INTERACTION: Jurnal Pendidikan Bahasa Vol. 13, No. 3; September 2026
P-ISSN: 2406-9558; E-ISSN: 2406-9566

Conversation Practice as a Strategy Enhancing Speaking Performance Among First-Semester Students of the Primary School Teacher Education Program at Universitas Halmahera

Prianka Lusy Koraag
pkoraag@gmail.com

Like Suoth
suothlike@gmail.com

Universitas Halmahera

ABSTRACT

This study investigated changes in overall speaking-performance scores among first-semester students of the Primary School Teacher Education (PGSD) Program at Universitas Halmahera after participation in structured conversation practice during the 2025/2026 academic year. A quantitative one-group pre-test/post-test design was used with 36 students. Speaking performance was operationalized through four dimensions, each scored on a five-point analytic rubric and converted to a 100-point total score. The mean score increased from 55.36 (SD = 2.37; range = 51.5 – 59.0) at pre-test to 75.19 (SD = 2.61; range = 71.0 – 79.0) at post-test, representing a mean increase of 19.83 points, or 35.83% relative to the pre-test mean. The Shapiro – Wilk test showed that paired difference scores were not normally distributed, $W = 0.929$, $p = 0.023$; consequently, the two-sided Wilcoxon signed-rank test was used. The test indicated a statistically significant difference between pre-test and post-test scores, $W = 0$, $z = -5.23$, $p < .001$, with a large within-cohort effect size ($r = 0.87$). The findings indicated that students' speaking-performance scores were higher after the instructional period. Because the study did not include a comparison group, the observed change cannot be attributed to conversation practice alone. The findings therefore constitute preliminary within-cohort evidence and should be examined in controlled or quasi-experimental studies with independent raters and stronger instrument-validation procedures.

Key words: Speaking Performance; Conversation Practice; Oral Interaction

INTRODUCTION

English proficiency is an important component of academic and professional preparation in higher education. For students in Primary School Teacher education (PGSD), the ability to communicate in English is particularly relevant because their future work may require them to explain ideas, interact with learners, and respond to changing educational demands. Speaking therefore represents not merely a language-learning target but also a practical competence that can support student's academic participation and future professional responsibilities. Speaking is a productive skill that requires learners to formulate and deliver language in real time. Successful oral communication depends on several related capacities, including

continuity of speech, intelligible pronunciation, appropriate responses to an interlocutor, and sufficient lexical resources. Consequently, speaking performance should be viewed as multidimensional rather than reduced to fluency alone. Research on task repetition has shown relationships between repeated oral performance and aspects of fluency and lexical use (Dawadi, 2019; Duong et al., 2023; Suzuki, 2021), while vocabulary knowledge is also associated with second-language speaking ability (Uchihara & Clenton, 2020).

Opportunities for meaningful interaction are also important. Peer interaction allows learners to listen, respond, negotiate meaning, and sustain exchanges rather than only reproduce isolated sentences (Soh et al., 2021). Students' willingness to communicate may be shaped by classroom conditions, and research has identified factors that can either support or inhibit participation in English-speaking activities (Liu et al., 2021; Ramli et al., 2021; Weda et al., 2021). Speaking anxiety may further constrain oral participation in EFL contexts (Andini et al., 2022; Hanafiah et al., 2022; Jannah et al., 2022; Yoskapela et al., 2022). These findings collectively suggest the value of classroom activities that provide repeated, structured, and supportive opportunities for oral production. In the local context, classroom observations reported in the original study indicated hesitation, extended pauses, limited vocabulary, pronunciation difficulties, and occasional switching to Indonesian when students struggled to express their ideas. These observations were treated as contextual information rather than quantified baseline evidence because they were based on classroom observations rather than standardized measurement. Structured conversation practice offers a plausible instructional route because it combines repeated oral production with interaction, vocabulary retrieval, pronunciation practice, role-play, and feedback.

Previous classroom studies have reported positive speaking-related outcomes following pair work, short conversation activities, guided conversation, and authentic online conversation (Mariam et al., 2022; Mariana et al., 2024; Tran et al., 2024; Yamazaki, 2023). At the same time, findings from a one-group pre-test/post-test design require caution because changes may also reflect testing effects, maturation, concurrent coursework, history, regression toward the mean, or rating effects. The research gap in this study is consequently not limited to the absence of evidence from Universitas Halmahera. Rather, the study examines how an intact first-semester PGSD cohort changes across two measurement occasions when speaking performance is represented by a composite of four observable dimensions: fluency, pronunciation, interaction skills, and vocabulary. This measurement-oriented focus is intended to distinguish overall speaking performance from fluency as a single construct.

