HAYEF: Journal of Education, 23; 1-12

DOI: 10.5152/hayef.2026.26063

Received: June 16, 2026
Revision Requested: June 29, 2026
Last Revision Received: July 9, 2026
Accepted: July 10, 2026
Publication Date: August 27, 2026

HAYEF: Journal of Education

ORIGINAL ARTICLE

Artificial Intelligence–Enhanced Virtual Reality in English as a Foreign Language Teaching: Instructor Experiences and Pedagogical Orchestration

Cemil Gökhan KARACAN[1], **Tuncer CAN**[2]

[1]Department of English Language and Literature, İstanbul Beykent University Faculty of Arts and Sciences, İstanbul, Türkiye
[2]Department of English Language Teaching, İstanbul University-Cerrahpaşa Hasan Ali Yücel Faculty of Education, İstanbul, Türkiye

What is already known on this topic?

- *Virtual reality can support language learning by simulating realistic communicative situations and increasing learner engagement.*
- *Artificial intelligence can support adaptive feedback, conversational practice, and individualized learning in educational settings.*
- *Existing research on immersive language learning has mainly emphasized learner outcomes, motivation, anxiety reduction, and technological affordances.*

What this study adds on this topic?

- *This study foregrounds instructors' pedagogical reasoning and classroom orchestration decisions in artificial intelligence–enhanced virtual reality lessons.*
- *It shows that instructors valued responsive artificial intelligence agents for back-and-forth language practice, while also identifying trust, infrastructure, and access as decisive constraints.*
- *It reframes artificial intelligence–enhanced virtual reality integration as a teacher-mediated pedagogical process rather than a simple technological upgrade.*

Abstract

Artificial intelligence is increasingly reshaping immersive educational environments; yet, empirical attention to its integration with virtual reality in English as a foreign language teaching remains limited. This qualitative descriptive study examined how English as a foreign language instructors perceived, implemented, and evaluated artificial intelligence–enhanced virtual reality activities in their teaching practices. The study was conducted in Türkiye with eight English as a foreign language instructors from two universities who had prior experience using virtual reality-based or virtual reality-compatible artificial intelligence–supported speaking, role-play, and communication activities. Data were collected through eight semi-structured interviews and five optional stimulated-recall sessions based on short lesson-recording excerpts or demonstration recordings. The data were analyzed through inductive qualitative content analysis supported by a large language model–assisted preliminary open-coding process, followed by manual code refinement, codebook development, peer debriefing, member checking, and intercoder reliability procedures. The analysis generated five interpretive themes: from scripted simulation to contingent conversation, lowering the social cost of speaking, classroom realism without real-world risk, orchestrating classroom visibility and fragile infrastructure, and pedagogically accountable artificial intelligence–virtual reality adoption. Instructors viewed artificial intelligence–enhanced virtual reality as pedagogically promising because, in their accounts, it enabled learners to converse with responsive computer-generated agents, increased engagement and attendance, supported lower-pressure speaking practice, and created opportunities for real-life communication. However, they also emphasized practical and pedagogical constraints related to institutional Wi-Fi authentication, mobile hotspot instability, headset management, screen-casting, projection quality, classroom visibility, task alignment, and the need for teacher mediation of automated feedback. The study suggests that the value of artificial intelligence–enhanced virtual reality in English as a foreign language teaching depends less on technological novelty than on instructional design, institutional readiness, and sustained teacher professional development.

Keywords: Artificial intelligence, English as a foreign language, immersive learning, teacher experience, virtual reality

Introduction

Virtual reality (VR) has become one of the most visible immersive technologies in language education because it can situate learners inside simulated communicative environments that would otherwise be difficult, costly, or anxiety-provoking to reproduce in ordinary classrooms (Hua & Wang, 2023; Kaplan-Rakowski & Gruber, 2023; Parmaxi, 2023). In English as a foreign language (EFL) teaching, this capacity is particularly relevant because learners often need opportunities to rehearse speaking, listening, and pragmatic interaction beyond the limits of textbook dialogues and teacher-led classroom exchanges. Al traditional digital tools usually present language input on a screen, VR places learners inside a simulated space where they can navigate, act, speak, and respond as if they were participating in a social situation.

