The Impact of AI-Driven Speaking Practice in VR on Foreign Language Anxiety Among Japanese EFL Learners

Yukie Saito, Chuo University, Japan

The IAFOR International Conference on Education in Hawaii 2026
Official Conference Proceedings

Abstract

This study examined whether speaking practice with AI-driven avatars in a virtual reality (VR) environment was associated with changes in foreign language anxiety (FLA) among Japanese university students preparing for a short-term study abroad program. Using a quantitative pre-experimental pre-test/post-test design, 10 of 12 participants provided complete matched questionnaire data. During one month prior to departure, students completed structured, self-directed speaking practice in Immerse VR by interacting with AI avatars. FLA was assessed before and after the intervention using a 24-item measure adapted from the Foreign Language Classroom Anxiety Scale (Horwitz et al., 1986) rated on a 5-point Likert scale (higher scores indicating greater anxiety). Mean FLA decreased from pre-test ($M = 3.304$, $SD = 0.594$) to post-test ($M = 2.471$, $SD = 0.880$), corresponding to a mean difference of 0.833 points. The pre–post change was statistically significant ($p = .001$) and the estimated effect size was large (Cohen's $d = 1.518$, direction indicating lower post-test anxiety). Internal consistency was high at both time points ($\alpha = .922$ pre; $\alpha = .967$ post). Exploratory analyses did not show a clear association between time spent in Immerse VR and anxiety change. Because there was no control group and the sample was small, it is difficult to determine whether the observed effects were truly caused by the intervention, and the findings cannot generalize widely. Future studies with larger, controlled designs could clarify whether AI-mediated VR speaking practice helps reduce anxiety before a study abroad program.

Keywords: virtual reality, foreign language anxiety, Japanese EFL students, AI avatars, study abroad preparation

iafor
The International Academic Forum
www.iafor.org

Introduction

The number of Japanese university students studying abroad has been increasing over the past few years, following a significant decline during the COVID-19 pandemic. To maximize the opportunity to study abroad, students are expected to be well prepared to communicate in English. However, foreign language anxiety (FLA) may prevent students from communicating in English even when they have sufficient linguistic knowledge. This is particularly salient for Japanese university students who learn English as a foreign language, where exposure to English is limited. The recent integration of virtual reality (VR) and generative AI may help address this challenge, particularly through VR-based AI avatars, because students can practice speaking with AI avatars at any time outside class.

In this study, Japanese university students practiced speaking with AI avatars on a VR language platform, Immerse VR, for one month prior to a four-week study abroad program. To evaluate whether this speaking practice can reduce their foreign language anxiety, pre- and post-intervention FLA questionnaires were collected. Using a quantitative matched pre-test/post-test design, this paper reports preliminary findings focusing specifically on changes in FLA. An exploratory analysis also considers whether total time spent on the platform (Time on Site; TOS) is related to the magnitude of FLA change. The study addresses the following research question: Does practicing speaking with AI avatars in a VR environment reduce foreign language anxiety among Japanese university students preparing to study abroad?

Literature Review

VR: Definition, Educational Potential and Application to Language Education

Virtual reality (VR) refers to a computer-generated, immersive three-dimensional environment designed to simulate realistic experiences, providing visual feedback that dynamically responds to user movements (Dionisio et al., 2013). By means of interactive simulations, VR allows participants to engage in settings that closely replicate real-world scenarios, thereby facilitating a high level of immersion. Given its advanced immersive features, VR is anticipated to have extensive applications across various fields, including retail, entertainment, education, industrial training, healthcare, fitness, and social engagement (Hamad & Jia, 2022). Notably, numerous studies have highlighted the potential of VR within educational environments (Hu-Au & Lee, 2017).

In language education, VR is often used to create situationally rich interaction that resembles everyday communication and may help maintain learners' attention and support retention (Scrivner et al., 2019). Research on virtual learning environments suggests that interactive and authentic contexts can promote autonomy and confidence, and may also alleviate anxiety by providing practice opportunities that are not easily available in daily life (Lin & Lan, 2015). Recent syntheses and empirical studies increasingly indicate that VR-supported activities can improve oral performance and related outcomes, including oral presentation performance and speaking development, together with affective benefits (Chen et al., 2022; Qiu et al., 2024). Task-based VR research further points to potential gains in willingness to communicate, fluency, pronunciation, and confidence (Ebadi & Ebadijalal, 2020; Liu & Hou, 2020). Simultaneously, review studies indicate that while motivational gains and reduced anxiety are commonly observed, the evidence supporting substantial and consistently replicable enhancements in speaking fluency remains inconclusive (Yudintseva, 2023).

