Indonesian Language Education and Applied Linguistics Reviews
Volume 2, Issue 1, 2026, 9-16
https://doi.org/10.66272/t2gadx67

MEDIA AKADEMIKA PUBLISHER
ISSN 3110-5904

Available online at:
https://journal.mediaakademika.id/index.php/ileal/article/view/76

Measuring Speaking Gains from HelloTalk Use in Eleventh Grade: Quasi-Experimental Evidence

Ardhita Chairunnisa[1*], Husnaini[2], Darmawan Budiyanto[3], Heru Setiawan
[1*]Universitas Tridinanti
**Corresponding Author:* ardhitachairunnisa@gmail.com

ARTICLE INFO

Received:
28 November 2025

Revised:
12 May 2026

Accepted:
23 May 2026

Keywords:
Speaking gain, HelloTalk

ABSTRACT

The objective of this study was to evaluate the effectiveness of the HelloTalk application in improving the speaking proficiency of eleventh-grade students at SMA Negeri 4 Palembang. A quantitative quasi-experimental design was employed, involving 120 students selected through purposive sampling and divided into experimental and control groups. The study used a pretest-posttest design to measure fluency, accuracy, grammar, mechanics, and pronunciation during an eight-week intervention with HelloTalk. The instrument's content validity was confirmed using expert judgment and Aiken's V formula, while data analysis was conducted through SPSS 24. The Kolmogorov-Smirnov test confirmed normal data distribution ($p > 0.05$). An independent sample t-test revealed a statistically significant difference in speaking performance between students taught using HelloTalk and those without, with a p-value of 0.000 (< 0.05). The findings indicate that HelloTalk substantially enhanced students' speaking proficiency, particularly in fluency and self-expression, after consistent use over eight weeks. These results highlight the application's positive impact on EFL learning and align with recent research promoting mobile and AI-assisted tools for language development. Thus, HelloTalk proves to be an effective platform to support formal classroom instruction in improving students' English-speaking skills.

How to Cite: Chairunnisa, A., Husnaini, Budiyanto, D., & Setiawan, H. (2026). Measuring Speaking Gains from HelloTalk Use in Eleventh Grade: Quasi-Experimental Evidence. *Indonesian Language Education and Applied Linguistics Reviews, 2*(1), 9-16. https://doi.org/10.66272/t2gadx67

Published by:
Media Akademika Publisher
mediaakademikapublisher@gmail.com

1. INTRODUCTION

The globalization era and rapid technological development have transformed how people interact, communicate, and learn. In particular, the integration of Artificial Intelligence (AI) into educational technologies has reshaped the dynamics of language learning, offering personalized and adaptive experiences beyond traditional classrooms. As learners increasingly rely on digital tools for academic and professional purposes, AI-powered platforms are becoming indispensable in bridging gaps between formal instruction and authentic communication practices. These tools not only provide real-time feedback but also create learning environments that simulate natural interactions, which are often missing in conventional EFL contexts.

Parallel to this development, the role of mobile-assisted language learning (MALL) has gained momentum. Applications such as HelloTalk, Duolingo, and Busuu exemplify how learners can access authentic interaction anytime and anywhere (Fauzi et al., 2024; Yorlanda & Abbas, 2023). Unlike traditional classroom settings where practice opportunities are limited, these platforms enable learners to engage with native speakers and peers worldwide, making learning more interactive, contextual, and meaningful. The growing adoption of MALL underscores a paradigm shift toward learner autonomy and self-directed learning, aligning with constructivist perspectives on language acquisition.

English proficiency, especially in speaking, remains essential for global communication, academic success, and professional advancement. Speaking competence allows learners to articulate ideas, negotiate meaning, and participate in multicultural environments where English functions as a lingua franca (Harmer, 2015). However, in many EFL contexts such as Indonesia, learners often encounter limited access to authentic communication partners and tend to rely heavily on grammar-based instruction (Nunan, 2015). This imbalance results in passive learning habits and insufficient oral proficiency, despite years of classroom exposure to English.

