THE EFFECTIVENESS OF STIMULER APPLICATION IN IMPROVING STUDENTS' ENGLISH SPEAKING SKILL AT MA NWDI KETANGGA

Kuspina Pebianti[1], Yulia Agustina[2], Siti Ayu Surayya[3]
Siti Maysuroh[4]
[1234]Department of English Language Education, Faculty of Language, Art, and Humanities, Hamzanwadi University
[1] febiantikusvina@gmail.com [2] rahestin@gmail.com,
[3] ssurayya@hamzanwadi.ac.id, [4] sitimaysuroh@hamzanwadi.ac.id

ABSTRACT

English speaking skill is essential yet challenging for EFL learners due to linguistic barriers and psychological constraints. This study aims to empirically investigate the effectiveness of the Stimuler Application as a learning medium in improving students' English speaking skills. Employing a quantitative true experimental design, specifically a pre-test and post-test control group design, this research involved 36 eleventh-grade students at MA NWDI Ketangga, divided equally into an experimental class and a control class. Data were collected through a speaking test assessing pronunciation, fluency, vocabulary, grammar, and comprehension, and subsequently analyzed using the Mann-Whitney U test. The results revealed a highly significant difference in speaking performance between the two groups (U = 0.0, Sig. < 0.001). The experimental class achieved a massive post-test mean score of 90.44, significantly outperforming the control class's mean of 71.56, with the most prominent improvements observed in the aspects of fluency and pronunciation. In conclusion, the Stimuler Application is highly effective in accelerating oral proficiency by providing high-frequency independent practice and immediate AI-driven diagnostic feedback, thereby successfully reducing speech hesitation and lowering the students' affective filter.

Keywords: EFL Learners, Mobile-Assisted Language Learning (MALL), Speaking Skill, Stimuler Application,

ABSTRAK

Keterampilan berbicara bahasa Inggris merupakan aspek esensial namun menantang bagi pembelajar EFL karena adanya hambatan linguistik dan kendala psikologis. Penelitian ini bertujuan untuk menguji secara empiris efektivitas Aplikasi Stimuler sebagai media pembelajaran dalam meningkatkan keterampilan berbicara bahasa Inggris siswa. Dengan menggunakan desain kuantitatif eksperimen murni (*pre-test and post-test control group design*), penelitian ini melibatkan 36 siswa kelas XI di MA NWDI Ketangga yang dibagi rata menjadi kelas eksperimen dan kelas kontrol. Data dikumpulkan melalui tes berbicara yang menilai aspek pelafalan, kelancaran, kosakata, tata bahasa, dan pemahaman, lalu dianalisis menggunakan

uji *Mann-Whitney U*. Hasil penelitian menunjukkan perbedaan performa berbicara yang sangat signifikan antara kedua kelompok (U = 0.0, Sig. < 0,001). Kelas eksperimen mencapai nilai rata-rata *post-test* sebesar 90,44, jauh melampaui kelas kontrol yang mencapai 71,56, dengan peningkatan paling menonjol pada aspek kelancaran dan pelafalan. Kesimpulannya, Aplikasi Stimuler sangat efektif dalam mengakselerasi kecakapan lisan dengan menyediakan praktik mandiri berfrekuensi tinggi dan umpan balik diagnostik langsung berbasis AI, sehingga berhasil mengurangi keraguan saat berbicara dan menurunkan kecemasan (*affective filter*) siswa.

Kata Kunci: Aplikasi Stimuler, Keterampilan Berbicara, Pembelajar EFL, Pembelajaran Bahasa Berbantuan Seluler (MALL).

A. Introduction

English speaking skill is an essential competency in English as a Foreign Language (EFL) learning that directly reflects students' communicative proficiency. In the context of secondary education, students are required not only to master grammar and vocabulary but also to participate actively and meaningfully in oral communication. However, mastering this productive skill is often the greatest challenge for learners as it requires the simultaneous integration of linguistic aspects, such as pronunciation and fluency, with psychological readiness during language production.