Recent developments in artificial intelligence (AI), especially large language models and conversational agents, have intensified the pedagogical significance of VR. Earlier VR language-learning environments were often constrained by pre-scripted dialogue trees and limited interactional pathways. In contrast, AI-supported agents can generate more responsive turns, adapt to learner input, and provide immediate feedback on language use (Godwin-Jones, 2025; Hemminki-Reijonen et al., 2025; Xiao et al., 2023). This shift changes the educational value proposition of VR: the core issue is no longer only whether learners feel immersed in a simulated environment, but whether they can engage in meaningful, contingent, and pedagogically useful communication within that environment.

Corresponding author: Cemil Gökhan Karacan **E-mail:** gokhankaracan@beykent.edu.tr

Cite this article as: Karacan, C. G., & Can, T. (2026). Artificial intelligence–enhanced virtual reality in English as a foreign language teaching: Instructor experiences and pedagogical orchestration. *HAYEF: Journal of Education*, 23, 0063, doi: 10.5152/hayef.2026.26063.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
Available online at https://hayefjournal.org

Augmented reality (AR), mixed-reality systems, and metaverse-oriented environments may later intersect with AI-driven language learning through object-aware interfaces, smart glasses, contextual overlays, and more persistent immersive spaces. Sarı (2022) similarly cautions that the educational value of metaverse-based language learning depends on whether immersive environments support authentic language use rather than becoming another buzzword in educational technology. However, the current classroom-ready evidence base and commercially visible language-learning use cases are more mature for VR-based conversational, role-play, and simulation environments than for widely available AI-enhanced AR language-teaching applications. Therefore, the present study focuses empirically on AI-enhanced VR in EFL teaching while acknowledging that future immersive classrooms may involve more fluid combinations of VR, AR, and AI.

Virtual Reality Meets Artificial Intelligence in English as a Foreign Language Teaching

Virtual reality is commonly defined as a computer-generated three-dimensional environment that users can enter and interact with through head-mounted displays or related devices (Marougkas et al., 2023). In language education, VR is associated with immersion, presence, contextualized practice, and opportunities for embodied interaction (Hua & Wang, 2023; Tai et al., 2022). These features can support vocabulary learning, speaking practice, listening comprehension, pronunciation work, and pragmatic rehearsal. A recurrent argument in VR-assisted language learning is that learners may feel more willing to communicate when the environment provides realism without the full social consequences of real-world performance, a point that resonates with willingness to communicate and foreign-language classroom anxiety research (Horwitz et al., 1986; Kaplan-Rakowski & Gruber, 2023; MacIntyre et al., 1998).

This focus also has a local scholarly lineage in English Language Teaching (ELT) research in Türkiye. Before the recent convergence of AI and VR, Can and Şimşek (2015) had already examined three-dimensional virtual learning environments in foreign language pre-service teacher education, showing that virtual worlds could extend language-teaching possibilities while still requiring careful pedagogical preparation, technical orientation, and teacher support. This study builds on this earlier immersive-learning trajectory by examining how AI-driven conversational responsiveness reshapes the instructional possibilities and classroom demands of VR-based EFL activities.

Artificial intelligence extends this potential by making the virtual environment more conversationally responsive. Before recent AI integration, many VR language-learning environments were limited by closed-ended content, pre-scripted dialogue, and restricted interactional pathways. In contrast, when learners interact with an AI-powered virtual assistant, tutor, character, or avatar, they are not merely selecting from fixed options; they may experience a more natural turn-taking sequence in which the virtual interlocutor responds to their utterances and the interaction is no longer confined to a narrow set of predefined communicative moves. In pedagogical terms, this may increase speaking volume, provide repeated low-stakes rehearsal, and allow learners to experiment with language before participating in more demanding real-world communication. This is especially important in EFL contexts where opportunities for authentic English interaction outside the classroom may be limited.

At the same time, AI-enhanced VR raises implementation and quality-control questions. Virtual reality requires hardware, stable internet connectivity, classroom management strategies, and time for setup. Artificial intelligence feedback introduces additional concerns about accuracy, appropriateness, transparency, and teacher oversight (Chen et al., 2022b; Zawacki-Richter et al., 2019). For instructors, the pedagogical challenge is therefore not simply to use a headset or an AI agent. It is to transform those affordances into a coherent language-learning task with clear objectives, manageable sequencing, reliable feedback, and post-activity consolidation.