To guide the analysis and provide a clear direction for the study, the research question and hypotheses were formulated based on the research objective and the one-group pre-test/post-test design. Is there a statistically significant difference in students' overall speaking-performance scores between the pre-test and post-test following participation in the structured conversation-practice program? H_0: There

is no statistically significant difference between the students' pre-test and post-test overall speaking-performance scores. H_1: There is a statistically significant difference between the students' pre-test and post-test overall speaking-performance scores.

LITERATURE REVIEW

Speaking performance in this study refers to the ability to formulate and communicate ideas orally in real time. It is represented by four dimensions. Fluency concerns continuity of speech and the management of hesitation and pauses. Pronunciation concerns intelligibility and the extent to which pronunciation supports meaning. Interaction skills involve initiating, responding to, maintaining, and developing an exchange. Vocabulary concerns the selection and use of words that are relevant and sufficient for communicating intended meaning. These dimensions are conceptually related but should not be treated as interchangeable; therefore, their combined score is interpreted as overall speaking performance.

Structured conversation practice is defined here as repeated opportunities to use English through guided interaction that progressively moves toward more independent communication. Its potential mechanisms include repeated oral production, retrieval of lexical items, peer interaction, feedback, and a gradual reduction in planning support. Task-repetition research suggests that repeated performance can affect aspects of L2 oral production, including fluency and lexical use (Dawadi, 2019; Duong et al., 2023; Hunter, 2024; Sun & Révész, 2021; Suzuki, 2021). In the present intervention, these mechanisms were operationalized through modelling, guided conversation, pair and small-group practice, role-play, question-and-answer activities, independent conversation, and feedback.

Each mechanism has a plausible relationship with the measured dimensions. Repeated oral production may provide additional opportunities to formulate responses and manage hesitation, supporting fluency. Interaction requires students to listen, respond, ask questions, and sustain an exchange, which is relevant to interaction skills. Repeated retrieval of words in familiar contexts may support lexical access and vocabulary use (Uchihara & Clenton, 2020). Teacher modelling and feedback provide opportunities to notice and reproduce pronunciation features. These mechanisms are theoretically plausible, but the present design does not allow the contribution of any single mechanism to be isolated.

Previous Studies and Research Gap

Existing studies generally indicate that conversation-oriented activities can support oral-language development, although contexts and outcome measures differ. Yamazaki (2023) examined pair-work activities in a Japanese university context, whereas Mariam et al. (2022) investigated short conversation activities and Mariana et al. (2024) examined guided conversation. Tran et al. (2024) considered authentic online conversation using video-conferencing tools. Together, these studies support

the pedagogical relevance of interaction-rich practice, but they do not establish that the same pattern will occur in a first-semester PGSD cohort at Universitas Halmahera. The present study addresses this contextual and measurement gap by examining change in a composite speaking-performance score while explicitly limiting causal interpretation.

METHOD

Design and Sample

The study involved 36 first semester students from one PGSD class at Universitas Halmahera during the 2025/2026 academic year. Initial classroom observation indicated that the students had difficulties with speaking. All 36 students participated in the intervention, attendance ranged from 34 to 36 students per meeting, and no student withdrew.

Instrument and Procedure

Speaking performance was assessed through an individual speaking task and a researcher developed analytic rubric. The rubric covered fluency, pronunciation, interaction skills, and vocabulary. Each dimension was scored from 1 (very limited) to 5 (very good), for a maximum raw score of 20. Scores were converted to a 100 point scale using the formula: Final Score = (Total Raw Score / 20) × 100. The descriptors addressed continuity of speech, intelligibility, interactional response, and lexical adequacy in the first semester PGSD context. Before scoring, two raters attended a 60 minute calibration session in which they reviewed the rubric and discussed sample performances. The raters were the researcher, who was also an English lecturer, and a second English lecturer experienced in speaking assessment. The second rater scored independently without access to the first rater's scores. Both raters knew whether a performance was from the pretest or posttest. The final score was the mean of their two scores. Performances were not audio or video recorded; scores were entered directly on the rubrics.

In the pretest, students introduced themselves and discussed their personal background, family, daily activities, and hobbies. They had approximately five minutes to prepare and could not use a written script. The researcher served as the interaction partner, asking main and brief follow up questions. The posttest followed the same individual format and used comparable familiar topics, including campus activities, daily routines, hobbies, family, and simple classroom or teaching situations. The exact questions differed to reduce memorization effects. Each test lasted approximately five to seven minutes.