To address these challenges, innovative pedagogical approaches must integrate authentic communication opportunities with classroom instruction. AI and mobile-based platforms can bridge this gap by enabling real-world practice within safe and supportive environments. They also motivate learners by offering instant corrections, gamification, and culturally diverse interactions. For high school students, who are often constrained by rigid curricula and exam-focused teaching, such tools provide valuable chances to build confidence, improve fluency, and develop communicative competence in English.

Despite the growing use of mobile and AI-mediated tools, research in Indonesian secondary education remains scarce. Most empirical studies have focused on higher education contexts, where learners have greater autonomy and access to technology (Damayanti et al., 2024). For example, Rosilah and Ulfa (2023) found that university students using HelloTalk improved their fluency and intercultural awareness. Similarly, Kasim et al. (2022) emphasized the potential of mobile-assisted tools in promoting collaborative learning. However, these findings may not fully represent the realities of high school classrooms, where technological access, teacher facilitation, and curriculum constraints differ significantly.

This study addresses that gap by evaluating the effectiveness of HelloTalk in enhancing speaking proficiency and exploring students' perceptions of its use in an Indonesian senior high school. Unlike previous research, this investigation combines performance-based outcomes with qualitative insights to capture not only improvements in oral proficiency but also learners' attitudes, challenges, and engagement levels. By situating the study in a secondary school context, the research contributes to filling a critical void in EFL literature, offering both theoretical enrichment and pedagogical implications. Ultimately, the findings are expected to provide practical insights for educators seeking to integrate innovative, AI-powered tools into classroom instruction, supporting a more interactive and authentic model of language learning (Stockwell & Hubbard, 2013; Kukulska-Hulme, 2020).

2. METHODS

This study utilized quantitative methodology trough quasi experimental approach to comprehensively evaluate the effectiveness of the HelloTalk application in enhancing the speaking ability of eleventh-grade students at SMA Negeri 4 Palembang. The research design enabled an in-depth examintation of HelloTalk's role in language development, while also capturing the students' experiences and perceptions regarding the application use in language learning (Creswell et al., 2018; Fraenkel et al., 2019 and Cohence et al., 2020).

This study employed a pretest-posttest design to measure the improvement in students' speaking ability (Brown et al. 2020). The pretest and posttest consisted a series of speaking tasks designed to assess the students' fluency, accuracy, and confidence in speaking English. The tasks encompassed a range of subjects related to daily life, personal experiences, and hypothetical scenarios. The pretest was administered at the beginning of the study, while the posttest was conducted after eight weeks of utilization of the HelloTalk application for the purpose language practice. The pretest and posttest results were utilized to assess the enhancement in students' speaking ability. The tests conducted by rubric of scoring tests that focused on fluency, accuracy, grammatical, mechanism, and pronunciation. These components are considered fundamental to assessing speaking proficiency. The results were analyzed using descriptive and inferential statistics, including paired sample t-tests. The objective of this analysis was to determine whether there was a significant improvement in students' speaking ability after using the application.

The participants in this study were 120 eleventh-grade students of SMA Negeri 4 Palembang, Indonesia, who were enrolled in the English language program. The students were

selected through the use of purposive sampling, a non-probability sampling technique that is frequently employed in educational research when researchers aim to select participants with specific characteristics relevant to the study (Etikan, 2021). The sampling was based on recommendations from their English teacher, focusing on students whose speaking scores were below the minimum competency standard, indicating the need for additional speaking support. Participants were also selected based on their availability and willingness to participate in the six-week intervention.

The data were collected through a tests designed to assess students' speaking abilities. The test content validity was established through experts judgement, and reliability wa determined using Aiken's V formula to calculate the degree of agreement among the experts (Aiken, 1985). The final Aiken's V score indicated that the test items were valid and suitable for measuring speaking performance. Additionally, the data were analyzed using the SPSS 24 application to ensure the statistical accuracy and reliability in the findings (Field, 2024).

3. RESULTS

3.1. Result of Normality Test

Normality of Using HelloTalk App to Speaking Ability

Testing the normality of the students' speaking ability using Kolmogorov-Smirnov test showed that the significant p values of the pretest and post-test scores in the control and experimental groups were higher than the required threshold of 0.05. Thus, it can be concluded that the data followed a normal distribution. This information is presented in table 1.