Instructor Experience

A substantial body of research has examined VR-assisted language learning in relation to learner motivation, engagement, anxiety, vocabulary learning, and performance outcomes, as well as learner strategy use, self-efficacy, and motivational conditions in technology-mediated learning (Chen et al., 2022a; Dikici & Altaş, 2024; Hua & Wang, 2023; Kıyançiçek & Çoban, 2024; Parmaxi, 2023; Tai et al., 2022). This literature is valuable, but it does not fully explain how instructors interpret AI-enhanced VR tools, how they decide whether a virtual task is pedagogically worthwhile, or how they manage the practical realities of implementation, partly because it is a recent development. Educational technology integration depends on tool affordances as well as teachers' beliefs, technological pedagogical knowledge, confidence, institutional context, and classroom decision-making (Ertmer & Ottenbreit-Leftwich, 2010; Mishra & Koehler, 2006; Tseng et al., 2022).

This gap is especially important for AI-enhanced VR because the technology appears impressive to both teachers and students. The sense of novelty may generate engagement, but novelty alone cannot guarantee language learning. Recent ELT-oriented work on AI integration similarly emphasizes that language educators may view AI tools as pedagogically useful while still requiring training, support, and critical guidance for classroom implementation (Can & Mangır, 2024). In AI-enhanced VR, this need becomes even more complex because instructors must decide whether students are actually using target forms, whether feedback is reliable, whether the VR scenario aligns with course outcomes, and whether the activity remains manageable in a physical classroom with limited devices. Instructors are therefore not passive users of AI-enhanced VR; they are pedagogical orchestrators who determine whether immersive and AI-mediated interaction becomes a meaningful language-learning opportunity.

Conceptual Framework: Teacher-Mediated Artificial Intelligence–Virtual Reality Orchestration

This study is informed by a teacher-mediated orchestration view of educational technology integration, where orchestration refers to the teacher's real-time management of multilayered learning activities, tools, classroom constraints, and learner participation within complex instructional settings (Dillenbourg, 2013). In this view, AI-enhanced VR is not treated as a self-sufficient learning solution; rather, it becomes pedagogically meaningful when immersive affordances, conversational AI, task design, infrastructure, and teacher judgment are aligned. The framework draws on technology integration scholarship emphasizing that technology use is shaped by teacher knowledge, confidence, beliefs, and contextual conditions (Ertmer & Ottenbreit-Leftwich, 2010; Mishra & Koehler, 2006), as well as VR-assisted language-learning research that foregrounds immersion, presence, situated practice, and reduced anxiety (Hua & Wang, 2023; Kaplan-Rakowski & Gruber, 2023; Parmaxi, 2023).

Within this framework, AI contributes conversational agency to VR by enabling virtual interlocutors to respond, ask follow-up questions, and provide immediate feedback. However, these affordances remain conditional. Teacher mediation is required to align VR activities with language objectives, orchestrate headset use, classroom visibility, and infrastructural constraints, monitor learner interaction, validate AI-generated feedback, and consolidate language learning after the immersive task. Figure 1 summarizes this conceptual framing by positioning instructor orchestration between VR affordances, AI conversational agency, contextual conditions, and pedagogically accountable classroom use.