The intervention comprised eight structured conversation practice meetings, held once a week for eight weeks. Each meeting lasted approximately 90 minutes, for a planned total of 720 minutes, or 12 hours. Instruction progressed from teacher modelling and guided practice to pair work, small group interaction, and more

independent speaking. Topics included self introduction, personal and family information, daily routines, hobbies, campus life, asking and answering questions, describing people and places, and everyday classroom or teaching situations. Materials included handouts, model dialogues, vocabulary lists, conversation prompts, role play cards, worksheets, and visual materials.

The researcher taught and facilitated the meetings by providing instructions and language models, organizing activities, monitoring participation, and giving feedback on fluency, pronunciation, interaction skills, and vocabulary. Students worked individually, in pairs, and in groups of approximately six. Each student was expected to engage in approximately 10 to 15 minutes of active speaking practice per meeting across the activities. Attendance was recorded at every meeting, while lesson plans and session notes were used to monitor implementation. No additional speaking intervention was intentionally introduced outside these sessions, although students continued their regular English coursework.

Data Analysis

Descriptive statistics for the overall pretest and posttest speaking scores included the sample size, minimum, maximum, mean, and standard deviation. The Shapiro–Wilk test was applied to the paired difference scores, calculated as posttest minus pretest. Because these differences were not normally distributed, a two sided Wilcoxon signed rank test was used, with statistical significance set at $\alpha = .05$. The rank based effect size was calculated as $r = |z|/\sqrt{N}$. The mean change and its 95% confidence interval were reported as descriptive estimates. Percentage change was used only as a descriptive index. The pretest and posttest means and standard deviations for the four rubric dimensions were presented descriptively. No separate inferential tests were conducted for the dimensions, so differences between them were not interpreted as statistically significant.

RESULT AND DISCUSSION

Table 1. Overall Speaking-Performance Scores

Measurement	N	Minimum	Maximum	Mean	SD
Pre-test	36	51.5	59.0	55.36	2.37
Post-test	36	71.0	79.0	75.19	2.61

The descriptive results show that the mean overall speaking-performance score rose from 55.36 at pre-test to 75.19 at post-test. The mean change was 19.83 points, equivalent to 35.83% relative to the pre-test mean. According to the study's interpretation categories, the group mean moved from the Fair category to the good category.

Table 2. Indicator-Level Descriptive Results

Speaking dimension	Pre-test M	Post-test M	Change	SD pre	SD post	Inferential test
Fluency	2.72	3.81	1.09	0.13	0.13	Descriptive only
Pronunciation	2.89	3.71	0.82	0.12	0.13	Descriptive only
Interaction skills	2.72	3.81	1.09	0.13	0.13	Descriptive only
Vocabulary	2.74	3.71	0.96	0.10	0.13	Descriptive only

All four dimensions showed higher descriptive means at post-test. Fluency increased by 1.09 points, pronunciation by 0.82, interaction skills by 1.09, and vocabulary by 0.96. These figures describe the observed pattern only. Because no separate inferential analyses were conducted for the four dimensions, the data do not justify concluding that one dimension improved statistically more than another.

Table 3. Normality and Paired Comparison

Analysis	Statistic	p	Interpretation
Shapiro–Wilk test of paired differences	W = 0.929	0.023	Difference scores were not normally distributed
Wilcoxon signed-rank test	W = 0	< .001	Post-test scores were significantly higher than pre-test scores

The Shapiro–Wilk test indicated non-normality of the paired difference scores, W = 0.929, p = 0.023. The Wilcoxon signed-rank test was therefore used as the primary inferential procedure. The test indicated a statistically significant difference between the two measurement occasions, W = 0, z = -5.23, p < .001. The mean observed change was 19.83 points, with a 95% confidence interval for the mean change of 19.44 to 20.23 points. The reported rank-based effect size was r = 0.87, which is large. All 36 participants had a higher post-test total than their pre-test total. These statistics establish a within-cohort difference across measurement occasions, not a causal estimate of the effect of conversation practice.

The principal finding is a clear within-cohort increase in overall speaking-performance scores. The mean rose from 55.36 to 75.19, a change of 19.83 points. The Wilcoxon result (W = 0, z = -5.23, p < .001) and the large rank-based effect size (r = 0.87) indicate that the post-test scores were consistently higher in the observed cohort. This pattern is compatible with the instructional period, but it should not be described as definitive evidence that conversation practice alone produced the change. The fluency pattern is consistent with literature on repeated oral production. Repetition can provide learners with additional opportunities to

retrieve language and perform communicative functions, potentially reducing hesitation and supporting aspects of fluency and lexical use (Dawadi, 2019; Duong et al., 2023; Hunter, 2024; Suzuki, 2021). The present sequence repeatedly moved students from modelling and guided practice toward less-supported interaction. Such repetition is a plausible mechanism for the observed change, although the design cannot distinguish the effect of repetition from feedback, task familiarity, or other components of the instructional sequence.