A critical factor in interpreting VR effects is the degree of immersion. High-immersion VR (HiVR), typically delivered via head-mounted displays, can produce stronger presence than low-immersion VR (LiVR) delivered via standard 2D displays (Kaplan-Rakowski & Gruber, 2019). HiVR has been reported to be more engaging and motivating than LiVR (Makransky & Lilleholt, 2018), but it may also introduce constraints such as motion sickness, eye strain, and higher cost (Dolgunsöz et al., 2018; Kaplan-Rakowski & Wojdynski, 2018). Accordingly, evaluating VR in language education may require attention not only to the presence of VR, but also to the level of immersion and the practical feasibility of sustained use over time.

AI and Generative AI for Speaking Practice

In parallel with VR, AI-supported language learning tools have expanded opportunities for interactive practice. The studies of AI-mediated instruction have reported improvements in speaking-related outcomes, motivation, and self-regulated learning compared with traditional instruction (Qiao & Zhao, 2023; Wei, 2023). Generative AI tools (e.g., ChatGPT) may enhance this potential by providing EFL learners with limited opportunities for English communication a means to practice speaking beyond the classroom. Recent studies report perceived benefits including expanded speaking opportunities and improvements in confidence or speaking self-efficacy (Muniandy & Selvanathan, 2024; Üstünbaş, 2024; Yıldız, 2024). Together, these lines of work suggest that AI may function as a conversational partner, potentially reducing fears often associated with speaking practice.

AI Avatars in VR: Early Evidence and Remaining Gaps for FLA

Several VR language platforms (e.g., Mondly VR, Ovation VR, Immerse VR) now incorporate AI avatars that allow learners to practice speaking in simulated real-world environments. Initial research suggests that avatar-based VR practice is often perceived as motivating and enjoyable, and may support confidence-building, though it can also increase fatigue or be affected by physical classroom constraints such as noise (Kawasumi & Ishii, 2023; Thrasher et al., 2024). Exploratory findings further suggest that VR practice combined with feedback may help reduce speaking anxiety in some contexts (Godefridi et al., 2021). Nevertheless, empirical evidence remains limited regarding whether speaking with AI avatars in immersive VR produces measurable reductions in foreign language anxiety (FLA), and how such changes relate to engagement metrics (e.g., time on platform).

Although the broader project of the present study collected multiple forms of data (e.g., speaking proficiency and learner perceptions), this paper reports preliminary results focusing specifically on FLA following a one-month AI-driven VR speaking intervention. This focused scope is intended to provide an initial estimate of the intervention's potential affective impact and to establish a foundation for subsequent analyses that integrate additional variables.

To support English learners before their study abroad program, speaking practice environments that reduce pressure are needed. Technology-mediated speaking practice may offer a low-stakes way for learners to rehearse communication in English. Building on this premise, the present study examines whether an AI-driven VR speaking intervention can reduce FLA among Japanese university students preparing to study abroad. Accordingly, this study investigates the efficacy of an AI-driven VR speaking intervention for Japanese university students preparing to study abroad. The primary research question is: Does practicing speaking with AI avatars in a VR environment reduce foreign language anxiety among Japanese university students preparing to study abroad?

Methodology

Research Design and Scope

This study employed a quantitative, pre-experimental design using a matched pre-test/post-test approach to examine changes in foreign language anxiety (FLA) following an AI-driven VR speaking intervention. Although the broader project collected additional data (e.g., speaking proficiency and learner perceptions), the present paper reports preliminary results focusing exclusively on FLA.

Participants

Twelve Japanese university students participated in a one-month ICT-focused study abroad program in Australia, where they took General English and IT-related courses. All participants were enrolled at a Japanese university and were preparing for an upcoming English-speaking environment abroad. Before the program, they practiced speaking with AI avatars in a VR environment. Matched pre- and post-intervention questionnaire data were obtained from 10 students.