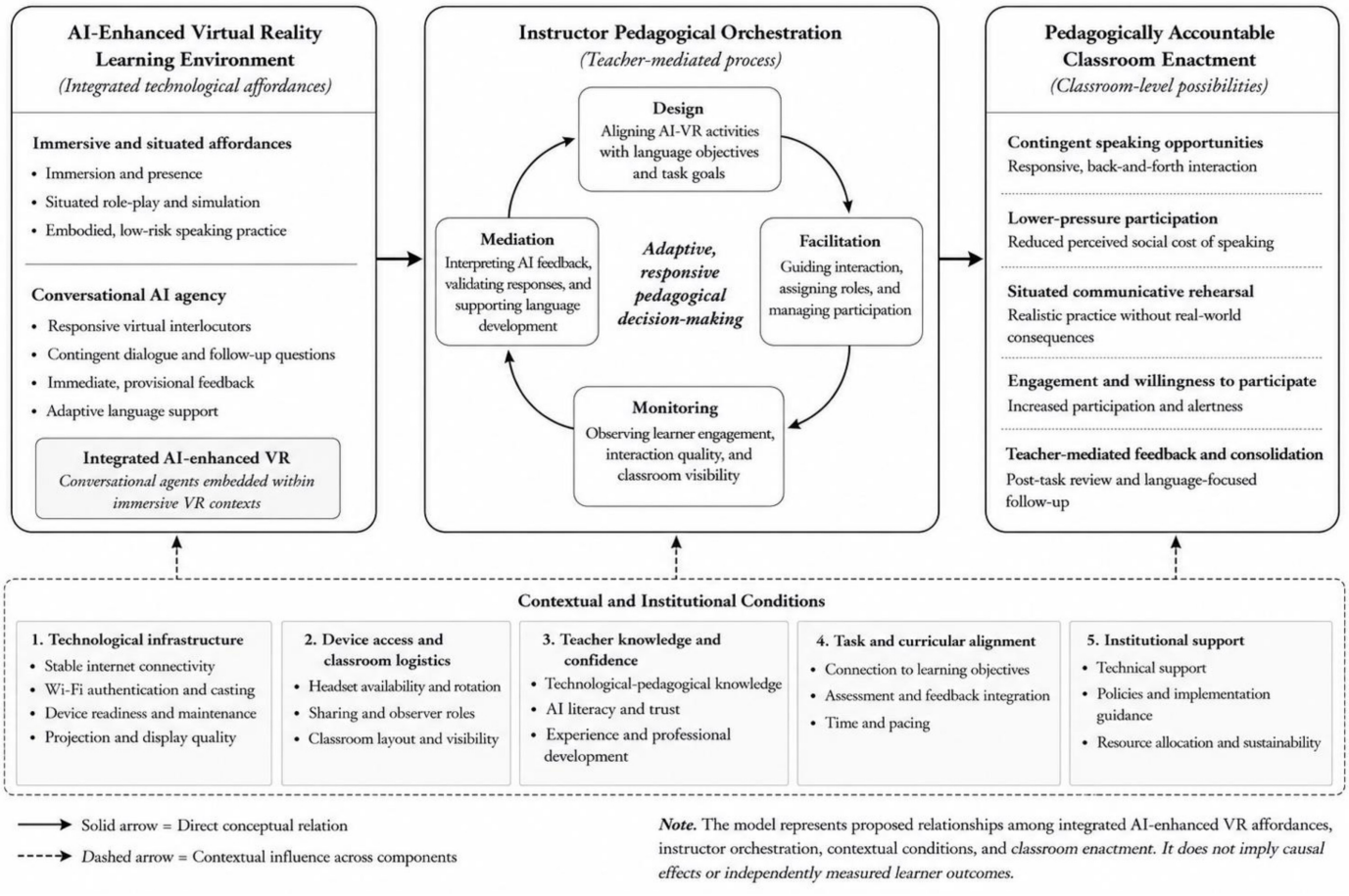

Figure 1.
Conceptual Model of Teacher-Mediated AI-VR Orchestration in EFL Teaching.

To address this gap, the current study examined EFL instructors' experiences, perceptions, and pedagogical reasoning concerning AI-enhanced VR activities. The study was guided by the following research questions:

1. How do EFL instructors perceive and implement AI-enhanced VR activities in their teaching practices?
2. What pedagogical, technical, and institutional conditions shape instructors' integration of AI-enhanced VR into EFL instruction?

Methods

Research Design

This study adopted a qualitative descriptive design (Sandelowski, 2000) to examine EFL instructors' experiences, perceptions, and pedagogical reasoning in relation to AI-enhanced VR lesson implementation. This design was chosen as the study aimed to generate a direct, practice-oriented account of how instructors interpreted and used an emerging educational technology in real classroom conditions rather than to test a predetermined theory, isolate the independent effects of specific technological components within AI-enhanced VR, or measure learning gains. The analysis followed inductive qualitative content analysis procedures, including repeated reading, meaning-unit identification, open coding, category development, abstraction, and interpretation (Elo & Kyngäs, 2008; Hsieh & Shannon, 2005).

Participants and Context

Eight EFL instructors from two universities in Türkiye participated in the study. Participants included five female and three male instructors, who ranged in age from 30 to 48 years (mean [M] = 38.6, standard deviation = 6.2), and had between 7 and 22 years of teaching experience (M = 12.4 years). They were recruited through purposive sampling because they had prior experience implementing AI-supported VR-based or VR-compatible speaking, role-play, and communication activities in EFL courses. Twelve instructors were initially contacted through departmental invitations, and eight volunteered to participate. No exclusion criteria were applied other than prior experience with AI-enhanced VR integration in language teaching. Participant selection followed a purposive, criterion-based logic aimed at identifying information-rich cases with direct experience of the phenomenon under investigation (Patton, 2015; Palinkas et al., 2015). In line with the principle of information power, the focused aim of the study, the specificity of the participant group, and the use of interview and stimulated-recall data supported a smaller, more specialized sample rather than a broad or statistically representative one (Malterud et al., 2016). This was particularly relevant because AI-enhanced VR use in EFL classrooms remains an emerging and unevenly distributed practice, making instructors with direct classroom implementation experience a limited but analytically valuable participant group. The participating institutions represented different higher education contexts, but institutional names and identifying details were removed to maintain blinded review and participant confidentiality. In the Results section, participants are identified only through anonymized codes (P1–P8). Departmental and institutional units are described generically as university EFL units to maintain blinded review and protect confidentiality.

