11th International Conference of the Immersive Learning Research Network
Practitioner Proceedings iLRN2025

DOI: https://doi.org/10.56198/8r82g517

Exploring Embodied Conversational Agents in Extended Reality for Language Learning

Daeun Han[1] and Jeeheon Ryu[1]

[1] Chonnam National University, Gwangju, Republic of Korea
han3761297@naver.com

Abstract. Embodied Conversational Agents (ECAs) integrated into Extended Reality (XR) environments have demonstrated transformative potential for second language (L2) acquisition by creating interactive, context-rich learning experiences. This study systematically reviews research from 2015 to 2024, following PRISMA guidelines, to evaluate the impact of ECAs on immersion, engagement, and cognitive learning outcomes. The findings reveal that ECAs effectively reduce learner anxiety through empathetic interactions, enhance cognitive engagement through adaptive feedback, and support task-based learning in realistic, immersive contexts. However, significant challenges remain, including limited cultural adaptability, declining learner motivation over time, and insufficient system flexibility to accommodate diverse learner needs. To address these challenges, this paper provides specific strategies for adapting ECAs to diverse cultural backgrounds and learning contexts. Additionally, the paper discusses personalized engagement strategies to sustain learner motivation over extended periods. By bridging these gaps, ECAs can maximize their potential to transform L2 education, delivering impactful, adaptive, and accessible language learning experiences.

Keywords: Embodied Conversational Agents, Immersive Learning Environments, Virtual Reality, Second Language Learning, Systematic Literature Review.

1 Introduction

The rapid advancement of immersive technologies, such as virtual reality (VR) and augmented reality (AR) has transformed educational practices by offering learners interactive experiences that bridge classroom instruction and real-world applications. Among these technologies, Embodied Conversational Agents (ECAs) have emerged as effective tools for facilitating dynamic, dialogue-driven L2 acquisition. ECAs leverage multimodal communication—speech, text, gestures, and facial expressions—to create natural, human-like interactions that promote active language practice [1].

While studies have separately explored the benefits of VR/AR technologies and ECAs, the combined potential of these technologies remains underexplored. By integrating ECAs into immersive environments, learners can experience contextualized, experiential language learning that reduces anxiety and fosters cognitive engagement [2]. Moreover, ECAs can serve as adaptive digital mentors, guiding learners in real time and offering immediate, context-aware feedback that enhances interaction quality. This adaptability is particularly beneficial for addressing diverse learning preferences and needs. Situated Learning Theory further emphasizes the importance of realistic contexts, where ECAs act as "digital humans" mediating task-based interactions [3]. However, gaps persist in understanding ECAs' cultural adaptability, long-term effects on retention, and strategies for maintaining motivation in immersive environments. Recent research highlights the importance of culturally adaptive ECAs to better support diverse learner backgrounds [4-5].

This review synthesizes existing research on ECAs in VR/AR-based language learning environments. It aims to identify key themes, strategies, and research gaps while proposing actionable directions for future innovation. Specifically, this review seeks to address the following research questions:

1. What are the key themes and findings in existing research on ECAs for L2 learning in immersive environments?
2. What are the commonalities and discrepancies in the application of ECAs for L2 learning?

3. What future research directions can address current gaps?

2 Theoretical Background

2.1 Key Themes and Findings in Existing Research

ECAs are AI-driven virtual agents designed to replicate human interaction through multimodal communication. Unlike traditional chatbots, ECAs provide learners with engaging, human-like interactions that enhance the emotional, social, and cognitive aspects of L2 learning [6]. Recent studies demonstrate that ECAs can dynamically adapt their communication styles to learners' proficiency levels and emotional states, effectively reducing language anxiety and fostering active participation [7]. For instance, Sikström et al. (2022) [4] provide a comprehensive review of how pedagogical agents adapt their communication strategies to foster deeper engagement and improved learning outcomes across diverse contexts.

Technological advancements, such as automatic speech recognition (ASR) and affective computing, enable ECAs to interpret non-native accents, emotional cues, and situational contexts with increasing accuracy [8-9]. These capabilities allow ECAs to provide tailored feedback and empathetic interactions, addressing diverse learner needs. For example, an ECA designed for East Asian learners may incorporate honorifics and indirect expressions, whereas one designed for Western learners may employ more direct communication styles.