The descriptive increase in interaction skills is also compatible with research emphasizing peer interaction. Pair and small-group activities require students to listen, respond, ask questions, and maintain an exchange, thereby creating more opportunities for oral participation than activities limited to isolated production (Soh et al., 2021). Research on willingness to communicate further suggests that classroom conditions influence students' readiness to use English (Liu et al., 2021; Ramli et al., 2021; Weda et al., 2021). The present study did not measure willingness to communicate, however, so any explanation involving increased willingness remains interpretive rather than an empirical finding. Vocabulary also increased descriptively. Students repeatedly encountered and used words and expressions in familiar conversational settings. This pattern is compatible with evidence that vocabulary resources are related to second-language speaking ability (Uchihara & Clenton, 2020) and with research connecting task repetition to lexical use (Duong et al., 2023). Nevertheless, the study assessed vocabulary as one component of a speaking rubric rather than through an independent vocabulary test. The result should therefore be described as improved vocabulary use within speaking performance, not as evidence of a general increase in vocabulary size.

Pronunciation likewise showed a descriptive increase. Teacher modelling and corrective feedback may have given students repeated opportunities to notice and reproduce target forms. This interpretation is plausible, but pronunciation ratings form part of a researcher-developed instrument and were not supported by a separate pronunciation measure. The observed change should therefore be treated as a descriptive pattern rather than a statistically isolated pronunciation effect. The overall pattern broadly converges with previous conversation-oriented studies. Mariam et al. (2022) and Mariana et al. (2024) reported improved speaking outcomes following conversation-based activities, while Yamazaki (2023) reported communication-related improvements associated with pair work. Tran et al. (2024) also emphasized meaningful conversational interaction in an authentic online setting. The present study adds evidence from an intact first-semester PGSD cohort at Universitas Halmahera, but the different settings, tasks, and measures mean that the findings should be interpreted as contextual rather than as a direct replication of those studies.

Several alternative explanations limit causal interpretation. Exposure to the pre-test format may have produced a testing or practice effect. Students may also have developed through maturation, regular English coursework, or speaking practice outside the intervention. Regression toward the mean is another possible

explanation when initial scores are relatively low. In addition, the researcher participated in teaching and assessment, and both raters knew the testing occasion; expectancy effects therefore cannot be excluded. Most importantly, the absence of a comparison group prevents separation of these explanations from the contribution of the conversation-practice program.

Measurement limitations should also be considered. The primary outcome combines four related but distinct dimensions, and each dimension is represented by a single 1–5 rating. The absence of recordings also prevents later re-scoring of performances. These issues do not erase the observed pre/post difference, but they reduce confidence in the precision and generalizability of the measurement. From a pedagogical perspective, the observed change suggests that structured conversation activities may be useful for creating regular opportunities for students to use English. Short dialogues, pair conversations, role-play, question-and-answer tasks, and familiar-topic discussions can be incorporated into PGSD speaking instruction. Future research should use a comparison or quasi-experimental group, larger and more diverse samples, independently scored recordings, stronger evidence for instrument validity, and more systematic fidelity monitoring. Such designs would provide a more defensible basis for estimating the specific contribution of conversation practice.

CONCLUSION

This study identified a statistically significant within-cohort difference in overall speaking-performance scores among 36 first-semester PGSD students at Universitas Halmahera. The mean score increased from 55.36 at pre-test to 75.19 at post-test, producing a mean change of 19.83 points. The Wilcoxon signed-rank test yielded $W = 0$, $z = -5.23$, $p <. 001$, with a large rank-based effect size ($r = 0.87$). Fluency, pronunciation, interaction skills, and vocabulary also showed descriptive increases. The results provide preliminary evidence of higher speaking-performance scores following the structured conversation-practice period. However, the improvement cannot be attributed to conversation practice alone because the study used a one-group pre-test/post-test design without a comparison group. Potential effects of testing, maturation, concurrent learning, regression toward the mean, and rater expectancy cannot be ruled out. Future studies should strengthen the research design by including a comparison group, validated assessment procedures, independent raters, and more systematic monitoring of implementation.

