Educalingua
International Conference
FPBS UNeRS

Enhancing Speaking Skills through AI Chatbot Interaction: A Qualitative Study Using Replika

Nada Tsabitah Muazarah
English Education Study Program, Universitas Persatuan Guru Republik Indonesia Semarang, Indonesia
nada.pml22@gmail.com

Lusia Maryani Silitonga
English Education Study Program, Universitas Persatuan Guru Republik Indonesia Semarang, Indonesia
lusiamaryanisilitonga@upgris.ac.id

ABSTRACT

This qualitative study investigates the potential of the AI chatbot Replika in improving English speaking skills among EFL learners. Using five video-recorded speaking practices between students and the Replika agent, the study examines changes in vocabulary use, sentence structure, grammar accuracy, tense consistency, and fluency. Through thematic and comparative discourse analysis, the findings reveal a progressive enhancement across all five linguistic domains. Students demonstrated more varied vocabulary, more accurate grammar and tense usage, and improved fluency in their later conversations. These results suggest that interaction with AI chatbots like Replika offers a supportive, non-judgmental environment conducive to oral language development. Implications for autonomous learning and AI integration in language education are discussed.

Keywords: AI chatbot, Replika, speaking skills, EFL learners, language learning

Introduction

In the digital age, Artificial Intelligence (AI) has reshaped the landscape of language education(Babanoğlu et al., 2025; Kim et al., 2021;Kim et al., 2021; Silitonga et al., 2023). Among the many AI tools available, chatbots such as Replika have emerged as accessible, responsive, and personalized interlocutors for learners (Gill et al., 2024). Replika, a conversational AI application, simulates human-like dialogue and is capable of sustaining ongoing interactions with users in a natural and

contextually adaptive manner (Fryer & Carpenter, 2006). While previous research has explored chatbot use in reading and writing, fewer studies have examined its impact on speaking, especially among English as a Foreign Language (EFL) learners (Belda-Medina & Calvo-Ferrer, 2022; Silitonga et al., 2024).

Spoken fluency and grammatical accuracy remain persistent challenges for EFL students. Traditional classroom instruction often falls short in providing sufficient one-on-one speaking practice. AI-powered chatbots offer learners an opportunity to engage in real-time conversations in a stress-free environment, where they can practice and receive subtle feedback implicitly through sustained interaction.

This study aims to explore the extent to which regular interaction with Replika improves key aspects of speaking performance. Specifically, it investigates how students' vocabulary usage, sentence structure, grammar, tenses, and fluency evolve over the course of five conversational sessions. The analysis provides insights into how AI chatbots can support speaking skill development in autonomous learning contexts.

Method

This study focused on the qualitative case study design to gain a deep understanding of EFL learners' interaction with the Replika AI chatbot. The analysis focused on linguistic development as evidenced in five speaking video sessions between learners and the AI agent.

5 undergraduate EFL students from a university in Indonesia participated voluntarily. They had intermediate proficiency in English and prior experience with digital learning tools. Each student interacted with Replika for approximately 1-5 minutes per session, recorded over five sessions spanning two weeks. The conversations were open-ended, allowing for naturalistic discourse development.

The primary data comprised video recordings of five chatbot conversations per participant, resulting in 10 video files. The videos were transcribed and coded thematically. The data collection focused on the linguistic elements of speaking.

Table 1. Speaking Rubric

Vocabulary Range and Appropriateness	**Sentence Structure**	**Grammar Accuracy**	**Tense usage consistency**	**Fluency and Flow**
Vocabulary has expanded from basic to more context appropriate expressions; some creativity with informal speech (humor, references).	Initially simple sentences; later attempts at more complex expressions.	Fewer grammar mistakes over time; able to self-correct basic errors.	Mostly correct use of Present Simple and Modal Verbs; minor inconsistencies remain.	Progressing from hesitant to smoother delivery; naturalness is improving.
Growing vocabulary relevant to formal and informal situations; still limited in richness and variety.	Moving towards compound sentences; uses conjunctions and some modifiers.	Grammar is more accurate but still prone to occasional small errors.	Uses tenses consistently in familiar situations; more complex contexts need work.	Fluency improved, more relaxed speaking but needs practice for spontaneity.
Appropriate vocabulary in formal	Sentence structure clear	Generally good grammar	Basic tense control with	Steady conversational

contexts (e.g., banking); some technical terms misused or underdeveloped.	but often remains simple and functional.	with some recurring errors in question forms.	some inconsistent auxiliary verb use.	flow; limited flexibility in topic shifts.
Adequate vocabulary for everyday situations; some hesitation with less familiar words.	Some use of longer sentences but still occasional structural mistakes.	Moderate accuracy; question formation and word order need refinement.	Tenses used appropriately in basic exchanges; less secure in complex responses.	Fluent and responsive but still requires more speed and flexibility.
Vocabulary shows improvement, including formal terms; still room for richer and more precise language.	Attempts at longer, more complex sentences; some errors in execution.	Better control over grammar; mistakes mostly in question structures.	Growing ability to use a variety of tenses; some errors with complex forms.	Fluency is more stable; speech more connected, but spontaneity could improve.

