

Artificial Intelligence-Driven Personalized Language Learning: Customizing Content and Feedback to Learners' Needs and Proficiency Levels

Mengkai Wang* and Yifei Gao
Beijing Foreign Studies University, China

Abstract

Artificial Intelligence (AI)-driven personalized language learning holds great promise for customizing content and feedback according to learners' needs and proficiency levels. This study explores the potential of AI, particularly large language models (LLMs), for enhancing personalized language learning experiences. The research examines how LLMs can be leveraged to provide dynamic, tailored content and real-time feedback, adapting to individual learners' progress and learning preferences. A group of 100 language learners participated in a 6-week experiment where they engaged with AI-powered platforms that adjusted the content's difficulty and offered personalized feedback based on their performance and proficiency levels for English grammar and vocabulary acquisition. The study used a randomized controlled trial design, assessing learners' language proficiency, motivation, engagement, and perceived effectiveness of the AI-assisted learning experience before and after the intervention. The results showed that LLMs significantly improved learners' language proficiency, with noticeable increases in their engagement and motivation. Learners reported higher satisfaction with the personalized feedback and content, appreciating the system's ability to adapt to their unique needs. However, challenges such as occasional misinterpretations by the AI and limited content variety were identified. This study contributes to

*Corresponding Author E-mail: tayloricy@bfsu.edu.cn & smxgaoyf@126.com

understanding the effectiveness of AI in creating personalized language learning environments and highlights the potential of integrating such technologies into mainstream language education.

Keywords: AI-driven, feedback, learner needs, proficiency levels, language learning, personalization

Introduction

Personalized learning has been a long-standing goal in education, particularly in language acquisition, where individual differences in proficiency, motivation, and learning strategies are profound (Dong et al., 2023). Traditional classroom-based instruction often struggles to accommodate these individual needs effectively. Recent advancements in artificial intelligence (AI), especially the development of large language models (LLMs) like GPT-4, have opened new possibilities for tailoring learning experiences dynamically to each learner (Prandi et al., 2023). LLMs are capable of generating human-like text, adapting content, and providing real-time feedback based on the users' input, making them particularly suited to personalized language learning environments.

Despite the excitement surrounding these technologies, empirical research into their effectiveness in educational settings, particularly in language learning, remains limited. Zheng et al. (2025) emphasize the need for robust experimental studies to explore how AI-driven platforms can be integrated into mainstream education to enhance learning outcomes. This study addresses this gap by investigating the impact of LLMs on language learners' proficiency, engagement, and motivation over a 6-week intervention.

Methodology

The study employed a randomized controlled trial design to assess the effectiveness of AI-driven personalized learning platforms. A total of 100 participants, aged 18 to 35, were identified as intermediate (CEFR B1 level) learners of English as a foreign language and recruited from a university language program. The CEFR (Common European Framework of Reference for Languages) is an international standard used to describe language proficiency levels, ranging from A1 (beginner) to C2 (mastery). The participants were randomly assigned to either an experimental group (AI-assisted learning) or a control group (traditional online learning without AI adaptation). Participants in the experimental group interacted with an AI-powered platform designed to adjust the content's difficulty and provide

personalized feedback based on real-time performance data. The platform, leveraging a model based on the GPT-4 architecture, was fine-tuned on a large corpus of learner texts with expert annotations to improve its accuracy in detecting grammatical errors and generating feedback. Learners completed adaptive grammar exercises (e.g., gap-filling, sentence transformations) and vocabulary-in-context tasks. In contrast, the control group used standard online language learning materials with static content. They accessed a fixed library of exercises covering the same grammatical and lexical topics as the experimental group. Feedback was limited to pre-programmed, generic responses, such as revealing the correct answer with a brief rule explanation. Pre- and post-intervention assessments were conducted using standardized language proficiency tests measuring grammar and vocabulary knowledge, self-reported motivation and engagement scales, and surveys on learners’ perceptions of the learning experience. Data were analyzed using mixed analysis of variance to assess changes over time between the two groups.

Results

An analysis of the proficiency scores, which measured grammar and vocabulary knowledge, is summarized in Table 1.

A significant group $\times$ time interaction effect was found ($F(1,98) = 23.45, p < 0.001$), indicating that participants in the AI-assisted group improved significantly more than those in the traditional learning group. Additionally, learners in the AI-assisted group reported higher levels of engagement and motivation. Figure 1 illustrates the changes in self-reported motivation scores.

Qualitative feedback revealed that learners appreciated the system’s ability to adapt to their performance, with comments highlighting the relevance of the exercises and the helpfulness of immediate, tailored feedback. However, several participants noted instances where the AI misinterpreted their responses, leading to occasional frustration.

Table 1. Descriptive statistics for pre- and post-test mean between groups

Group	Pre-Test Mean (SD)	Post-Test Mean (SD)	Gain Score (SD)
AI-assisted learning	72.3 (8.4)	82.7 (7.1)	10.4 (3.9)
Traditional learning	71.8 (7.9)	76.5 (7.5)	4.7 (3.5)

SD, standard deviation.