2.2 ECA Strategies and Social Presence in Immersive Environments

Social presence—the perception of ECAs as responsive and engaging entities—is critical in immersive L2 learning environments. ECAs that exhibit intentionality, empathy, and real-time feedback foster trust and sustain learner motivation [10]. Immersive environments, such as VR and AR, enhance these interactions by providing authentic, task-based learning contexts. For instance, learners can participate in simulated real-world situations, such as ordering food in a restaurant or engaging in business negotiations, where ECAs offer corrective feedback and adaptive language modeling [11].

Moreover, Sikström et al. (2024) [5] highlight how high school teachers and students perceive the role of ECAs in scaffolding learning tasks and adapting feedback based on learners' real-time needs, emphasizing their transformative potential in education. By integrating gamification elements, such as achievement badges, leaderboards, and real-time feedback, ECAs can further enhance motivation and ensure sustained engagement.

However, maintaining learner motivation over time remains a challenge, as the novelty of immersive technologies may wane with repeated use [12]. To counteract this, ECAs should incorporate adaptive difficulty levels and personalized learning trajectories that align with individual learner progress [13].

2.3 Addressing Challenges and Enhancing ECA Integration

To fully realize the potential of ECAs in immersive L2 learning, critical challenges must be addressed. Cultural and linguistic adaptability is essential for ECAs to effectively engage diverse learner populations. For instance, ECAs designed for East Asian learners may incorporate indirect communication patterns and hierarchical politeness strategies, while those for Western learners might adopt a more explicit and egalitarian approach. Additionally, ECAs tailored for younger learners can include gamification elements, such as adaptive storytelling, achievement badges, and interactive role-playing activities, whereas those for adults may focus on structured conversational simulations and industry-specific language applications. ECAs can thus ensure that their interactions are both contextually relevant and pedagogically effective, ultimately leading to improved engagement and learning outcomes across diverse learner demographics. Culturally responsive ECAs can tailor their feedback and communication styles to align with learners' norms and expectations, fostering inclusivity and relevance [14].

Further, integrating advanced affective computing technologies enables ECAs to interpret and respond to emotional states, such as frustration or excitement, creating supportive learning environments that enhance retention and motivation. By recognizing non-verbal cues, such as hesitation, prolonged pauses, or fluctuations in speech intonation, ECAs can dynamically adjust their responses—offering encouragement, slowing down explanations, or rephrasing content to scaffold language learning more effectively.

Technological advancements in natural language processing (NLP) and multimodal interaction design are equally crucial. High-fidelity ECAs that seamlessly integrate real-time feedback, synchronized gestures, and voice modulation provide learners with more immersive and realistic experiences. For instance, ECAs that incorporate synchronized lip movements, culturally adaptive gestures, and prosodic modulation can significantly enhance authenticity, leading to improved engagement and deeper immersion in language learning environments. These

features enable ECAs to support complex, context-rich scenarios that mirror real-world interactions, maximizing their educational impact. Future research should focus on refining these technologies to enhance adaptability across diverse learner profiles, ensuring that ECAs effectively support varying proficiency levels, learning styles, and socio-cultural backgrounds. By addressing these challenges, ECAs can evolve into culturally responsive and emotionally intelligent tools, fostering more inclusive and engaging immersive language learning experiences across diverse learner populations.

3 Methodology

A systematic review adhering to PRISMA guidelines was conducted to ensure transparency and rigor. Academic databases such as IEEE Xplore, Scopus, and ScienceDirect were searched using keywords like “conversational agents,” “extended reality,” and “second language acquisition.” From an initial pool of 1,908 articles, seven studies met the inclusion criteria for in-depth analysis. The inclusion criteria focused on peer-reviewed studies examining the impact of ECAs on L2 learning in immersive environments, emphasizing measurable outcomes such as engagement, fluency, and retention. Studies published in the last decade (2015–2024) were prioritized to ensure the inclusion of recent advancements in ECA development.

The review process involved multiple steps, including title and abstract screening, full-text evaluation, and thematic synthesis. Studies were included based on their relevance to ECAs in immersive language learning, alignment with the research questions, and methodological rigor. Exclusion criteria eliminated articles lacking empirical data, focusing solely on technical implementations or unrelated to L2 acquisition. This systematic approach ensured that the findings were grounded in high-quality research.

