Journal of Language
Volume 8 Number 2 pp: 343-350
e-ISSN: 2685-8878 | p-ISSN: 2655-9080
DOI: https://doi.org/10.30743/jol.v8i2.14234

ELSA Speak in Improving the Students' Speaking Competence in Learning English

Puan Suri Mira Annisa[1], Rafika Dewi Nasution[2], Johannes Jefria Gultom[3], Fanny Rizki Sembiring[4]
English Department, Faculty of Languages and Arts, Universitas Negeri Medan, Indonesia
E-mail: puansuri@unimed.ac.id

Article Info
Article history:
Received: month dd, yyyy
Accepted: month dd, yyyy
Published: month dd, yyyy

Keywords:
ELSA Speak;
AI-assisted language learning;
speaking Competence;
learning English

ABSTRACT
This study investigates the effectiveness of integrating the ELSA Speak as an artificial intelligence (AI)-based language learning application to enhance students' speaking competence following an eight-week course intervention. A quantitative, one-group pre-test–post-test design was used with 32 undergraduate students enrolled in a Receptive Oral Language Skills course. Speaking performance was assessed in five reported dimensions: pronunciation, fluency, vocabulary, grammar, and confidence/comprehensibility. The reported mean scores increased from pre-test to post-test in all five dimensions. Pronunciation showed the largest gain (from 65.2 to 84.5; +19.3 points), followed by confidence (64.9 to 83.1; +18.2), grammar (66.2 to 82.5; +16.3), fluency (66.7 to 80.4; +13.7), and vocabulary (68.1 to 81.8; +13.7). A post-intervention questionnaire also indicated favorable perceptions: across the ten reported items, 78.1%–84.4% of respondents selected "Strongly Agree," while the remaining 15.6%–21.9% selected "Agree." Because the available results contain descriptive summaries only, no claim of statistical significance is made. The findings show improvement over the intervention period and positive learner perceptions, but the one-group design cannot establish that ELSA Speak alone caused the changes. The results support considering ELSA Speak as a supplementary practice tool alongside teacher-guided communicative activities, while further controlled research is needed.

Corresponding Author:
Puan Suri Mira Annisa
Politeknik Mandiri Bina Prestasi Medan, Indonesia
E-mail: puansuri@unimed.ac.id

1. Introduction

Oral communication is an important component of university students' and professional preparation. In English as a Foreign Language (EFL) settings, however, students may have limited opportunities to use English beyond the classroom and may experience difficulties with pronunciation, fluency, vocabulary retrieval, and confidence during spoken interaction. These challenges make it useful to explore supplementary forms of guided practice that can extend opportunities for speaking beyond scheduled class time.

AI-supported language-learning applications offer one possible form of supplementary practice. ELSA Speak provides speech-focused activities and automated feedback intended to help learners identify aspects of their pronunciation that may need attention. Its potential pedagogical value lies in the opportunity to repeat speaking tasks and receive feedback during individual practice. Nevertheless, feedback from a pronunciation-oriented application should not be treated as equivalent to a comprehensive assessment of speaking competence, which also involves fluency, lexical and grammatical resources, comprehensibility, and the ability to communicate meaning in context.

Previous work on computer-assisted language learning and AI-supported language learning has reported potential benefits for language practice, learner autonomy, and pronunciation development (Hanafiah et al., 2022; Woo & Choi, 2021). Studies of ELSA Speak have also considered its role in pronunciation learning (Aulia & Santosa, 2025; Winduastuti et al., 2026). However, positive perceptions or improvements in pronunciation do not by themselves establish gains across broader speaking dimensions. In addition, when an application is integrated with regular classroom instruction, observed changes may reflect the combined learning experience rather than the application in isolation.

The present study addresses this issue by describing pre-test and post-test speaking scores and students' post-intervention perceptions after an eight-week intervention combining ELSA Speak with classroom activities. The study does not include a control group; therefore, it is intended to document changes during the intervention rather than to establish a causal effect attributable solely to the application.

The study expected to examine the use of Elsa Speak during the course and assessed the students' speaking competences

2. Literature Review

2.1 AI-supported language learning and ELSA Speak

AI-supported language-learning tools can provide repeated practice and automated feedback, potentially allowing learners to focus on specific features of their performance. Woo and Choi (2021) reviewed AI technologies used in language learning, while Hanafiah et al. (2022) examined computer-assisted language learning in relation to vocabulary, speaking skills, and foreign-language speaking anxiety. These studies provide a broader rationale for examining technology as a supplement to instruction, but findings from computer-assisted learning generally should not automatically be attributed to a particular application.