Data Collection

Data were collected between September 15, 2025 and January 15, 2026 during the Fall semester of the 2025–2026 academic year. The data collection process included semi-structured interviews, optional stimulated-recall sessions, and classroom recordings of AI-enhanced

VR practices; however, these recordings were used as prompts during stimulated recall rather than as independently coded video data.

The interview protocol included open-ended questions about participants' experiences with AI-enhanced VR activities, perceived benefits for language learning, classroom implementation decisions, technical constraints, feedback reliability, learner engagement, and expectations for future use. Each interview lasted approximately 30–45 minutes and was conducted individually. The interview protocol is provided as Supplementary Material 1.

During stimulated-recall sessions, five instructors (P1, P3, P4, P6, and P7) viewed a short recording excerpt from a VR-based lesson or demonstration activity they had conducted and narrated their pedagogical decision-making for approximately 12–20 minutes. The remaining three instructors did not participate in stimulated recall because they either did not have a shareable recording excerpt or their institution did not allow for such recording. These sessions were used to elicit more situated accounts of instructional choices, classroom monitoring, learner reactions, technical breakdowns, and post-task consolidation. The stimulated-recall protocol is provided as Supplementary Material 2. All interviews and stimulated-recall sessions were audio-recorded with consent. Before analysis, recordings were transcribed verbatim using Otter.ai, manually checked for accuracy, anonymized, and stored in a password-protected digital folder. For analysis, the transcripts were first prepared as verbatim records. In reporting participant quotations, a lightly edited verbatim approach was used: filler sounds, false starts, repeated words, and minor speech disfluencies were removed where necessary to improve readability, while the substantive meaning, tone, and analytic relevance of participants' accounts were preserved.

Data Analysis

The unit of analysis was the meaning unit, defined as a phrase, sentence, or short passage that conveyed a relevant idea about instructors' perceptions, classroom implementation decisions, pedagogical reasoning, challenges, or future expectations regarding AI-enhanced VR in EFL teaching. After repeated reading of the transcripts, a large language model (LLM) was used to generate preliminary open-code suggestions by identifying recurring ideas, tentative categories, and possible thematic clusters. In line with recent methodological discussions on the use of LLMs to support qualitative textual analysis, these outputs were treated as analytical prompts rather than final codes (Tai et al., 2024; Than et al., 2025). In reporting participant evidence, quotations were attributed using anonymized participant codes and data-source labels such as P2, semi-structured interview, or P4, stimulated recall.

Large language model–assisted preliminary coding was used as an analytic support procedure. After all transcripts had been anonymized, ChatGPT-5.5 Thinking version was used to generate preliminary open-code suggestions during the early stage of analysis. The prompts asked the model to identify recurring meaning units, tentative codes, possible category labels, and relationships among instructors' accounts of AI-enhanced VR implementation. The prompt structure followed a staged format: first, identification of meaning units; second, generation of provisional open codes; third, clustering of related codes; and fourth, suggestion of possible thematic patterns. These outputs were not treated as findings or final codes. All AI-generated suggestions were reviewed, revised, merged, renamed, rejected, or retained by the researchers through manual qualitative analysis. Suggestions that were too broad, repetitive, insufficiently grounded in participant accounts, or inconsistent with the research questions were set aside, while data-grounded suggestions were retained, renamed, merged, or refined through constant comparison. Final coding decisions, theme construction, interpretation, and reporting remained under full researcher control.