In reality, many EFL students still experience significant difficulties in producing spoken language naturally, ranging from high anxiety about making mistakes to a lack of interactive practice opportunities. Various previous studies confirm that linguistic barriers and psychological obstacles are the main factors contributing to students' low participation in speaking activities (Kusmaryani & Tanjung, 2023; Franscy & Ramli, 2022). Similar problems are specifically identified at MA NWDI Ketangga, where students' speaking performance is heavily constrained by limited language exposure and the sub-optimal availability of technology-based learning media capable of facilitating independent practice.

Bridging these issues is highly urgent both scientifically and practically, making the utilization of educational technology such as the Stimuler Application an innovative

solution. Stimuler provides a digital speaking practice platform specifically designed to train fluency, pronunciation, and vocabulary use through digital interaction mechanisms. Various preliminary studies indicate that the use of this application positively influences students' learning motivation and is proven effective in yielding greater improvements in speaking scores compared to conventional teaching methods (Sari & Rahman, 2025; Denistiani, 2025a).

Although the effectiveness of various language learning applications such as ELSA, Cake, and Hello English has been widely explored in previous research (Daulay et al., 2023; Sari et al., 2025; Sari et al., 2026), quantitative studies utilizing a true experimental design that specifically examine the Stimuler Application remain highly limited. The majority of previous studies tend to focus on analyzing students' perceptions and experiences, rather than proving causality and technical effectiveness. Therefore, this study aims to empirically investigate the significance of the effectiveness of the Stimuler Application as a learning medium in improving the English speaking skills of students at MA NWDI Ketangga.

The novelty of this article lies in the implementation of a true experimental design to precisely measure the effectiveness of the Stimuler Application on five fundamental aspects of speaking: pronunciation, fluency, vocabulary, grammar, and comprehension. The results of this study are expected to provide theoretical contributions that enrich the literature in the field of Mobile-Assisted Language Learning (MALL), while simultaneously offering practical implications for English educators in integrating digital media to create an adaptive, interactive, and student-centered learning ecosystem

B. Methods

This study employed a quantitative method utilizing a true experimental design, specifically the pre-test and post-test control group design, to investigate the effectiveness of the Stimuler Application in enhancing students' English speaking skills. This design was selected as it allows for a rigorous causal comparison between an experimental group receiving the technological intervention and a

control group receiving conventional instruction. The research was conducted at MA NWDI Ketangga, located in East Lombok, West Nusa Tenggara, over a period of five months from May to September 2026.

The population of this research comprised students of MA NWDI Ketangga in the 2026 academic year. Through the cluster random sampling technique, two classes were selected as the research sample, yielding a total of 36 participants. Class X A, consisting of 18 students, was assigned as the experimental class taught using the Stimuler Application, while Class X B, also comprising 18 students, served as the control class taught without the application.

Data were collected through a speaking test administered as both a pre-test and a post-test. The students' speaking performance was evaluated across five key indicators: pronunciation, fluency, vocabulary, grammar, and comprehension. Scoring was conducted using a 5-point Likert scale rubric adapted from the IELTS Speaking Band Descriptors. To ensure the quality of the measuring instrument, content validity was established through expert judgment by evaluating the test instructions, speaking prompts, and scoring rubric prior to implementation. Furthermore, inter-rater reliability was applied, involving two competent raters to minimize subjectivity and ensure consistency in assessing the students' oral performances. The data collection procedure systematically involved administering the pre-test to gauge initial speaking abilities, implementing the respective treatments to both groups, and concluding with the post-test to measure developmental outcomes.

The quantitative data obtained from the speaking tests were analyzed using the Statistical Package for the Social Sciences (SPSS) version 22. The analysis procedure commenced with calculating descriptive statistics, including mean scores and standard deviations, to depict the speaking achievement of each group. Prerequisite testing was then conducted, utilizing the Shapiro-Wilk test to assess data normality and Levene’s test to determine the homogeneity of variances. Finally, hypothesis testing was executed using an Independent Samples t-test at a significance level of 0.05 to ascertain whether a statistically significant

difference existed between the speaking skills of the experimental and control groups.