Intervention and Procedure

The intervention consisted of one month of structured speaking practice using Immerse VR, a VR language-learning platform. Participants interacted with AI-powered avatars multiple times and were instructed to practice daily at home on their PCs (LiVR) during the one-month period in preparation for studying abroad. A pre-intervention questionnaire was administered before the practice period, and a post-intervention questionnaire was administered after the one-month intervention. Usage data from Immerse VR—recorded as Time on Site (TOS)—were obtained to examine the relationship between changes in foreign language anxiety (FLA) and TOS.

Foreign Language Anxiety (FLA) Questionnaire

FLA was assessed using a questionnaire adapted from the Foreign Language Classroom Anxiety Scale (FLCAS; Horwitz et al., 1986). The adapted instrument consisted of 24 items rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). For analysis, item responses were averaged to compute a mean FLA score for each participant at each time point, with higher scores indicating greater anxiety. Internal consistency reliability was examined using Cronbach's alpha at both administrations.

Data Analysis

Descriptive statistics (means and standard deviations) were calculated for pre- and post-intervention FLA scores. A matched pre-test/post-test comparison was conducted to evaluate whether the change in FLA was statistically significant. Effect size was calculated using Cohen's d to quantify the magnitude of change. Scale reliability was assessed using Cronbach's alpha. Exploratory correlational analyses examined (a) the association between change in FLA and time on site (TOS) and (b) the relationship between baseline FLA and TOS. Statistical significance was evaluated at an alpha level of .05.

Results

Matched pre- and post-intervention questionnaire data were available for 10 participants. Foreign language anxiety (FLA) was operationalized as the mean score across Items 1–24 on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree), with higher scores indicating greater anxiety. Internal consistency of the 24-item scale was excellent at both time points (Cronbach's α = .922 at pre-test; α = .967 at post-test; N = 10), supporting the reliability of the adapted measure in this sample.

Descriptive Change in FLA From Pre-test to Post-test

Descriptive statistics indicated a decrease in FLA following the intervention (Table 1). Mean FLA decreased from M = 3.304 (SD = 0.594) at pre-test to M = 2.471 (SD = 0.880) at post-test, corresponding to a mean change (post − pre) of −0.833 points (approximately 25% of the pre-test mean). The distribution shifted downward as well: the observed range moved from 2.542–4.250 at pre-test to 1.125–3.875 at post-test. The reduction was also evident at the upper end of the distribution, with the post-test maximum (3.875) lower than the pre-test maximum (4.250), suggesting fewer extremely high anxiety scores after the intervention.

Table 1
Pre- and Post-test FLA Scores

Scale	N	Pre M	Pre SD	Post M	Post SD	Range (Pre)	Range (Post)
FLA (Q1–Q24)	10	3.304	0.594	2.471	0.880	2.542–4.250	1.125–3.875

Despite the overall decrease, post-test scores were more dispersed than pre-test scores (SD: 0.594 → 0.880), indicating greater heterogeneity at post-test. This pattern suggests that participants did not converge on a similar post-intervention anxiety level; rather, some showed large decreases while others showed modest decreases, resulting in a wider spread of post-test outcomes even as the group mean declined.

Inferential Test of Pre–Post Difference

A paired-samples t test showed that FLA decreased significantly from pre-test to post-test, $t(9) = -4.79$, $p = .0010$. The mean change (post−pre) was −0.833, with a 95% confidence interval of [−1.225, −0.439]. The within-participant effect size was large ($d_z = -1.52$), indicating a substantial reduction relative to within-participant variability.

Given the small sample size and the potential sensitivity to parametric assumptions, a Wilcoxon signed-rank test was conducted as a nonparametric robustness check. The Wilcoxon test corroborated the reduction, $W = 0$, $p = .00098$, providing convergent evidence that the pre–post decrease in FLA is not dependent on the parametric model.

Individual-Level Pattern

To assess whether the group-level effect reflected a general pattern rather than being driven by a small subset of participants, individual pre–post changes were examined. All 10 participants showed lower FLA at post-test than at pre-test (i.e., $\Delta FLA < 0$ for every participant), indicating a directionally consistent decrease across the sample. However, the magnitude of change varied

substantially, ranging from −0.13 to −2.00 points, consistent with the increased post-test variability noted above.

The largest reduction was observed for Student I (ΔFLA = −2.00), whereas the smallest reduction was observed for Student E (ΔFLA = −0.13). The median change was −0.77, with an interquartile range (IQR) of approximately 0.65, suggesting moderate-to-large typical decreases but notable individual differences. Using practical thresholds, 7 of 10 participants showed decreases of ≥ 0.50 points and 4 of 10 showed decreases of ≥ 1.00 point, indicating that the mean reduction was not driven by a single extreme case but reflected decreases across multiple participants.