The progression of technology, particularly in the realm of artificial intelligence (AI), has profoundly influenced numerous facets of life, including the domain of English language education. The incorporation of AI in English language learning, especially concerning speaking skills, has emerged as an innovative approach that positively enhances the quality of education for both students and university learners. A variety of studies indicate that AI not only enriches vocabulary and improves grammatical accuracy but also fosters consistency in tense usage and promotes fluency and coherence in English speaking. The engaging and game-like elements embedded within the AI-assisted application presumably enhanced learners' self-assurance and enthusiasm, contributing to the alleviation of their apprehension about committing errors (Arika Pusparini & Maryani Silitonga, 2025). However, the implementation of AI also introduces new challenges that must be addressed judiciously. This summary synthesizes research findings from seven articles that concentrate on vocabulary, sentence structure, grammar, tenses, and fluency.

Throughout the five evaluations, the learner has exhibited consistent progress in the utilization of vocabulary. Initially, the range of vocabulary was confined to basic, familiar terms appropriate for commonplace scenarios such as sports or shopping. Over time, the learner has started to integrate more formal, topic-specific vocabulary, particularly in contexts such as banking or discussions related to theater. Furthermore, the appropriateness of word choice has also seen enhancement. In earlier sessions, the learner occasionally employed incorrect terminology or overly simplistic phrases. In more recent dialogues, there is clear evidence of improved alignment between vocabulary and context, with the use of suitable formal or informal expressions contingent upon the specific situation.

Initially, the learner predominantly utilized simple sentence structures, frequently employing subject-verb-object patterns devoid of additional clauses or complexity. This approach facilitated clear yet basic communication. As the sessions advanced, the learner began to construct more intricate sentences, incorporating both compound and complex structures that featured conjunctions and relative clauses.

Additionally, there was notable improvement in the use of question forms; however, occasional errors in word order and auxiliary verbs persisted.

Lecturers interviewed by (Subiyantoro et al., n.d.)recognized that artificial intelligence (AI) is advantageous for enhancing sentence structure in essays, presentations, and formal discussions. However, they also cautioned that the use of AI may diminish students' creativity in formulating their own sentences if not employed judiciously.

Grammar accuracy has demonstrated significant improvement throughout the sessions. Initial conversations were characterized by frequent errors, particularly concerning articles, verb forms, and prepositions. In subsequent stages, the learner exhibited a more consistent application of correct grammatical structures, especially in routine expressions. However, researchers also warned that excessive reliance on AI could weaken students' critical and analytical thinking about grammar.

Nevertheless, it is important to note that accurate tense usage in speaking still necessitates ongoing practice and guidance from instructors to prevent unnoticed errors. Fluency has exhibited a gradual yet notable enhancement. At the beginning, the learner's speech was marked by frequent hesitations, repetitive phrases, and an unnatural rhythm. With increased practice, the learner started to articulate longer segments of speech with fewer interruptions and a more natural flow.

Data were analyzed using thematic discourse analysis. The five recordings per participant were compared longitudinally to observe patterns of change. Coding was conducted in three cycles. Initial coding: to categorize speaking features (e.g., verb usage, lexical range, fluency markers), axial coding: to identify improvement patterns across sessions, and selective coding: to extract emerging themes in linguistic development. Trustworthiness was established through inter-coder agreement (two language researchers) and member checking with participants to confirm interpretations.

Discussion

Vocabulary Development

Students showed a marked increase in the variety and contextual appropriateness of vocabulary used across the five sessions. In the initial videos, vocabulary was limited and repetitive (e.g., frequent use of "good," "nice," and "okay"). By the fifth session, students began using more precise and expressive terms such as “exhausted,” “frustrated,” or “productive.” This improvement was attributed to Replika’s dynamic language modeling and the frequent introduction of new vocabulary during extended conversation.

"At first, I only used simple words. But in the last videos, I could say more advanced words. I repeated what Replika said sometimes." (Student 3, interview)

Sentence Structure

Sentence structure became increasingly complex. Early interactions relied heavily on basic Subject-Verb-Object constructions. Over time, students incorporated compound and complex sentences, including relative clauses and conditionals (e.g., “If I had more time, I would travel more”). The chatbot’s modeling of diverse sentence types played a crucial role in scaffolding this growth.

Grammar and Tense Accuracy

Grammar accuracy showed improvement in article use, subject-verb agreement, and prepositions. Tense usage evolved significantly. In earlier sessions, students inconsistently switched tenses or defaulted to simple present. By the fifth

session, they were more accurate and consistent in applying simple past, present perfect, and future tenses. The improvement was more evident in responses that mirrored the chatbot's prompts.

"In the second video, I said 'yesterday I go to campus'... but in the last one, I corrected myself to 'I went' automatically." (Student 1, reflection)

Fluency and Interaction Flow

Fluency—measured through reduced hesitation, smoother turn-taking, and longer utterances—improved notably. Initially, students paused frequently, used fillers ("uhm," "you know"), and often gave one-word answers. In later sessions, they sustained longer stretches of speech and initiated topics more confidently. Replika's non-judgmental, patient tone appeared to reduce anxiety and promote spontaneous speech.