ELSA Speak is designed to support English pronunciation practice through speech-recognition technology and feedback on spoken production. Its feedback and repeatable exercises may help learners notice pronunciation features and practice them independently. Research specifically addressing ELSA Speak has discussed pronunciation development and learner experience (Aulia & Santosa, 2025; Winduastuti et al., 2026). The current study extends this area by reporting changes in several speaking-score dimensions as well as questionnaire responses. Because classroom activities occurred alongside application use, the study evaluates the combined intervention as implemented, not the isolated effect of ELSA Speak.

2.2 Speaking competence

Speaking competence is multidimensional. Pronunciation concerns the production of sounds and the use of stress, rhythm, and intonation; fluency concerns the flow and continuity of speech; vocabulary concerns the range and appropriateness of lexical choices; and grammar concerns the accuracy and complexity of structures used to convey meaning. Comprehensibility reflects how readily listeners can understand a speaker. These dimensions overlap during communication but should remain analytically distinct.

Vocabulary knowledge is associated with second-language speaking proficiency because speakers need lexical resources to formulate messages (Uchihara & Clenton, 2023). Pronunciation can affect intelligibility, while fluency and interactional practice are important for communicating in real time. Communicative activities such as speaking clubs and debates may also support oral performance (Aditya et al., 2024; Tarigan & Lubis, 2024). Consequently, an application that emphasizes pronunciation may be most directly aligned with pronunciation practice, whereas broader speaking development may require complementary classroom tasks.

Journal of Language
Volume 8 Number 2 pp: 343-350
e-ISSN: 2685-8878 | p-ISSN: 2655-9080
DOI: https://doi.org/10.30743/jol.v8i2.14234

Confidence and speaking anxiety are affective constructs rather than direct substitutes for linguistic performance. A learner may feel more confident without demonstrating equivalent gains in fluency or grammatical accuracy, and self-reported confidence does not establish a reduction in anxiety unless anxiety is measured with an appropriate instrument. In this study, confidence appears as a reported speaking-score dimension and as a questionnaire topic; the manuscript does not provide a separate validated measure of speaking anxiety. Therefore, conclusions about anxiety reduction are not drawn.

2.3 Theoretical perspective

Krashen's Input Hypothesis (1985) emphasizes the role of comprehensible input in language acquisition. In the present intervention, modeled spoken English and listening-pronunciation activities may have provided input for learners to notice and practice speech features. However, the study did not directly measure input comprehension or isolate input as a causal mechanism.

Long's Interaction Hypothesis (1996) highlights the role of interaction and negotiation of meaning in second-language development. ELSA Speak's automated feedback can provide a form of learner–system feedback, but it is not equivalent to interpersonal interaction or negotiation of meaning with a teacher or peer. Classroom speaking and presentation activities are therefore important parts of the intervention because they provide opportunities to use language communicatively. These theories inform the interpretation of the combined instructional approach; they are not directly tested by the study design.

3. Methodology

3.1 Research design

The study used a quantitative, one-group pre-test–post-test pre-experimental design. Participants completed a speaking pre-test, took part in an eight-week course intervention that combined ELSA Speak-supported practice and classroom speaking activities, and then completed a post-test. The design permits a descriptive comparison of scores before and after the intervention. Because there was no control group or random assignment, changes cannot be attributed to ELSA Speak alone. Other possible explanations include ordinary course instruction, practice effects from taking the test twice, maturation, and other learning experiences during the eight-week period.

3.2 Participants and sampling

The participants were 32 undergraduate students aged 18–21 enrolled in a Receptive Oral Language Skills course. They had completed introductory English courses, had basic smartphone literacy, and volunteered to participate. Purposive sampling was used because the participants were enrolled in the course in which the intervention was implemented. The sample came from one course at one institution; therefore, the findings should not be generalized to other student populations without further research.

Participants received information about the study's purpose, procedures, confidentiality, and voluntary nature, and written informed consent was obtained. The manuscript should identify the relevant institutional ethics-review procedure and approval or exemption reference, if applicable, before publication.