Platform Use (Time on Site; TOS): Descriptive Statistics and Exploratory Associations

Time on Site (TOS) captured how long each participant engaged with the VR platform outside of class. In the matched sample (N = 10), engagement varied widely. Participants spent an average of 8.86 hours on the platform (SD = 7.78), with a median of 7.32 hours, indicating that the mean was influenced by a small number of high-use observations. TOS ranged from 0.78 to 28.83 hours, suggesting a right-skewed distribution and pronounced between-participant differences in exposure.

Exploratory correlational analyses examined whether TOS was related to (a) the magnitude of anxiety reduction (ΔFLA) and (b) baseline anxiety (pre-test FLA). Because TOS showed substantial dispersion and included an extreme high-exposure participant, both Pearson's r (linear association; outlier-sensitive) and Spearman's ρ (rank-based; more robust to non-normality and outliers) were reported. TOS was not associated with anxiety reduction (ΔFLA) in this sample (Spearman's ρ = 0.042, p = .907; Pearson's r = −0.132, p = .715).

For baseline anxiety (pre-test FLA), Pearson's correlation indicated a positive association with TOS (r = 0.663, p = .037), whereas Spearman's correlation was not statistically significant (ρ = 0.604, p = .065). This discrepancy may reflect sensitivity to distributional features such as right-skew and/or outlier influence. Substantively, the direction of the association is consistent with the possibility that participants with higher initial anxiety tended to spend more time on the platform, but this interpretation remains tentative.

Discussion

This preliminary study examined whether a one-month AI-driven VR speaking intervention was associated with reduced foreign language anxiety (FLA) among Japanese university students preparing for study abroad. Overall, FLA decreased significantly from pre-test to post-test with a large within-participant effect, and all participants showed reductions. This pattern is broadly consistent with prior work suggesting that technology-mediated and VR-supported environments can provide low-stakes opportunities for practice that foster confidence and reduce anxiety by minimizing immediate social evaluation (e.g., Chen et al., 2022; Lin & Lan, 2015; Qiu et al., 2024). In the Japanese EFL context described in the Introduction—where spontaneous speaking opportunities are limited and fear of negative evaluation can constrain participation—an AI-avatar environment may function as a psychologically safer "rehearsal space," helping learners acclimate to real-time speaking demands before going abroad.

At the same time, participants' post-test FLA scores were more dispersed than pre-test scores, indicating heterogeneous responses to the intervention. Nonetheless, the overall pattern was a

clear reduction in FLA from pre- to post-test, suggesting that even low-immersion VR (LiVR) can be associated with lower foreign language anxiety. This combination of a reliable mean decrease alongside increased variability aligns with reviews noting that affective outcomes in VR are not uniformly large or consistent and may depend on learner characteristics and implementation conditions (e.g., Yudintseva, 2023). The immersion literature offers one possible perspective: high-immersion VR (HiVR) can heighten presence and engagement (Kaplan-Rakowski & Gruber, 2019; Makransky & Lilleholt, 2018), but it can also introduce constraints such as fatigue, discomfort, and classroom practicality issues (Dolgunsöz et al., 2018; Kaplan-Rakowski & Wojdynski, 2018). In this context, the present LiVR approach appears sufficient to reduce FLA on average, yet the wider spread of post-test scores suggests that the magnitude of anxiety reduction may vary across learners. Importantly, these findings raise the possibility that lowering FLA may not require HiVR; for some learners, more accessible LiVR practice may provide comparable affective benefits, although direct HiVR–LiVR comparisons are needed to determine whether immersion level moderates anxiety-related outcomes.

Exploratory analyses relating platform engagement (Time on Site; TOS) to outcomes further suggest that the time that learners spent on the platform to practice speaking with AI avatars is not related to anxiety reduction. Despite wide variation and a right-skewed TOS distribution, TOS showed no evidence of association with the magnitude of anxiety reduction (ΔFLA). This finding is compatible with the idea, highlighted in VR and AI-assisted learning research, that the quality and structure of practice may matter as much as (or more than) raw exposure time. It also resonates with the emerging generative-AI literature suggesting that conversational AI can support self-efficacy and self-regulated learning (e.g., Muniandy & Selvanathan, 2024; Qiao & Zhao, 2023; Wei, 2023), which are not fully captured by total time spent on a platform. Accordingly, richer engagement indicators such as number of sessions, task completion, and interaction depth may be needed to identify which aspects of AI-avatar practice are most closely tied to reduced anxiety.