C. Result and Discussion

Result

Tabel 1 Descriptive Statistics of the Pre-test and Post-test Scores

Group	N	Pre-test Mean	Post-test Mean	Mean Gain	Std. Deviation (Post)
Experimental	18	63.33	90.44	27.11	**5.96**
Control	18	61.00	71.56	10.56	**3.11**

Table 1 demonstrates that both groups commenced the study with nearly identical speaking abilities, as indicated by the baseline pre-test means of 63.33 for the experimental group and 61.00 for the control group. However, following the intervention, the experimental class achieved a massive mean gain of 27.11 points, resulting in a post-test mean of 90.44. In contrast, the control group only experienced a moderate gain of 10.56 points, ending with a post-test mean of 71.56. This 18.88-point difference at the post-test stage clearly indicates that the Stimuler Application accelerated speaking proficiency at more than double the rate of conventional classroom instruction.

Table 2. Frequency of Students by Post-test Score Category

Category	Score Range	Experimental Class	Control Class
Very Good	85 - 100	15	**0**
Good	70 - 84	3	**17**
Fair	55 - 69	0	**1**
Poor	40 - 54	0	**0**
Very Poor	0 - 39	0	**0**

The distribution of post-test scores in Table 2 illustrates a stark contrast in skill mastery. After the treatment, 83.3% (15 out of 18) of the

students in the experimental class escalated into the "Very Good" category, demonstrating high-level speaking proficiency. Conversely, the control class peaked entirely within the "Good" (17 students) and "Fair" (1 student) tiers, completely failing to produce any student in the highest category. This distribution proves that the application does not merely improve scores marginally, but effectively pushes the majority of learners to the upper limits of the evaluation scale.

Table 3. Post-test Mean Score by Speaking Aspect

Speaking Aspect	Experimental	Control	Mean Difference
Fluency	4.69	3.64	**+1.05**
Pronunciation	4.56	3.53	**+1.03**
Vocabulary	4.50	3.64	**+0.86**
Grammar	4.47	3.58	**+0.89**
Comprehension	4.39	3.50	**+0.89**

Table 3 details the origin of the post-test score gap by breaking it down into specific speaking aspects. The experimental class systematically outperformed the control class across all five metrics. The most significant disparities emerged in Fluency (4.69 vs. 3.64) and Pronunciation (4.56 vs. 3.53). This data specifically confirms that the Stimuler application is highly effective in serving its core function: drilling phonological accuracy and reducing speech hesitation through continuous task repetition.

Table 4. Hypothesis Testing (Mann-Whitney U Test)

Comparison	U Value	Z Value	Sig. (2-tailed)
Experimental vs. Control	0.0	-5.222	**< 0.001**

Because preliminary analyses revealed the data violated normal distribution parameters, the non-parametric Mann-Whitney U test was applied to ensure strict methodological validity. As shown in Table 4, the test produced a U value of 0.0 and a Z value of -5.222 with a significance level of < 0.001. A U value of 0.0 is an extreme statistical rarity indicating

zero overlap between the two datasets; empirically, the lowest-scoring student using the Stimuler application still achieved a higher speaking score than the highest-scoring student taught via conventional methods.

Overall, the results conclusively reject the null hypothesis (H0) and confirm the alternative hypothesis (Ha). The data proves that there is a highly significant, empirical difference in English speaking skill between the two groups. The Stimuler Application proved to be vastly superior to conventional teaching methods, driving massive improvements across all evaluated speaking metrics—most notably in fluency and pronunciation—and successfully elevating the majority of the students to the highest speaking proficiency tier.

Discussion

The empirical superiority of the experimental group confirms the efficacy of integrating Artificial Intelligence (AI) within Mobile-Assisted Language Learning (MALL) frameworks. The distinct enhancement in fluency validates the cognitive theory of language acquisition regarding task repetition. By forcing continuous oral production and active engagement, the Stimuler Application effectively reduces cognitive hesitation during speech, compensating for the inherent limitation of individual speaking time in conventional classroom settings.