Table 1. One-Sample Kolmogorov-Smirnov Test

		Pretest Experimental Class	Pretest Control Class	Posttest Experimental Class	Posttest Control Class
N		60	60	60	60
Normal Parameters[a,b]	Mean	65.55	54.02	80.78	69.72
	Std. Deviation	6.598	8.243	7.614	7.121
Most Extreme Differences	Absolute	.104	.112	.113	.088
	Positive	.057	.056	.042	.079
	Negative	-.104	-.112	-.113	-.088
Test Statistic		.104	.112	.113	.088
Asymp. Sig. (2-tailed)		.173[c]	.060[c]	.053[c]	.200[c,d]

a. Test distribution is Normal.
b. Calculated from data.
c. Lilliefors Significance Correction.
d. This is a lower bound of the true significance.

The normality of the data was evaluated using the One-Sample Kolmogorov Smirnov Test for all pre-test and post-test scores in the experiment and control class (N=60).

Pre-Test Experiment Class:

The mean score for the pre-test experiment class was 65.55, with a standard deviation of 6.598. Absolute difference = 0.104, positive difference = 0.057, and negative difference = -0.104 were the most pronounced discrepancies between the observed and predicted cumulative distributions. With a 2-tailed asymptotic significance value of 0.173, the test statistic was 0.104. A normal distribution of the data is indicated by a significance value greater than 0.05 ($p > 0.05$).

Post-Test Experimental Class:

The standard deviation was 7.614 and the mean score was 80.78. Absolute difference = 0.113, positive difference = 0.042, and negative difference = -0.113 were the most pronounced discrepancies between the observed and predicted cumulative distributions. The asymptotic significance value (2-tailed) was 0.053, and the test statistic was 0.113. Even though this result is

quite near the threshold, it is still over 0.05, indicating that the data had a normal distribution (p > 0.05).

Pre-Test control Class:

The standard deviation was 8.243 and the mean score was 54.02. Absolute difference = 0.112, positive difference = 0.056, and negative difference = -0.112 were the most dramatic differences. The asymptotic significant value (2-tailed) was 0.060, and the test statistic was 0.112. Despite being very near the significance level, this result shows that the data were normally distributed (p > 0.05).

Post-Test Control Class:

The mean score for the post-test control class was 69.72, with a standard deviation of 7.121. Absolute difference = 0.088, positive difference = 0.079, and negative difference = -0.088 were the most dramatic differences. With a 2-tailed asymptotic significance value of 0.200, the test statistic was 0.088. Given that this value is significantly higher than the 0.05 cutoff, the control group's post-test results appear to have a normal distribution (p > 0.05).

All pre-test and post-test score distributions followed a normal distribution, according to the results of the normality test for both the experimental and control classes (p > 0.05). The experimental class's mean score improved from 65.55 on the pre-test to 80.78 on the post-test. In a similar vein, the control group's mean score improved from 54.02 on the pre-test to 69.72 on the post-test. The assumption of normalcy for further statistical analysis was confirmed by the Kolmogorov-Smirnov tests, which yielded all significance values over 0.05.

3.2. Result of Homogeneity Test

The homogeneity check through SPSS 24 in students' speaking ability using levene Statistic shows that the p value obtained from the test conducted on the variance of all pre-test and post-test scores is higher than 0.05. This shows that stutendts' speaking ability in the pretest and post-test groups can be said to be homogeneous. The significant p-value of the reading test also exceeded the threshold for significant at 0.05 level.

Table 2. Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
Score	Based on Mean	1.696	3	236	.169
	Based on Median	1.394	3	236	.245
	Based on Median and with adjusted df	1.394	3	231.484	.245
	Based on trimmed mean	1.659	3	236	.176

Pre-Test:

The Levene Statistic was 1.696 with degrees of freedom (df1 = 3, df2 = 236), and a significance value of 0.169. This indicates that the variances of the pre-test scores are equal across groups (p > 0.05).

Post-Test:

The Levene Statistic was 1.659 with degrees of freedom (df1 = 3, df2 = 236), and a significance value of 0.176. This result also confirms equal variances for the post-test scores across groups (p > 0.05).

Based on these results, the assumption of homogeneity of variances is satisfied for both pre-test and post-test scores.