REFERENCES

Albelihi, H. H. M. (2022). Intermediate English as a foreign language learners' formulaic language speaking proficiency: Where does the teaching of lexical chunks figure? *Frontiers in Psychology*, 13, 949675. https://doi.org/10.3389/fpsyg.2022.949675

Andini, V., Aridah, A., & Setiawan, I. (2022). Exploring the factors and strategies used to cope with speaking anxiety of EFL students during online learning. *Borneo Educational Journal*, *4(2)*, 121–130. https://doi.org/10.24903/bej.v4i2.1093

Dawadi, S. (2019). Effects of task repetition on EFL oral performance. *ITL – International Journal of Applied Linguistics*, *170(1)*, 3–23. https://doi.org/10.1075/itl.16025.daw

Duong, P.-T., Montero Perez, M., Nguyen, L.-Q., Desmet, P., & Peters, E. (2023). The impact of input, input repetition, and task repetition on L2 lexical use and fluency in speaking. *Studies in Second Language Learning and Teaching*, *13(1)*, 101–124. https://doi.org/10.14746/ssllt.29727

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*, 5500077. https://doi.org/10.1155/2022/5500077

Hasan, M. K., Fakih, A.-H., & Seraj, P. M. I. (2022). The effect of technology-assisted language programme on vocabulary learning among EFL students at the tertiary level. *Heliyon*, *8(8)*, e10313. https://doi.org/10.1016/j.heliyon.2022.e10313

Hunter, A.-M. (2024). The poster carousel in the ESL classroom: What happens to learners' L2 fluency during same and parallel-task repetition? TESOL Quarterly, 58(1), 479–510. https://doi.org/10.1002/tesq.3257

Jannah, N., Noni, N., & Talib, A. (2022). Students' speaking anxiety in EFL classroom of senior high school. *Journal of Excellence in English Language Education*, *1(3)*. https://doi.org/10.26858/joeele.v1i3.35438

Liu, F., Vadivel, B., Rezvani, E., & Namaziandost, E. (2021). Using games to promote English as a foreign language learners' willingness to communicate: Potential effects and teachers' attitude in focus. *Frontiers in Psychology*, *12*, 762447. https://doi.org/10.3389/fpsyg.2021.762447

Mariam, S., Kepirianto, C., Raharjo, R., & Ma'mun, N. (2022). Integrating short conversation and English conversation applications to improve students' speaking skills. *Journal of English Teaching and Learning Issues*, *5(2)*.

Mariana, A., Jabu, B., & Muliati, A. (2024). Improving the students' speaking ability by applying guided conversation strategy. *Journal of Excellence in English Language Education*, *3(4)*. https://doi.org/10.26858/joeele.v3i4.67086

Ramli, K., Hidayah, J., Edy, S., & Esmianti, F. (2021). Factors of students' willingness and unwillingness to speak English in the classroom. *Journal of English Education and Teaching*, *5(1)*, 95–109. https://doi.org/10.33369/jeet.5.1.95-109

Soh, S. B., Tam, S. S., & Nikitina, L. (2021). Peer interaction in the L2 classroom: A study among Malaysian ESL learners. *Issues in Language Studies*, *10(1)*, 131–150. https://doi.org/10.33736/ils.2592.2021

Sun, B., & Révész, A. (2021). The effects of task repetition on child EFL learners' oral performance. *Canadian Journal of Applied Linguistics*, *24(2)*, 30–47. https://doi.org/10.37213/cjal.2021.31382

Suzuki, Y. (2021). Optimizing fluency training for speaking skills transfer: Comparing the effects of blocked and interleaved task repetition. *Language Learning*, *71(1)*, 100–132. https://doi.org/10.1111/lang.12433

Tran, N., Hoang, D. T. N., Gillespie, R., Yen, T. T. H., & Phung, H. (2024). Enhancing EFL learners' speaking and listening skills through authentic online conversations with video conferencing tools. Innovation in Language Learning and Teaching. https://doi.org/10.1080/17501229.2024.2334809

Uchihara, T., & Clenton, J. (2020). Investigating the role of vocabulary size in second language speaking ability. *Language Teaching Research*, *24(4)*, 540–556. https://doi.org/10.1177/1362168818799371

Weda, S., Atmowardoyo, H., Rahman, F., Said, M. M., & Sakti, A. E. F. (2021). Factors affecting students' willingness to communicate in EFL classroom at higher institution in Indonesia. *International Journal of Instruction*, *14(2)*, 719–734. https://doi.org/10.29333/iji.2021.14240a

Yamazaki, K. (2023). Enhancing EFL learners' communication skills in English classes at a Japanese university. *TELES Journal*, *43*, 181–196. https://doi.org/10.57539/telesjournal.43.0_181

Yoskapela, Susanty, Bahing, & Ristati. (2022). A study on speaking anxiety of English Education Study Program students at the University of Palangka Raya. *EBONY: Journal of English Language Teaching, Linguistics, and Literature*, *2(1)*, 14–27. https://doi.org/10.37304/ebony.v2i1.4065