.v5i2.1118

Arikunto, S., Saputra, W. N. E., Khoirunnisa, D. A., & Sawai, R. P. (2023). Development and validation of teacher competency perception scale in Indonesia: The Rasch analysis. *Journal of Professional Teacher Education*, *1*(1), 11–21. https://doi.org/10.12928/jprotect.v1i1.491

Aryani, F., Aryani, F., Romadlon, F. N., Romadlon, F. N., Nurcahyo, A. D., & Nurcahyo, A. D. (2025). Vocational High School Students' Difficulties in Practicing Speaking. *Prominent*, *8*(1), 25–38. https://doi.org/10.24176/pro.v8i1.13204

Ayuk Puput Wilujing, & Dina Karikawati. (2025). The Use of ELSA Speak Application to Improve Pronunciation Accuracy of Ninth Graders at SMPN 1 Gondang. *Semantik : Jurnal Riset Ilmu Pendidikan, Bahasa Dan Budaya*, *3*(2), 340–347. https://doi.org/10.61132/semantik.v3i2.2145

B, F., Rachmawati, R., & Priastuti, D. N. (2023). Students' Anxiety While Reading Aloud During English Learning At The TKJ Eleventh Grade Of SMK Nurul Yaqin Tlagah. *Journal of Education and Research*, *2*(1), 60–74. https://doi.org/10.56707/jedarr.v2i1.158

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation hardback with audio CDs (2): A course book and reference guide*. Cambridge University Press.

Crystal, D., & Alan, C. L. (2023). *A dictionary of linguistics and phonetics*. John Wiley & Sons.

Derwing, T. M., & Munro, M. J. (2022). Pronunciation Learning and Teaching. In *The Routledge Handbook of Second Language Acquisition and Speaking* (pp. 147–159). Routledge. https://doi.org/10.4324/9781003022497-14

Esling, J. H. (2024). Phonetics, Articulatory. In *International Encyclopedia of Language and Linguistics* (pp. 509–527). Elsevier. https://doi.org/10.1016/B978-0-323-95504-1.00114-9

English Language Teaching Institute of Symbiosis. (2025, January 21). *The Importance of pronunciation in spoken English.* ELTIS. The Importance of Pronunciation in Spoken English

Gilakjani, A., Ahmadi, S., & Ahmadi, M. (2011). Why is Pronunciation So Difficult to Learn? *English Language Teaching*, *4*(3). https://doi.org/10.5539/elt.v4n3p74

Gilakjani, A. P. (2012). A study of factors affecting EFL learners' English pronunciation learning and the strategies for instruction. *International Journal of Humanities and Social Science*, *2*(3), 119–128.

Gracia, E. (2025). Teaching ESL pronunciation to international teaching assistants with the ELSA Speak app. *Language Learning & Technology*, *29*(1), 1–13. https://doi.org/10.64152/10125/73658

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64–74. https://doi.org/10.1119/1.18809

Halliday, M. A. K. (1978), Language as social semiotic: The social interpretation of language and meaning. Edward Arnold.

Hayder Abdul Hasan Hamza, Noora Adnan Tawfeeq, & Fatimah Jammal Al-Quraishi. (2025). Identifying Pronunciation Errors among Iraqi EFL University learners. *مجلة الباحث*, 3(*43*, 491–476). https://doi.org/10.63797/bjh.v43i3.3258

Hoeriyah, S (2024). *The Effectiveness of ELSA Speak toward students' pronunciation in the Eleventh grade of SMKN 2 Purwokerto* [Thesis, Prof. K.H Saifuddin Zuhri Islamic Islamic University]. Puwokerto. Faculty of Tarbiya and Teacher Training.

Jeanne Ellis Ormrod. (2011). *Educational Psychology: Developing Learner*. Pearson Education Inc.

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

Kosasih, M. M. (2021). Factors Affecting Indonesian Students in Learning English Pronunciation. *International Research in Higher Education*, *6*(3), 13. https://doi.org/10.5430/irhe.v6n3p13

Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. Pergamon Press Inc. https://www.researchgate.net/publication/242431410

Kumalasari, C., & Fitri Wulandari. (2025). The Perception of Using ELSA Speak Application to Improve Speaking Skills of Informatics Students. *Journal of English Language Teaching and English Linguistics*, *10*(2), 23–33. https://doi.org/10.31316/eltics.v10i2.6201

Kumar, R., & Sathyanathan, K. (2025). Pronunciation: Stress, Syllables and Aspiration. In *Academic and Professional English* (pp. 19–30). Springer Nature Singapore. https://doi.org/10.1007/978-981-95-3264-3_2

Kurniawan, I. W. A. (2024). English language and its importance as global communication. *Samā Jiva Jnānam (International Journal of Social Studies)*, *2*(1), 51–57.