3.3 Intervention and procedure

The intervention lasted eight weeks and integrated ELSA Speak with classroom speaking instruction. The reported activities included a pre-test, ELSA Speak practice sessions, listening and pronunciation activities, presentation practice, and a post-test with

post-intervention data collection. The application was used as a supplementary practice tool, while classroom activities provided opportunities to apply spoken English in course tasks.

The intervention was implemented over eight weeks as part of a Receptive Oral Language Skills course. It combined ELSA Speak with teacher-guided classroom activities. The procedure included an initial speaking pre-test, practice using ELSA Speak, listening and pronunciation activities, classroom speaking and presentation practice, and a post-test administered at the end of the intervention. ELSA Speak was used as a supplementary learning resource rather than as a replacement for classroom instruction. The classroom component provided opportunities for students to apply their spoken English in communicative tasks.

3.4 Speaking assessment

Speaking performance was assessed with an analytic rubric described as adapted from Brown (2007) and the IELTS Speaking Band Descriptors. The rubric covered pronunciation, fluency, vocabulary, grammar, and comprehensibility/confidence. Each component was reportedly scored on a 1–20 scale, yielding a maximum total of 100. However, the component means reported in Table 1 are above 20. The scoring scale, conversion procedure, and table values must therefore be checked against the original rubric and rating records before submission. Two experienced English lecturers independently rated students' speaking performances.

The rubric was informed by Brown (2007) and the IELTS Speaking Band Descriptors. The final manuscript should explain how the criteria were selected and adapted to the speaking tasks and learning context. Two experienced English lecturers independently rated the students' speaking performances.

3.5 Questionnaire

A questionnaire was administered after the intervention to collect students' perceptions of the learning experience, comprehension of the learning process, motivation, confidence, independent learning, clarity of instructions, pronunciation practice, critical thinking, perceived language improvement, and overall perceived usefulness. The available summary reports ten items and two response categories: Strongly Agree and Agree.

The questionnaire was intended to document students' perceptions of the learning experience, including motivation, confidence, independent learning, pronunciation practice, perceived improvement, and perceived usefulness.

3.6 Data analysis

The available speaking-test results were summarized using pre-test means, post-test means, and mean-score differences. Questionnaire results were summarized as frequencies and percentages for the two response categories reported in the source manuscript. The reported percentages correspond to counts out of 32 participants. The source manuscript does not provide participant-level pre-test and post-test scores, standard deviations, test statistics, p-values, confidence intervals, or effect sizes; inferential significance therefore cannot be verified from the available information.

If participant-level paired scores are available, conduct an appropriate paired analysis after checking the distribution of the paired differences and the relevant assumptions. Report the selected test, test statistic, degrees of freedom where applicable, exact p-value, confidence interval, and an appropriate effect size. If raw data are unavailable, retain descriptive language and do not claim statistical significance. All statistics must be calculated from the original data.

Journal of Language
Volume 8 Number 2 pp: 343-350
e-ISSN: 2685-8878 | p-ISSN: 2655-9080
DOI: https://doi.org/10.30743/jol.v8i2.14234

4. Results and Discussion

4.1 Results

4.1.1 *Pre-test and Post-test Speaking Scores*

Table 1 presents the reported mean scores before and after the intervention. Mean scores increased in all five dimensions. Pronunciation had the largest reported increase (+19.3 points), followed by confidence (+18.2), grammar (+16.3), fluency (+13.7), and vocabulary (+13.7). These are descriptive changes; without participant-level data and inferential analysis, they should not be described as statistically significant.

Speaking dimension	Pre-test mean	Post-test mean	Mean-score difference
Pronunciation	65.2	84.5	+19.3
Fluency	66.7	80.4	+13.7
Vocabulary	68.1	81.8	+13.7
Grammar	66.2	82.5	+16.3
Confidence	64.9	83.1	+18.2

Note. n = 32. The table reproduces the descriptive means reported in the source manuscript. Standard deviations and inferential statistics were not supplied.

4.1.2 *Questionnaire Responses*

The questionnaire summary reported only two response categories, Strongly Agree and Agree. Based on the reported percentages and n = 32, the corresponding counts are shown in Table 2. These responses indicate favorable perceptions among the respondents, but they do not independently establish objective learning gains or causal effectiveness.