4 Results and Discussion

4.1 Key Themes and Findings in Existing Research

This systematic review identified three primary themes related to the integration of ECAs in immersive L2 learning environments:

(1) Communication and Scaffolding: ECAs reduce learner anxiety through empathetic interactions and enhance self-regulation with adaptive feedback. These strategies align with Social Agency Theory, emphasizing relational communication to foster trust and engagement [7, 1].
(2) Immersive Capabilities: The use of VR and AR technologies amplifies the benefits of ECAs by providing realistic, task-based contexts that improve language immersion and retention [11, 9].
(3) Multimodal Interaction: The use of synchronized gestures, facial expressions, and adaptive personalities by ECAs enhances both cognitive and emotional engagement, making interactions more dynamic and impactful [3].

4.2 Commonalities and Discrepancies in ECA Employment

The studies reviewed revealed shared insights into the use of ECAs for immersive L2 learning. Feedback mechanisms and immersive engagement emerged as critical factors. Adaptive feedback, such as personalized hints or corrective responses, fosters motivation and supports self-regulation. For example, Bian and Zhou (2022) [12] demonstrated that feedback tailored to learners' proficiency levels significantly enhances cognitive engagement. Additionally, immersive environments enabled by VR and AR provide context-rich scenarios that simulate real-world interactions, which are indispensable for task-based language learning [11]. These elements underscore the potential of ECAs to deepen learner engagement and improve knowledge retention.

Despite these shared benefits, the studies revealed significant differences in ECA implementation. A key area of divergence is communication modalities. While most studies highlight the importance of multimodal communication, incorporating gestures and facial expressions, others prioritize verbal interactions to enhance fluency and vocabulary acquisition [9]. Differences also emerge in the role of ECA personality traits. Sajjadi et al. (2019) [14] found that extroverted ECAs boost intrinsic motivation, whereas Bian and Zhou (2022) [12] observed that choleric ECAs enhance extrinsic motivation through performance-oriented tasks. These findings suggest the importance of flexible ECA designs that accommodate diverse learner preferences and goals.

Discrepancies also arise in the integration of ECAs across varying educational contexts. For instance, high school learners and university students exhibit differences in technological access and expectations. While Ke et al. (2016) [10] emphasize the importance of high fidelity in immersive tools for university learners, other studies

focus on the adaptability of ECAs for younger audiences. Addressing these contextual variations is essential to broadening the applicability of ECAs in language education.

4.3 Future Directions

Current ECAs often adopt a one-size-fits-all approach, overlooking learners' diverse needs. Future research should prioritize developing ECAs capable of real-time personalization. By leveraging advancements in natural language processing (NLP) and affective computing, ECAs could dynamically adapt their feedback, interaction styles, and personality traits to align with individual learner preferences, emotional states, and proficiency levels. Such adaptability is critical for sustaining engagement and enhancing learning outcomes.

For global effectiveness, ECAs must recognize and respond to cultural nuances in communication. Designing region-specific teaching strategies and localized content can improve learners' connection to and benefit from ECA-mediated interactions [13]. Future research should examine how cultural adaptability affects learner trust and motivation, especially in cross-cultural educational contexts, to ensure ECAs are inclusive and culturally sensitive.

While ECAs demonstrate effectiveness in short-term engagement, their long-term impact on retention and learner autonomy remains underexplored. Future research should investigate the long-term retention effects of ECAs through longitudinal studies assessing language retention over time and evaluating how learners maintain language proficiency after exposure to immersive ECA environments. Additionally, gamification elements such as achievement badges, personalized learning paths, and adaptive difficulty levels could be integrated to sustain engagement over time.

Current research predominantly focuses on traditional learner groups, such as university students, leaving underrepresented populations like adult learners, older adults, and individuals with disabilities underserved. Developing inclusive ECAs that cater to diverse demographics is essential. For instance, multimodal ECAs with visual, auditory, and tactile feedback can support learners with sensory impairments, while simplified interfaces can enhance accessibility for older learners. Expanding research to include these groups will enhance the relevance and accessibility of ECAs.

To overcome existing technical limitations, ECAs must integrate advanced multimodal interactions, including synchronized gestures, facial expressions, and real-time speech recognition. Improving visual realism and minimizing interaction lag are also crucial for enhancing immersion and engagement. These advancements will enable ECAs to simulate lifelike conversational dynamics, improving their effectiveness in real-world language learning.