The researchers then conducted manual coding through constant comparison. Codes were refined, merged, renamed, or removed as the researchers moved between transcripts, coded segments, and emerging categories. A consolidated codebook was developed, including 18 final codes, operational definitions, inclusion and exclusion criteria, and representative evidence slots. Using MAXQDA, the full dataset was re-coded with the finalized codebook. This process generated five interpretive themes reflecting instructors' accounts of AI-enhanced VR: from scripted simulation to contingent AI conversation, lowering the social cost of speaking, classroom realism without real-world risk, orchestrating classroom visibility and fragile infrastructure, and toward pedagogically accountable AI-VR adoption.

To strengthen analytic dependability, a second coder independently coded 20% of the dataset using the finalized codebook. Intercoder agreement was calculated in MAXQDA using Cohen's kappa across the 18-code codebook, with the meaning unit as the coded unit, resulting in a coefficient of 0.82. Coding disagreements were discussed until consensus was reached, and the codebook was refined where necessary. The final themes were developed through researcher interpretation, code comparison, peer debriefing, member checking, and repeated movement between coded data extracts and the research questions.

Trustworthiness, Reflexivity, and Ethical Considerations

Trustworthiness was addressed through credibility, transferability, dependability, and confirmability (Lincoln & Guba, 1985; Nowell et al., 2017; Stahl & King, 2020). Credibility was supported through multiple qualitative data sources, including semi-structured interviews and stimulated-recall sessions where available. Member checking was conducted with four selected participants to verify the semantic accuracy of key interpretations and to ensure that the themes reflected participants' intended meanings rather than researcher or LLM-generated assumptions, while peer debriefing was used to examine emerging codes, thematic decisions, and alternative explanations.

Transferability was supported by providing contextual information about the participants, sampling logic, teaching experience, and prior AI-enhanced VR implementation background. Rather than claiming statistical generalizability, the study offers sufficient contextual detail to allow readers to judge the relevance of the findings for similar EFL, teacher education, and educational technology contexts.

Dependability was strengthened through a documented analytic procedure. This included verbatim transcription, manual checking of transcripts, anonymization before analysis, LLM-assisted preliminary open-code suggestions, manual code refinement, codebook development, MAXQDA-supported re-coding, and an intercoder agreement check on 20% of the dataset. These procedures created an audit trail of key analytic decisions and supported the consistency of the analysis.

Confirmability was addressed through reflexive memoing and researcher review of all AI-generated analytical suggestions. Reflexive memos were used to document assumptions, interpretive decisions, and potential researcher biases related to educational technology, language teaching, and AI-supported analysis. Large language model outputs were treated only as preliminary analytical prompts; final coding decisions, theme construction, interpretation, and reporting remained under researcher control. This helped ensure that the findings were grounded in participants' accounts rather than in researcher preference or automated output. Reflexivity was addressed through an explicit researcher positionality statement and ongoing memoing during data analysis.

The research team approached the study from the position of EFL teacher educators and educational technology researchers with prior interest in immersive learning, digital pedagogy, and AI-supported language teaching. This background supported sensitivity to the

pedagogical possibilities of AI-enhanced VR, but it also created a risk of overemphasizing technological promise. To manage this risk, the researchers used reflexive memoing to document assumptions, explicitly searched for negative and contradictory cases during coding, and separated technology enthusiasm from evidence of pedagogical value. Peer debriefing and member checking further helped challenge premature interpretations and keep the analysis grounded in participants' accounts.

The study also sought to enhance transparency, reporting quality, and methodological coherence (Tracy, 2010). Transparency was supported by clearly describing the sampling logic, data collection procedures, unit of analysis, coding process, LLM role, codebook development, and intercoder agreement procedure. Methodological coherence was addressed by aligning the qualitative descriptive design, research questions, purposive sampling strategy, interview and stimulated-recall data, and inductive qualitative content analysis with the study's aim of producing a direct, practice-oriented account of instructors' experiences with AI-enhanced VR in EFL teaching.

Ethical committee approval was received from the Scientific Research and Publication Ethics Committee of Istanbul Aydın University (Approval no: 162512, Date: August 1, 2025). Written informed consent was obtained from all participating instructors. Transcripts were anonymized; identifying institutional or personal information was removed from the analytic dataset, and the data were stored in a password-protected digital folder.

Results

The analysis generated five interpretive themes concerning instructors' experiences with AI-enhanced VR in EFL teaching. Across the dataset, instructors expressed cautious optimism. They were often impressed by the possibility that learners could talk back and forth with a computer-generated virtual agent that was not a real person yet could still sustain a meaningful conversational exchange. At the same time, instructors repeatedly emphasized that the pedagogical value of such activities depended on task design, teacher orchestration, infrastructure, and equitable access to functioning equipment. Table 1 presents an overview of the themes, core patterns, and pedagogical meanings.