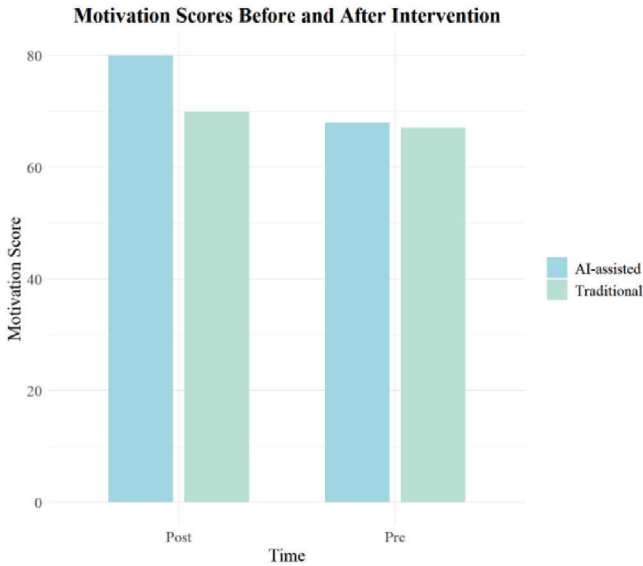


Figure 1. Motivation scores pre- and post-intervention for AI-assisted and traditional groups.

Discussion

The results of this study support the growing belief that AI-driven personalized learning can significantly enhance language acquisition outcomes. The substantial gains in proficiency scores and the heightened engagement and motivation reported by the participants align with prior theoretical expectations about the benefits of personalization (Christodoulou & Angeli, 2022). One major strength of the AI-powered system was its responsiveness to individual progress. Learners felt that the exercises were “just right” in difficulty, a hallmark of effective scaffolding in educational theory (Vygotsky, 1978). The enhanced proficiency can be attributed to this adaptive difficulty combined with personalized feedback; immediate corrections and tailored suggestions helped the learners consolidate their knowledge and address errors more efficiently than static learning materials.

Nonetheless, the study also uncovered challenges. Misinterpretations by the AI, although relatively infrequent, disrupted the learning flow for some participants (e.g., Ji et al., 2023). This highlights a limitation of current LLM-based systems: while powerful, they are not infallible and require careful tuning and perhaps hybrid human-in-the-loop models to mitigate errors. Moreover, the content variety offered by the LLM system was somewhat limited, as noted

by several participants. Learners reported that although the specific sentences and words were novel, the underlying exercise structures (e.g., multiple choice, gap-filling) were sometimes repetitive, which could impact long-term engagement (O'Neill & Satar, 2024). Future systems might incorporate larger and more diverse corpora to enhance content richness.

Conclusion

This study demonstrates the promising potential of LLMs in creating personalized language learning environments that effectively enhance learners' proficiency, engagement, and motivation. Proficiency was enhanced by providing learners with tailored feedback on specific errors and adaptively challenging tasks that aligned with their individual learning pace. While there are challenges to address, particularly regarding misinterpretations and the content's diversity, the findings suggest that the integration of AI in language education could be a transformative development. Future research should explore longer intervention periods, different age groups, and hybrid models combining AI and human feedback to further refine and optimize personalized language learning experiences.

References

- Christodoulou, A. & Angeli, C. (2022). Adaptive learning techniques for a personalized educational software in developing teachers' technological pedagogical content knowledge. *Frontiers in Education*, 7, 789397. <https://doi.org/10.3389/educ.2022.789397>
- Dong, B., Bai, J., Xu, T., & Zhou, Y. (2023). Large language models in education: A systematic review. In *Proceedings of the 6th International Conference on Computer Science and Technologies in Education* (pp. 131–134). <https://doi.org/10.1109/CSTE62025.2024.00031>
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), 1–38. <https://doi.org/10.1145/3571730>
- O'Neill, E., & Satar, M. (2024). Exploring the impact of task repetition on learner engagement in chatbot-mediated language learning. *Computer Assisted Language Learning*, 37(3), 214–235. <https://doi.org/10.1080/09588221.2023.xxxxxxx>
- Prandi, C., Valva, A., Salomoni, P., Roccetti, M. & Mirri, S. (2023). Incidental language and culture learning through mobile technologies: A multi-case

- study. *Multimedia Tools and Application*, 82, 46039–46063. <https://doi.org/10.1007/s11042-023-16095-1>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Zheng, Y. B., Zhou, Y. X., Chen, X. D., & Ye, X. D. (2025). The influence of large language models as collaborative dialogue partners on EFL English oral proficiency and foreign language anxiety. *Computer Assisted Language Learning*, 1–27. <https://doi.org/10.1080/09588221.2025.2453191>