These findings strongly align with and expand upon previous literature. While Denistiani (2025a) proved the effectiveness of the Stimuler Application in a vocational school setting, this study extends its proven applicability to the Islamic senior high school context, demonstrating its versatility across different educational environments. Furthermore, whereas Sari and Rahman (2025) highlighted the positive perceptions and experiences of high school students using Stimuler, this research provides the definitive quantitative causality that backs those subjective perceptions. The acceleration in phonological accuracy also mirrors the findings of Sharma et al. (2024) and Kusmaryani & Tanjung (2023) regarding mobile AI tools, specifically corroborating that AI-driven immediate diagnostic feedback corrects errors significantly faster than

delayed, generalized feedback from a human teacher.

From a socio-psychological perspective, the intervention successfully lowered the students' affective filter. A persistent barrier in EFL classrooms is the paralysis caused by high anxiety and the fear of peer judgment (Franscy & Ramli, 2022). The application functioned as a non-judgmental environment, enabling students to build self-confidence through private, low-stress practice. Interestingly, the intervention also triggered passive-to-active linguistic mapping; the necessity to process audio prompts and formulate coherent responses allowed students to unconsciously solidify their grammatical and comprehension skills without explicit traditional instruction.

Despite these robust outcomes, a critical analysis necessitates acknowledging the potential presence of a "Novelty Effect". The sudden integration of advanced smartphone-based AI into the standard curriculum likely generated a temporary surge in motivation and enthusiasm, which may have catalyzed the rapid progress. While conventional instruction remains functional and necessary for foundational language building, it simply lacks the high-frequency practice mechanics afforded by this specific technological intervention.

D. Conclusion

This research establishes that the integration of the Stimuler Application is significantly more effective than conventional instructional methods in developing students' English speaking skills. The quantitative evidence definitively shows that consistent exposure to the application's AI-driven platform dramatically accelerates oral proficiency, driving massive improvements across fundamental speaking metrics, most notably in minimizing speech hesitation (fluency) and enhancing phonological accuracy (pronunciation). Ultimately, the intervention successfully elevated the vast majority of the learners to the highest speaking proficiency tier, confirming the application's robust capability to facilitate active language production.

Theoretically, these findings validate the core premises of Mobile-Assisted Language Learning (MALL)

and cognitive theories of language acquisition, demonstrating that high-frequency task repetition coupled with immediate AI diagnostic feedback effectively reduces the cognitive gap in error correction. Furthermore, it highlights the technology's psychological benefit in lowering the affective filter, allowing students to build self-confidence in a non-judgmental, low-anxiety environment. Practically, this study provides a clear, scalable blueprint for English educators: AI applications should not be viewed as replacements for human instruction, but rather strategically integrated as structured, out-of-class extensions to overcome the inherent constraints of limited classroom speaking time.

Scientifically, this research contributes essential empirical rigor to the current literature by shifting the focus from subjective learner perceptions to objective, causal evidence of algorithmic efficacy. By employing a true experimental design, this study provides concrete proof of the Stimuler Application's pedagogical value in secondary EFL contexts. To further validate and expand upon these findings, future research should incorporate delayed post-tests to measure the long-term cognitive retention of the acquired speaking skills. Moreover, exploring larger, demographically diverse samples and conducting direct comparative analyses between various AI-driven speaking platforms will be vital in identifying the most optimal technological frameworks for diverse educational settings.

REFERENCES

Aini, N., Amalia, F., & Ningrum, A. S. B. (2022). Improving Students' Speaking Skill Using Hello English Application as a Medium of Learning from Home. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, 10(1), 730–745. https://doi.org/10.24256/ideas.v10i1.2533

Baron, R. (2020). Students' Perception on Online Application in Speaking Skill. *VELES Voices of English Language Education Society*, 4(2), 213–221. https://doi.org/10.29408/veles.v4i2.2543

Bozorgian, H., & Kanani, S. M. (2017). Task Repetition on Accuracy and Fluency: EFL Learners Speaking Skill. *International Journal of English Language and Literature Studies*, 6(2), 42–53. https://doi.org/10.18488/journal.23.2017.62.42.53

British Council, IDP IELTS, & Cambridge Assessment English. (2025).