For baseline anxiety, Pearson's correlation suggested that higher pre-test FLA was associated with greater TOS, whereas Spearman's estimate did not reach conventional significance. Interpreted cautiously, this tentative pattern is consistent with the possibility that more anxious learners sought out more practice because they perceived a greater need to prepare. However, given the small sample and outlier sensitivity, this should be treated as hypothesis-generating. Future work should examine whether baseline anxiety moderates engagement and outcomes (i.e., whether learners with higher initial anxiety both practice more and benefit differently).

These findings should be interpreted with caution in light of several limitations. First, the study employed a pre-experimental design without a control group; therefore, the observed reductions cannot be attributed uniquely to the intervention. Second, the sample was small and context-specific, limiting generalizability and the precision of correlational estimates. Engagement analyses may also have been influenced by the right-skewed TOS distribution. Third, although internal consistency was high, the adapted FLA measure may reflect multiple facets of anxiety. Future studies should triangulate self-report results with speaking-specific anxiety measures, behavioral indicators (e.g., willingness to communicate), and qualitative reflections on perceived evaluative pressure when interacting with AI avatars. Addressing these issues is particularly important because, despite the growing availability of AI-avatar VR platforms, empirical evidence on measurable FLA change—and its relationship to engagement—remains limited.

Conclusion

This study investigated whether a one-month AI-driven VR speaking intervention was associated with reduced foreign language anxiety (FLA) among Japanese university students preparing for study abroad. Using a matched pre-/post-test design (N = 10 with complete questionnaires), FLA decreased significantly from pre-test to post-test with a large within-participant effect, and all participants showed reductions. However, the magnitude of improvement varied across individuals, and post-test scores were more dispersed than pre-test scores. Platform engagement (Time on Site; TOS) also varied widely and showed a right-skewed distribution (N = 10), but exploratory analyses found no clear association between TOS and the magnitude of anxiety reduction (ΔFLA). A tentative pattern suggested that higher baseline FLA may be associated with greater platform use (Pearson significant; Spearman not), although this finding should be interpreted cautiously given the small sample and outlier sensitivity. Overall, the findings provide preliminary evidence that AI-avatar VR speaking practice may help reduce anxiety prior to study abroad, particularly in contexts such as Japanese EFL where opportunities for spontaneous English speaking are limited.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

The author declares that ChatGPT, a generative AI software, was used in proofreading and refining the language used in the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The author further declares that, apart from ChatGPT, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

References

Chen, B., Wang, Y., & Wang, L. (2022). The effects of virtual reality-assisted language learning: A meta-analysis. *Sustainability, 14*(6), Article 3147. https://doi.org/10.3390/su14063147

Dionisio, J. D. N., Burns, W. G., & Gilbert, R. (2013). 3D virtual worlds and the metaverse: Current status and future possibilities. *ACM Computing Surveys, 45*(3), 1–38. https://doi.org/10.1145/2480741.2480751

Dolgunsöz, E., Yıldırım, G., & Yıldırım, S. (2018). The effect of virtual reality on EFL writing performance. *Journal of Language and Linguistic Studies, 14*(1), 278–292. http://www.jlls.org/index.php/jlls/article/view/869

Ebadi, S., & Ebadijalal, M. (2020). The effect of Google Expeditions virtual reality on EFL learners' willingness to communicate and oral proficiency. *Computer Assisted Language Learning, 35*(8), 1975–2000. https://doi.org/10.1080/09588221.2020.1854311

Godefridi, I., Suñer, F., Leblanc, C., & Meunier, F. (2021). Using virtual reality and peer feedback to reduce L2 speaking anxiety: An exploratory study. In N. Zoghlami, C. Brudermann, C. Sarré, M. Grosbois, L. Bradley, & S. Thouësny (Eds.), *CALL and professionalisation: Short papers from EUROCALL 2021* (pp. 100–105). Research-publishing.net. https://doi.org/10.14705/rpnet.2021.54.1316