By addressing these areas, ECAs can evolve into sophisticated, adaptive tools that meet the diverse needs of global learners. Such progress will not only amplify the educational impact of ECAs but also contribute to more inclusive, culturally responsive, and technologically advanced language learning environments.

5 Conclusion

This systematic review highlights the transformative potential of ECAs in immersive L2 learning. By integrating adaptive communication strategies with immersive technologies, ECAs support diverse learner needs, fostering personalized and inclusive educational experiences.

The review identifies three central themes in ECA-mediated language education. First, Communication and Scaffolding illustrate how ECAs balance relational communication to reduce anxiety with task-oriented scaffolding to improve instructional performance. Second, Immersive Learning Capabilities underscore the importance of virtual presence and interactive feedback in VR/AR environments, enhancing engagement, motivation, and real-time skill development. Third, Multimodal and Personality-Driven Communication highlights the role of synchronized gestures, facial expressions, and adaptive personalities in fostering immersion, social presence, and learner satisfaction.

Despite these benefits, critical gaps remain. Emotional engagement through affective computing represents a promising avenue for sustaining learner motivation. Longitudinal studies are needed to evaluate the long-term impact of ECAs on language retention and proficiency. Additionally, culturally adaptive ECAs that cater to diverse learner backgrounds and communication styles are essential for broader global applicability.

By addressing these challenges, ECAs can evolve into transformative tools for bridging classroom instruction and real-world communication. Leveraging advancements in AI, VR, and affective computing, ECAs can deliver adaptive, inclusive, and impactful learning experiences. This review lays the groundwork for future research, offering insights to guide ECA development and implementation in immersive language education. With

continued technological advancements, ECAs promise to redefine language learning, creating engaging and personalized environments for learners worldwide.

References

1. Kim, Y., Baylor, A. L.: Research-based design of pedagogical agent roles: A review, progress, and recommendations. International Journal of Artificial Intelligence in Education 26, 160–179 (2016).
2. Mayer, R. E., Sobko, K., Mautone, P. D.: Social cues in multimedia learning: Role of speaker's voice. Journal of Educational Psychology 95(2), 419–430 (2003).
3. Schroeder, N. L., Adesope, O. O.: A systematic review of pedagogical agents' persona. Journal of Research on Technology in Education 46(3), 229–251 (2014).
4. Sikström, P., et al.: How pedagogical agents communicate with students: A two-phase systematic review. Computers & Education 188, 104564 (2022).
5. Sikström, P., et al.: Pedagogical agents communicating and scaffolding students' learning: High school teachers' and students' perspectives. Computers & Education 222, 105140 (2024).
6. Cassell, J., Bickmore, T., Campbell, L., Vilhjalmsson, H., Yan, H.: Human conversation as a system framework: Designing embodied conversational agents. In: Embodied Conversational Agents, pp. 29–63. MIT Press (2000).
7. Ayedoun, E., Hayashi, Y., Seta, K.: A conversational agent to encourage willingness to communicate in SLA. Procedia Computer Science 60, 1433–1442 (2015).
8. Bear, E., Bodnar, S., Chen, X.: Learner and linguistic factors in commercial ASR use for spoken language practice: A focus on form. In Proceedings of the 9th Workshop on Speech and Language Technology in Education (SLaTE), pp. 161–165 (2023).
9. Griol, D., Sanchis, A., Molina, J. M., Callejas, Z.: Developing enhanced conversational agents for social virtual worlds. Neurocomputing 354, 27–40 (2019).
10. Ke, F., Lee, S., Xu, X.: Teaching training in a mixed-reality integrated learning environment. Computers in Human Behavior 62, 212–220 (2016).
11. Fryer, L. K., Nakao, K., Thompson, A.: Chatbot learning partners: Connecting learning experiences, interest, and competence. Computers in Human Behavior 93, 279–299 (2019).
12. Bian, Y., Zhou, C.: Motivation effect of animated pedagogical agent's personality and feedback strategy types on learning in virtual training environment. Virtual Reality & Intelligent Hardware 4(2), 153–172 (2022).
13. Lan, Y.-J.: Immersion, interaction, and experience-oriented learning: Bringing virtual reality into FL learning. Language Learning & Technology 24(1), 1–15 (2020).
14. Sajjadi, P., Hoffmann, L., Cimiano, P., Kopp, S.: A personality-based emotional model for embodied conversational agents: Effects on perceived social presence and game experience of users. Entertainment Computing 32, 100313 (2019).