3.3. Result of Paired Sample t-Test

The results of the paired sample t-test can be seen in Table 3 below:

Table 3. Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Pretest Experimental Class - Posttest Experimental Class	-15.233	9.414	1.215	-17.665	-12.801	-12.53	59	.000
Pair 2	Pretest Control Class - Posttest Control Class	-15.700	11.932	1.540	-18.782	-12.618	-10.19	59	.000

We used the paired samples t-test to look at the scores from the pre-test and post-test in both the experimental and control groups.

Pair 1: Pre-test vs. Post-test (Experimental Class)

Considering a standard deviation of 9.41 and a standard error of the mean of 1.215, the experimental class's mean difference between pre-test and post-test scores was -15.233. The difference's 95% confidence interval was between -17.665 and -12.801. With degrees

of freedom (df = 59), the t-value was -12.534, and the two-tailed significance value was 0.000. The results show that the difference between the pre-test and post-test scores is statistically significant ($p < 0.05$). The results show that after receiving treatment in the experimental class, pupils' scores significantly improved.

Pair 2: Pre-test vs. Post-test (Control Class)

Based a standard deviation of 11.93 and a standard error of the mean of 1.540, the control class's mean difference between pre-test and post-test scores was -15.700. The difference's 95% confidence interval was between -18.782 and -12.618. With degrees of freedom (df = 59), the t-value was -10.192, and the two-tailed significance value was 0.000. Additionally, this result shows a statistically significant difference ($p < 0.05$) between the pre-test and post-test scores, indicating that the control group students also improved during the learning session. These results showed that scores improved statistically significantly from the pre-test to the post-test for both the experimental and control groups.

3.4. Result of Independent Sample t-Test

When evaluating their hypotheses, the researchers used the results of statistical analysis to help them make sure they were answering the study topic. The purpose of the test was to ascertain whether the speaking abilities of the students who received instruction at SMA Negeri 4 Palembang using the HelloTalk program and those who did not differed significantly. The following were the hypotheses for this test:

Ho: Students that received instruction at SMA Negeri 4 Palembang using the HelloTalk app did not significantly differ in their speaking abilities from those who did not.

Ha: Students that received instruction at SMA Negeri 4 Palembang using the HelloTalk app differed significantly from those who did not in terms of their speaking abilities.

Testing criteria:
If p-value < 0.05, then Ha is accepted and Ho is rejected.
If p-value > 0.05, then Ho is accepted and Ha is rejected.

Based on the initial research question, the researchers examined whether there were significant differences in speaking ability among tenth-grade students who used the HelloTalk app compared to those who did not at SMA Negeri 4 Palembang. The analysis revealed a significant difference (p = 0.309) which exceeds the 0.05 significance level. Therefore, the null hypothesis is accepted, indicating no significant difference in speaking ability between the two groups.

Further details of the independent sample t-test calculations are presented in Table 4.

Table 4. Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Nilai	Equal variances assumed	.044	.834	8.223	118	.000	11.067	1.346	8.401	13.732
	Equal variances not assumed			8.223	117.477	.000	11.067	1.346	8.401	13.732

4. DISCUSSIONS

After analyzing all the information, the researchers have reached a variety of interpretations related to the effects of Hellotalk app to improve their students' speaking ability. Based on the results of the independent sample t-test, it was found that there was a statistically significant difference in speaking ability between students who were taught using HelloTalk and those who were not, with a p-value of 0.000 (< 0.05). This indicates that the HelloTalk application positively impacted the speaking performance of the eleventh-grade students at SMA 4 Palembang.

This result is consistent with earlier research by Rosilah & Ulfa (2024) and Damayanti et al. (2024), which discovered that HelloTalk may greatly increase learner confidence and fluency. Additionally, the experimental group's kids demonstrated a notable gain in their ability to articulate concepts, pronounce words correctly, and use proper grammar. They frequently found it difficult to answer confidently or have smooth interactions prior to utilizing HelloTalk. Their speaking performance did, however, significantly improve after eight weeks of regular use of the app, especially in the areas of fluency and self-expression.