Daulay, S. H., Siregar, D. Y., & Sulistia, N. (2023). Hello English Application to Improve the Students' Speaking Skills. *Edulitics (Education, Literature, and Linguistics) Journal*, 8(2), 21–28. https://doi.org/10.52166/eduliticcs.v8i2.4798

Denistiani, S. A. (2025a). The Effectiveness of Stimuler Application as Learning Media to Improve Students' Speaking Skills at Vocational School. *MATCHA: Journal of Modern Approaches to Communication, Humanities, and Academia*, 1(1), 21–30. https://doi.org/10.70152/matcha.v1i1.135

Denistiani, S. A., Asipi, L. S., & Nopiyadi, D. (2025). The Effectiveness of Using Stimuler Application and Vocabulary Mastery on Speaking Skills in Cirebon.

Eldania, I. (2024). The Effect of Cake Application on Students' Speaking Skill At the Eleventh Grade.

Franscy, F., & Ramli, R. (2022). Problems Encountered by Indonesian EFL Learners in Mastering Speaking Skills. *Pioneer: Journal of Language and Literature*, 14(1), 1–14. https://doi.org/10.36841/pioneer.v14i1.1176

Hasbi, M., & Nursaputri, E. (2024). Using ELSA Speak Application as A Medium to Improve English Speaking Skills. *International Journal of Language, Humanities, and Education*, 7(1). *(Halaman atau DOI belum tercantum)*

Kusmaryani, W., Musthafa, B., & Purnawarman, P. (2019). The influence of mobile applications on students' speaking skill and critical thinking in English language learning. *Journal of Physics: Conference Series*, 1193(1). https://doi.org/10.1088/1742-6596/1193/1/012008

Kusmaryani, W., & Tanjung. (2023).

Normawati, A., Nugrahaeni, D. A., Kusuma Hadi Manggolo, N. S., & Susanto, A. I. F. (2023). EFL Learners' Difficulties in Speaking English. *English Language and Education Spectrum*, 1(1), 1–9. https://doi.org/10.53416/electrum.v1i1.116

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

Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, 14. https://doi.org/10.3389/fpsyg.2023.1255594

Sabila, A., Cahyono, B. Y., & Anugerahwati, M. (2024). Asynchronous Video Feedback as a Game-Changer for Online Speaking Classes. *Voices of English Language Education Society*, 8(2), 360–370. https://doi.org/10.29408/veles.v8i2.24708

Sari, K., & Rahman, B. I. (2025). Stimuler Application For

Speaking Class: Senior High School Students' Experiences. *Jle: Journal of Literate of English Education Study Program*, 6(1), 25–38.

Sari, K., et al. (2026).

Shafipour, R., Baten, R. A., Hasan, M. K., Ghoshal, G., Mateos, G., & Hoque, M. E. (2018). Buildup of speaking skills in an online learning community: a network-analytic exploration. *Palgrave Communications*, 4(1). https://doi.org/10.1057/s41599-018-0116-6

Sharma, S., Mhasakar, M., Mehra, A., Venaik, U., Singhal, U., Kumar, D., & Mittal, K. (2024). Comuniqa: Exploring Large Language Models For Improving English Speaking Skills. In *COMPASS 2024 - Proceedings of the ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies* (Vol. 1). Association for Computing Machinery. https://doi.org/10.1145/3674829.3675082

Suryadi, S., & Erlangga, F. (2024). Enhancing English Speaking Skills through Learning Media and Discourse Marker Mastery: A Quasi-Experimental Study. *Voices of English Language Education Society*, 8(1), 271–281. https://doi.org/10.29408/veles.v8i1.25314