Lai Minh Thu. (2025). Applying Speech Recognition Technology In Teaching Speaking Skill At The University Of Labour And Social Affairs. *Tạp Chí Khoa Học Trường Đại Học Mở Hà Nội*, 43. https://doi.org/10.59266/houjs.2025.748

Lenneberg, E.H. (1967). *Biological Fondations of Language.* New York: Wiley.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

McLeod S. (2020). *The Oxford Handbook of Speech Development in Languages of the World* (S. McLeod, Ed.). Oxford University Press. https://doi.org/10.1093/9780191964909.001.0001

Ngân, M. A. P. T. B., Thủy, M. A. L. T., Nguyệt, M. A. N. T. M., Hân, M. A. T. P. B., & Huệ, Ph. D. P. T. (2024). The Application of Elsa Speak Software in English Teachingat Can Tho University of Technology – A Case Study. *International Journal of Social Science and Human Research*, *07*(10). https://doi.org/10.47191/ijsshr/v7-i10-29

Nur Fauziah, Zalzabil Zakiah Aororah, Repi Anjani Ramadhany, & Sitti Maryam Hamid. (2024). The Impact of The English Language Speech Assistant (ELSA) Application on Students' Pronunciation Skills. *Interaction: Jurnal Pendidikan Bahasa*, *11*(2), 193–203. https://doi.org/10.36232/interactionjournal.v11i2.26

Nuraini, W. R., Setyawan, M. A., & Pambayun, R. N. (2024). The effect of elsa speak on pronounciation of english speaking at vocational high school. *English Language Teaching Methodology*, *4*(3), 499–511. https://doi.org/10.56983/eltm.v4i3.1683

Nurul Eka Musnianti, Fitria Putri Fernanda, Retno Wulandari, Nadiyatul Hanina, & Wulan Wangi. (2025). The Effect of ELSA Speak Application on Students' Vowels Pronunciation Ability at SMK PGRI 1 Giri Banyuwangi. *Jurnal Bina Ilmu Cendekia*, *6*(2), 184–189. https://doi.org/10.46838/jbic.v6i2.1063

Richard P. Halgin, & Susan Krauss Whitbourne. (2007). *Abnormal Psychology: Clinical Perspectives on Psychological Disorders*. McGraw-Hill.

Rinaepi, R., Triwardani, H. R., & Azi, R. N. (2022). The Effectiveness of Elsa Speak Application to Improve Pronunciation Ability. *Jurnal Fakultas Keguruan Dan Ilmu Pendidikan*, *3*(1), 28–33.

Rismawati, D., Suryana, Y., & Agustiana, V. (2022). The Effectiveness Of Elsa Speaking Application In Improving English Pronunciation. *The Proceedings of English Language Teaching, Literature, and Translation (ELTLT)*, *10*(1), 177–184. https://proceedings.unnes.ac.id/eltlt/article/view/1329

Roach, P. (2009). *English phonetics and phonology paperback with audio CDs (2): A practical course*. Cambridge university press.

Rothgerber, J. (2022). The Causes of Learner Pronunciation Problems in English. In *Research Anthology on Applied Linguistics and Language Practices* (pp. 1495–1513). IGI Global. https://doi.org/10.4018/978-1-6684-5682-8.ch072

Samad, I. S., & Ismail, I. (2020). ELSA speak application as a supporting media in enhancing students' pronunciation skill. *Majesty Journal*, *2*(2), 1–7.

Sitompul, E. D., Hasibuan, M., Nasution, S. O., Aulia, W., & Santoso, D. (2024). Student's Difficulties In Mastering English Pronunciation. *Jurnal Ilmu Sosial, Humaniora Dan Seni*, *2*(6), 639–641. https://doi.org/10.62379/jishs.v2i6.1863

Swain, M. (1993). The Output Hypothesis: Just Speaking and Writing Aren't Enough. *The Canadian Modern Language Review*, *50*(1), 158–164. https://doi.org/10.3138/cmlr.50.1.158

: Students' Voices. *Scope : Journal of English Language Teaching*, *8*(2), 288. https://doi.org/10.30998/scope.v8i2.22451

Tabain, M. (2020). The phonetic properties of consonants. In *The Routledge Handbook of Phonetics* (pp. 264–288). Routledge. https://doi.org/10.4324/9780429056253-11

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Widagsa, R. (2020). Intonation Of English Declarative Sentence Produced By English Department Students. *Journal of English Language Teaching and English Linguistics*, *1*(1). https://doi.org/10.31316/eltics.v1i1.382

Wiranda, A., Lestari, T. D., & Lubis, Y. (2023). The Role Of English Consonants In Language Development : Acquisition And Articulation. *Jurnal Multidisiplin Dehasen (MUDE)*, *2*(3). https://doi.org/10.37676/mude.v2i3.4393

Zuhairoh, Z., Syafa'ah, N., & Kurniati, D. (2024). Content and Face Validity Analysis on 9th grade final test items for secondary school level. *Prominent*, *7*(1), 21–28. https://doi.org/10.24176/pro.v7i1.10608