No.	Questionnaire topic*	Strongly Agree, n (%)	Agree, n (%)	Total n
1	New learning experience	27 (84.4%)	5 (15.6%)	32
2	Ability to follow the learning process	25 (78.1%)	7 (21.9%)	32
3	Learning motivation	25 (78.1%)	7 (21.9%)	32
4	Confidence in speaking English	27 (84.4%)	5 (15.6%)	32
5	Independent learning	25 (78.1%)	7 (21.9%)	32
6	Visual clarity and instructional guidance	27 (84.4%)	5 (15.6%)	32
7	Pronunciation and intonation practice	27 (84.4%)	5 (15.6%)	32
8	Critical-thinking stimulation	25 (78.1%)	7 (21.9%)	32
9	Perceived overall language improvement	25 (78.1%)	7 (21.9%)	32
10	Perceived overall usefulness of ELSA Speak	27 (84.4%)	5 (15.6%)	32

Across the ten items, all reported responses were either Strongly Agree or Agree. The proportion selecting Strongly Agree ranged from 78.1% to 84.4%. The highest reported proportion (84.4%, 27 of 32 students) appeared for items concerning the learning experience, confidence, clarity of the interface and guidance, pronunciation and intonation, and overall perceived usefulness. The remaining items received 78.1% Strongly Agree responses (25 of 32 students). Because no neutral or negative responses are reported, the authors should verify whether the original instrument offered additional response options and whether the table represents all responses.

4.2 Discussion

The descriptive results show higher post-test mean scores than pre-test means across all five reported dimensions. The largest increase occurred in pronunciation (+19.3 points). This pattern is compatible with the focus of ELSA Speak on pronunciation practice and

immediate feedback. Repeated opportunities to produce speech, notice feedback, and retry a task may help learners attend to sound production, stress, and intonation. Nevertheless, the design cannot determine whether the observed pronunciation change resulted from the application, classroom teaching, repeated testing, or a combination of these factors.

The increases in fluency and vocabulary were both 13.7 points, while grammar increased by 16.3 points. These changes may reflect the wider course activities, including listening-pronunciation work and presentation practice, as well as students' continued English learning during the study period. Since ELSA Speak is not a substitute for extended communicative interaction, the results should not be interpreted as evidence that application practice alone produces equivalent gains in every dimension of speaking competence.

The reported confidence score increased by 18.2 points. This result should be interpreted cautiously because the manuscript does not fully explain how confidence was operationalized within the speaking rubric. Confidence is affective and should be distinguished from observable language performance. Likewise, the questionnaire's favorable responses suggest that students perceived the application positively, but perceptions are not direct evidence of proficiency gains. The current manuscript does not report a separate validated measure of speaking anxiety; therefore, it cannot conclude that anxiety decreased.

The findings are broadly consistent with literature suggesting that technology-supported practice can supplement language learning and support learner autonomy (Hanafiah et al., 2022; Woo & Choi, 2021). They are also relevant to studies focusing specifically on ELSA Speak and pronunciation practice (Aulia & Santosa, 2025; Winduastuti et al., 2026). However, comparisons should remain cautious because the current manuscript does not provide enough methodological detail to establish equivalence between interventions, assessment instruments, or study populations. The present results add a descriptive account of changes across multiple score dimensions and students' reported perceptions in one university course.

Krashen's Input Hypothesis provides a possible interpretation of the listening and pronunciation activities: students may have had repeated exposure to spoken models. Long's Interaction Hypothesis also helps explain why classroom speaking and presentation tasks may complement individual application use. However, automated corrective feedback should not be equated with negotiation of meaning in human interaction. The combined intervention is therefore pedagogically plausible, but the study does not test either theory or identify the contribution of each component.

Several limitations shape the interpretation. First, the sample consisted of 32 students from one course at one institution, limiting generalizability. Second, the absence of a control group prevents attribution of the observed changes to ELSA Speak alone. Third, the manuscript currently reports means without dispersion statistics or inferential analyses. Fourth, information about intervention dosage, rubric validation, inter-rater reliability, and questionnaire validation is incomplete. Finally, the questionnaire summary reports only positive categories; the original questionnaire and response records should be checked to ensure that the presentation is complete. These limitations should be addressed transparently rather than covered by stronger claims.