From Scripted Simulation to Contingent Artificial Intelligence Conversation

The first theme concerned instructors' perception that AI changed the communicative status of VR. Participants did not describe AI-enhanced VR merely as a visual simulation, a game-like environment, or an attractive classroom novelty. Instead, they focused on learners' ability to speak with a responsive computer-generated interlocutor. Several instructors contrasted this with earlier digital role-play experiences in which learners followed predetermined dialogue pathways, clicked fixed options, or responded to scripted prompts. In AI-enhanced VR, by contrast, the virtual interlocutor could respond to learner input, continue the exchange, ask follow-up questions, and provide language support.

This shift was interpreted as a movement from practice inside a simulated space to interaction with a simulated social actor. One instructor described the virtual character as no longer "just a character standing there" because it responded in a way that "matched the student's sentence" (P3, semi-structured interview). The value of the activity therefore did not lie only in visual immersion. It lay in the perception that the student had entered an unfolding exchange in which the next turn depended partly on what the learner had actually said.

> The surprising part was that it was not just a character standing there. The student said something, and the assistant answered in a way that matched the student's sentence. Then it asked another question. At that moment, it became more like a conversation than a VR activity. (P3, semi-structured interview)

This account shows that instructors did not interpret AI-enhanced VR merely as a more attractive version of conventional digital role-play. What appeared pedagogically significant was the shift from controlled response selection to interactional contingency. Learners had to listen, interpret, and respond to an unfolding exchange rather than complete a closed practice sequence. This also increased the amount of language learners had to produce because the AI agent could continue the interaction through follow-up questions.

Stimulated-recall data added a more situated view of this process. While watching a classroom excerpt, one instructor noticed that the learner was "not waiting for my correction" but was instead listening to the avatar and preparing a response (P6, stimulated recall). This observation suggests that AI-enhanced VR may temporarily redistribute interactional control: the teacher remains pedagogically responsible, but the immediate conversational turn can be carried by the virtual interlocutor.

> When I watched the recording again, I noticed that the learner was not waiting for my correction. He was listening to the avatar,

Table 1.
Overview of Analytical Themes, Core Patterns, Participant Evidence, and Pedagogical Meanings

Analytical Theme	Core Pattern	Illustrative Participant Evidence	Pedagogical Meaning
From scripted simulation to contingent AI conversation	AI changes VR from a pre-scripted environment into a responsive conversational space.	P3 semi-structured interview; P6 stimulated recall	AI-enhanced VR may expand speaking practice through simulated but responsive interaction.
Lowering the social cost of speaking	Students appear more relaxed, engaged, and willing to participate when speaking to virtual agents.	P2 semi-structured interview; P4 stimulated recall	VR can function as an affective scaffold for reluctant speakers.
Classroom realism without real-world risk	Teachers use VR to stage realistic service, travel, workplace, and social interactions inside the classroom.	P1 stimulated recall; P5 semi-structured interview	Simulated scenarios can support situated rehearsal before real-world communication.
Orchestrating classroom visibility and fragile infrastructure	Institutional Wi-Fi, hotspot instability, casting, display quality, and headset orchestration shape whether VR becomes a shared classroom activity.	P1, P2, P5, P6, and P7 semi-structured interviews; P4 stimulated recall	Infrastructure and classroom visibility are pedagogical variables, not merely logistical issues.
Toward pedagogically accountable AI-VR adoption	Teachers value AI feedback but insist on human mediation, validation, and task alignment.	P3 semi-structured interview; P6 stimulated recall	AI-enhanced VR should augment teacher judgment rather than replace it.

Semi-structured interview excerpts were used primarily to represent participants' general perceptions, rationales, and retrospective evaluations, whereas stimulated-recall excerpts were used to illuminate situated decision-making while instructors revisited concrete classroom moments.
Note: AI = Artificial intelligence; VR = Virtual reality.

> thinking, and then trying to answer. The conversation moved in a direction that I had not planned sentence by sentence. (P6, stimulated recall)

The theme therefore points to a central affordance of AI-enhanced VR: it can create a simulated but responsive interactional space. Instructors did not claim that this replaced authentic human communication. Rather, they viewed it as a bridge between solitary language practice and more socially demanding forms of communication with peers, teachers, or real-world interlocutors.