Hamad, A., & Jia, B. (2022). How virtual reality technology has changed our lives: An overview of the current and potential applications and limitations. *International Journal of Environmental Research and Public Health, 19*(18), 11278. https://doi.org/10.3390/ijerph191811278

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Hu-Au, E., & Lee, J. J. (2017). Virtual reality in education: A tool for learning in the experience age. *International Journal of Innovation in Education, 4*(4), 215–226. https://doi.org/10.1504/IJIIE.2017.10012691

Kaplan-Rakowski, R., & Gruber, A. (2019). Low-immersion versus high-immersion virtual reality: Definitions, classification, and examples with a foreign language focus. In *Proceedings of the Innovation in Language Learning International Conference 2019* (pp. 552–555). Pixel. https://conference.pixel-online.net/files/ict4ll/ed0012/FP/6143-SLA4232-FP-ICT4LL12.pdf

Kaplan-Rakowski, R., & Wojdynski, T. (2018). Students' attitudes toward high-immersion virtual reality assisted language learning. In P. Taalas, J. Jalkanen, L. Bradley, & S. Thouësny (Eds.), *Future-proof CALL: Language learning as exploration and encounters—Short papers from EUROCALL 2018* (pp. 124–129). Research-publishing.net. https://doi.org/10.14705/rpnet.2018.26.824

Kawasumi, S., & Ishii, Y. (2023). Comparative study of English learning using virtual reality and a smartphone application. *Bulletin of the Faculty of Education, Chiba University, 71*, 99–105. https://doi.org/10.20776/S13482084-71-P99

Lin, T.-J., & Lan, Y.-J. (2015). Language learning in virtual reality environments: Past, present, and future. *Educational Technology & Society, 18*(4), 486–497. https://www.jstor.org/stable/jeductechsoci.18.4.486

Liu, X., & Hou, S. (2020). Effects of VR on users' speaking performance and cognitive load: Design of a VR-assisted English-speaking application on mobile phones. In *Proceedings of the 2020 International Conference on Intelligent Design (ICID)* (pp. 128–131). IEEE. https://doi.org/10.1109/ICID52250.2020.00035

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development, 66*(5), 1141–1164. https://doi.org/10.1007/s11423-018-9581-2

Muniandy, J., & Selvanathan, M. (2024). ChatGPT, a partnering tool to improve ESL learners' speaking skills: Case study in a public university, Malaysia. *Teaching Public Administration, 43*(1), 4–20. https://doi.org/10.1177/01447394241230152

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*, Article 1255594. https://doi.org/10.3389/fpsyg.2023.1255594

Qiu, X., Shan, C., Yao, J., & Fu, Q. (2024). The effects of virtual reality on EFL learning: A meta-analysis. *Education and Information Technologies, 29*(2), 1379–1405. https://doi.org/10.1007/s10639-023-11738-0

Scrivner, O., Madewell, J., Buckley, C., & Perez, N. (2019). Best practices in the use of augmented and virtual reality technologies for SLA: Design, implementation, and feedback. In M. L. Carrió-Pastor (Ed.), *Teaching language and teaching literature in virtual environments* (pp. 55–72). Springer. https://doi.org/10.1007/978-981-13-1358-5_4

Thrasher, T., Kaplan-Rakowski, R., Ovsiannikova, U., Meyr, J., & Yuan, Y. (2024). "I can talk to Spanish speakers in Illinois!": Student perspectives on AI-avatar role plays in virtual reality. In *WorldCALL2023: Conference Proceedings* (pp. 157–165). https://doi.org/10.22492/issn.2759-1182.2023.18

Üstünbaş, Ü. (2024). Hey, GPT, can we have a chat?: A case study on EFL learners' AI speaking practice. *International Journal of Modern Education Studies, 8*(1), 91–107. https://doi.org/10.51383/ijonmes.2024.318

Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology, 14*, Article 1261955. https://doi.org/10.3389/fpsyg.2023.1261955

Yıldız, C. (2024). ChatGPT integration in EFL education: A path to improved speaking self-efficacy. *Novitas-ROYAL (Research on Youth and Language), 18*(2), 167–182. https://doi.org/10.5281/zenodo.13861137

Yudintseva, A. (2023). Virtual reality affordances for oral communication in English as a second language classroom: A literature review. *Computers & Education: X Reality, 2*, 100018. https://doi.org/10.1016/j.cexr.2023.100018