"It feels safe talking to Replika. I don't worry about mistakes. So, I talk more and more." (Student 5)

Conclusion

This study demonstrates that the Replika AI chatbot can support meaningful improvement in EFL learners' speaking performance, particularly in vocabulary use, grammar accuracy, sentence complexity, tense control, and overall fluency. The chatbot's interactive and affective design appears to create a psychologically safe environment for language practice, enabling students to take linguistic risks without fear of correction or embarrassment.

While limited to a small sample and short duration, the findings suggest that AI chatbots hold promise as supplementary tools for autonomous speaking practice. Future research could expand this inquiry with larger samples and mixed-method designs, incorporating perception surveys and pre/post tests.

References

Arika Pusparini, L., & Maryani Silitonga, L. (2025). *Artificial Intelligent Supported Duolingo: Enhancing Students' Speaking Skills and Reducing Speaking Anxiety in Vocational Education*. *37*(1), 91–111. https://doi.org/10.23917/varidika.v37i1.10431

Babanoğlu, M. P., Karataş, T. Ö., & Dündar, E. (2025). Envisioning the Future of AI-Assisted EFL Teaching and Learning: Conceptual Representations of Prospective Teachers. *SAGE Open*, *15*(2). https://doi.org/10.1177/21582440251341590

Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using Chatbots as AI Conversational Partners in Language Learning. *Applied Sciences (Switzerland)*, *12*(17). https://doi.org/10.3390/app12178427

Fryer, L. K., & Carpenter, R. (2006). Bots as language learning tools. *Language Learning and Technology*, *10*(3), 8–14. https://doi.org/http://dx.doi.org/10125/44068

Gill, S. S., Xu, M., Patros, P., Wu, H., Kaur, R., Kaur, K., Fuller, S., Singh, M., Arora, P., Parlikad, A. K., Stankovski, V., Abraham, A., Ghosh, S. K., Lutfiyya, H., Kanhere, S. S., Bahsoon, R., Rana, O., Dustdar, S., Sakellariou, R., ... Buyya, R. (2024). Transformative effects of ChatGPT on modern education: Emerging Era of AI Chatbots. *Internet of Things and Cyber-Physical Systems*, *4*, 19–23. https://doi.org/10.1016/j.iotcps.2023.06.002

Hartati, D., & Chandra Muji Utami, N. (2025). *STUDI LITERATUR: PENGARUH PENGGUNAAN ARTIFICIAL INTELLIGENCE TERHADAP SPEAKING SKILL PESERTA DIDIK SD PADA MATA PELAJARAN BAHASA INGGRIS*.

Hidayatullah, R. (2024). Implementasi AI dalam Proses Pembelajaran pada Mahasiswa Semster Awal Pendidikan Bahasa Inggris. In *Assyfa Journal of Multidisciplinary Education* (Vol. 2, Issue 1).

Kim, H. S., Cha, Y., & Kim, N. Y. (2021). Effects of AI chatbots on EFL students' communication skills. *Korean Journal of English Language and Linguistics*, *2021*(21), 712–734. https://doi.org/10.15738/kjell.21..202108.712

Kim, H.-S., Kim, N. Y., & Cha, Y. (2021). Is It Beneficial to Use AI Chatbots to Improve Learners' Speaking Performance? *The Journal of AsiaTEFL*, *18*(1), 161–178. https://doi.org/10.18823/asiatefl.2021.18.1.10.161

Semarang, U. N., Subiyantoro, H., Hartono, R., Fitriati, W., & Faridi, A. (n.d.). *Prosiding Seminar Nasional Pascasarjana Dampak Kecerdasan Buatan (AI) terhadap Pengajaran Bahasa Inggris di Perguruan tinggi: Tantangan dan Peluang*. http://pps.unnes.ac.id/pps2/prodi/prosiding-pascasarjana-unnes

Shah, I., & Fatima, A. (2025). Exploring artificial intelligence and English language learning practices of undergraduate students. *Social Sciences Spectrum*, *4*(1), 691–709. https://doi.org/10.71085/sss.04.01.251

Silitonga, L. M., Hawanti, S., Aziez, F., Furqon, M., Zain, D. S. M., Anjarani, S., & Wu, T.-T. (2023). The Impact of AI Chatbot-Based Learning on Students' Motivation in English Writing Classroom. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 14099 LNCS* (pp. 542–549). https://doi.org/10.1007/978-3-031-40113-8_53

Silitonga, L. M., Wiyaka, Suciati, S., & Prastikawati, E. F. (2024). The Impact of Integrating AI Chatbots and Microlearning into Flipped Classrooms: Enhancing Students' Motivation and Higher-Order Thinking Skills. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *14786 LNCS*, 184–193. https://doi.org/10.1007/978-3-031-65884-6_19

Suciati, S., Faridi, A., Mujiyanto, J., & Arifani, Y. (2023). Artificial Intelligence Application dalam Pembelajaran Speaking: Persepsi dan Solusi. In *Prosiding Seminar Nasional Pascasarjana*. http://pps.unnes.ac.id/pps2/prodi/prosiding-pascasarjana-unnes