This development backs up Harmer's (2015) theoretical viewpoint, which highlighted the need of incorporating meaningful engagement and communicative practices into language learning. In a similar vein, Nunan (2015) maintained that learner-centered, real-world communication-focused language instruction is essential. By giving students access to real-world language contexts, the HelloTalk app helped close the gap that EFL learners frequently encounter in non-native contexts (Godwin-Jones, 2018).

Therefore, it can be said that HelloTalk not only inspired the students and produced a more engaging and learner-centered environment, but it also enhanced their speaking abilities in terms of quantifiable performance. These findings corroborate the increasing amount of research that suggests using mobile and AI-powered platforms to enhance speaking development in formal EFL learning environments.

5. CONCLUSION

The purpose of this study was to assess how well the HelloTalk app improved the speaking skills of eleventh-grade students at SMA Negeri 4 Palembang. The use of HelloTalk was determined to considerably increase students' speaking abilities based on the findings of statistical analysis, including normality, homogeneity, paired sample t-tests, and independent sample t-tests. When taught via HelloTalk, students in the experimental group outperformed those in the control group in terms of fluency, accuracy, and confidence. The results are consistent with earlier research and professional opinions highlighting the advantages of AI-powered and mobile-assisted language learning resources. HelloTalk helped close the gap between classroom education and real-world communication by offering a dynamic, authentic, and learner-centered environment. This study contributes to the growing evidence supporting the integration of technology into English language teaching. Teachers are encouraged to consider mobile learning applications like HelloTalk as effective tools for improving students' oral communication skills, especially in settings where access to native speakers and real-world practice is limited

6. REFERENCES

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

Brown, H. D., Abeywickrama, P., & Saito, Y. (2020). *Language Assessment: Principles and Classroom Practices (3rd ed.).* Pearson Education.

Cohen, L., Manion, L., & Morrison, K. (2020). *Research Methods in Education (8th ed.*). Routledge.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Damayanti, A., Hifdil, M., & Hamdani, A. B. (2024). Indonesian EFL students' perception on Hello Talk application in fostering speaking. *English Language Teaching for EFL Learners*, 6(1), 18-32.

Etikan, I. (2021). Sampling and sampling methods. *Biometrics & Biostatistics International Journal, 5(*6), 00149.

Fauzi, I., Daulay, S. H., Dewi, U., & Irham, N. R. (2024). Indonesian elementary students' achievement in learning English speaking through viewing Hello Talk: The effectiveness. *Jurnal Inovasi Teknologi Pendidikan*, *11*(3), 246-255.

Field, A. (2024). *Discovering Statistics Using IBM SPSS Statistics (6th ed.).* Sage Publications.

Godwin-Jones, R. (2018). *Using mobile devices in the language classroom*. Cambridge University Press.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2019). *How to design and evaluate research in education* (10th ed.). McGraw-Hill Education.

Harmer, J. (2015). *The practice of English language teaching* (5th ed.). Pearson Education.

Kasim, A. A. M., Darus, N. A., Lee, N. A. A., Subramaniam, A. L. B., & Januin, J. (2022). Online collaborative performance in group-based tasks among learners of higher education. *Studies in English Language and Education*, 9(3), 948-966.

Kukulska-Hulme, A., & Lee, H. (2020). Intelligent assistants in language learning: An analysis of features and limitations. In K.-M. Frederiksen, S. Larsen, L. Bradley, & S. Thouësny (Eds.), *CALL for widening participation: Short papers from EUROCALL 2020* (pp. 172-176). Research-publishing.net. https://doi.org/10.14705/rpnet.2020.48.1184

Nunan, D. (2015). *Teaching English to speakers of other languages: An introduction*. Routledge.

Rosilah, I., & Ulfa, S. M. (2024). The use of Hello Talk application for speaking skills: A case study of online language learners for senior high school. *Edulitics (Education, Literature, and Linguistics) Journal*, *9*(1), 33-40.

Stockwell, G., & Hubbard, P. (2013). Some emerging principles for mobile-assisted language learning. *The international research foundation for english language education*, *2013*, 1-15.

Yorlanda, N., & Abbas, M. (2022). Improving Students' Pronunciation, Vocabulary, Fluency, and Comprehension through HelloTalk Application. *Journal of Applied Linguistics and Education, 6*(3), 234–247.