6. Conclusion

In this group of 32 undergraduate students, mean post-test scores were higher than pre-test means across pronunciation, fluency, vocabulary, grammar, and confidence after an eight-week intervention combining ELSA Speak-supported practice with classroom activities. Pronunciation showed the largest reported mean increase (+19.3 points), followed by confidence (+18.2), grammar (+16.3), fluency (+13.7), and vocabulary (+13.7). Questionnaire

Journal of Language
Volume 8 Number 2 pp: 343-350
e-ISSN: 2685-8878 | p-ISSN: 2655-9080
DOI: https://doi.org/10.30743/jol.v8i2.14234

summaries also showed favorable perceptions, with 78.1%–84.4% of respondents selecting Strongly Agree across the ten reported items.

These findings document positive descriptive changes and favorable student perceptions during the intervention period; they do not establish statistical significance or demonstrate that ELSA Speak alone caused the changes. For teaching practice, ELSA Speak may be considered a supplementary resource for repeated pronunciation practice and feedback, provided it is combined with teacher guidance and communicative tasks that develop fluency, vocabulary, grammar, and interactional competence. Educators should not use application feedback as the sole measure of overall speaking competence.

Future studies should use larger and more diverse samples, include a comparison group or a controlled design, report complete descriptive and inferential statistics, and establish the reliability and validity of speaking assessments and questionnaires. Research should also document intervention dosage and examine pronunciation, overall speaking performance, confidence, and speaking anxiety as distinct outcomes.

References

Aditya, N., Prayogo, D., & Anggara. (2024). Learning beyond the classroom: English clubs and students' speaking competence in an Indonesian EFL context. Jurnal Tahuri, 21(2), 137–154. https://doi.org/10.30598/tahurivol21issue2page137-154

Aulia, L. G. Y. M., & Santosa, M. H. (2025). AI in language education: Systematic review of ELSA Speak's impact on English. Jurnal Pendidikan Bahasa Inggris Indonesia, 13(1). https://doi.org/10.23887/jpbi.v13i1.3771

Brown, H. D. (2007). Principles of language learning and teaching (5th ed.). Pearson Education.

Crystal, D. (2003). English as a global language (2nd ed.). Cambridge University Press.

Daflizar, D., Sulistiyo, U., & Kamil, D. (2022). Language learning strategies and learner autonomy: The case of Indonesian tertiary EFL students. LEARN Journal: Language Education and Acquisition Research Network, 15(1), 257–281.

Hanafiah, W., Aswad, M., Sahib, H., Yassi, A. H., & Mousavi, M. S. (2022). The impact of CALL on vocabulary learning, speaking skill, and foreign language speaking anxiety: The case study of Indonesian EFL learners. Education Research International, 2022, Article 5500077. https://doi.org/10.1155/2022/5500077

Kılıç, A., & Uzun, L. (2022). The relationship between tertiary level EFL learners' attitudes towards English and technology and their autonomy levels. Journal of Language Education and Research, 8(2), 227–248. https://doi.org/10.31464/jlere.1073392

Krashen, S. D. (1985). The input hypothesis: Issues and implications. Longman.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), Handbook of second language acquisition (pp. 413–468). Academic Press.

Rose, H., McKinley, J., & Galloway, N. (2021). Global Englishes and language teaching: A review of pedagogical research. Language Teaching, 54(2), 157–189. https://doi.org/10.1017/S0261444820000518

Sun, L. (2023). Cultivating critical global citizens through secondary EFL education: A case study of mainland China. Literacy, 57, 249–261. https://doi.org/10.1111/lit.12314

Tarigan, R. M., & Lubis, Y. (2024). Qualitative findings on the impact of debate techniques on English language speaking competence. The Journal of Educational Development, 12(2). [Complete page range and verify publication details.]

Uchihara, T., & Clenton, J. (2023). The role of spoken vocabulary knowledge in second language speaking proficiency. The Language Learning Journal, 51(3), 376–393. https://doi.org/10.1080/09571736.2022.2080856

Winduastuti, N. K., Utami, I. G. A. L. P., & Artini, N. N. (2026). The integration of ELSA Speak on English pronunciation skills in EFL university students: A systematic literature review. Edu Society: Jurnal Pendidikan, Ilmu Sosial dan Pengabdian kepada Masyarakat, 5(3), 2212–2221. https://doi.org/10.56832/edu.v5i3.2534

Woo, J. H., & Choi, H. (2021). Systematic review for AI technology used in language learning. International Journal of Contents, 17(3), 1–11.

Zhang, J. (2023). EFL teachers' digital literacy: The role of contextual factors in their literacy development. Frontiers in Psychology, 14, Article 1153339. https://doi.org/10.3389/fpsyg.2023.1153339