Lowering the Social Cost of Speaking

The second theme concerned instructors' perceptions of learner engagement, attendance, and reduced speaking pressure. Unlike the earlier version of VR as a visual novelty, AI-enhanced VR was described as changing the emotional conditions of classroom speaking. One instructor noted that the lesson had "a different energy in the room" and that even quieter students wanted to participate because the activity "felt less formal" (P2, semi-structured interview). Instructors interpreted students' willingness to speak as connected to curiosity about the headset and to the lower interpersonal risk of speaking to a virtual agent.

Instructors consistently perceived AI-enhanced VR as being associated with higher participation, especially among students who were usually reluctant to speak in traditional classroom settings. Because students were interacting with a virtual agent rather than being evaluated directly by a teacher or exposed to the immediate judgment of classmates, they appeared more willing to try, make mistakes, and continue speaking. In instructors' accounts, the virtual interlocutor created pressure to respond, but that pressure was experienced as less socially threatening.

> When they knew VR was part of the lesson, more students came and they were more alert. There was a different energy in the room. But it was not only excitement; even quiet students wanted to try speaking because it felt less formal. (P2, semi-structured interview)

This extract shows that engagement was not described as simple entertainment. The instructor connected attendance and alertness with a classroom atmosphere in which speaking felt less formal. The perceived value of reduced pressure was especially visible in relation to error tolerance. Instructors observed that learners were more willing to produce imperfect language when the immediate audience was a virtual agent.

Stimulated-recall excerpts made this affective shift more concrete. In one recorded moment, a participant observed that a student who would normally stop after one mistake "smiled and tried again" while speaking to the virtual person (P4, stimulated recall). The activity for the instructor appeared to normalize hesitation, repair, and repetition. Rather than waiting until they could produce a fully accurate utterance, students attempted to communicate and then adjusted their language through feedback, repetition, or teacher-led consolidation after the VR task.

> In the video, I can see that she hesitates first, but then she continues. Normally she stops after one mistake. Here, because she was talking to the virtual person, she smiled and tried again. That was important for me. (P4, stimulated recall)

However, instructors did not equate engagement with learning automatically. They were aware that students could become excited because of the headset, the novelty of the environment, or the visual realism of the task. For this reason, several participants emphasized the need to connect VR activities to explicit language goals, such as using target vocabulary, practicing functional expressions, rehearsing interview responses, or developing turn-taking strategies. Engagement was therefore treated as a useful condition for learning, not as sufficient evidence that learning had occurred.

Classroom Realism Without Real-World Risk

The third theme involved the creation of real-life communicative experiences inside the physical classroom. Instructors valued AI-enhanced VR because it allowed them to simulate situations that would normally be difficult to stage in an ordinary EFL classroom. These included service encounters, interviews, travel situations, workplace exchanges, informal conversations, and problem-solving tasks. One participant summarized this value by describing a VR moment as "an airport problem" in which the student needed to "ask, clarify, and respond" (P1, stimulated recall).

This sense of realism was not understood as perfect authenticity. Rather, instructors described it as pedagogically useful rehearsal. The classroom remained physically the same, but the communicative frame changed. Students could temporarily inhabit a cafe, airport, hotel, meeting room, or social setting and practice language in relation to a situated goal. For instructors, the value was strongest when the VR task involved a clear communicative purpose rather than simply exploring a virtual environment.

> Watching the excerpt, I can see why the task worked. The classroom did not physically change, but for the student the situation became an airport problem. He needed to ask, clarify, and respond. It gave the language a real purpose. (P1, stimulated recall)

The quotation illustrates how VR helped instructors relocate real-world communicative situations into the classroom without requiring real-world exposure. The pedagogical value here lies not in perfect authenticity, but in situated rehearsal: learners could act within a recognizable social scenario while still remaining in a protected instructional environment. This was particularly useful in EFL contexts where opportunities for regular English interaction outside the classroom may be limited.

Several instructors also emphasized that realism depended on task purpose. A visually rich environment was not enough if students were merely looking around. One instructor explained that the cafe scenario worked because students were "not only saying isolated sentences" but had to order, respond, and repair the conversation after misunderstanding (P5, semi-structured interview). In this sense, realism was treated as interactional and functional, not only visual.

> The cafe scenario worked well because students were not only saying isolated sentences. They had to order, respond to a question, and repair the conversation when they misunderstood. That kind of context is difficult to create with a worksheet. (P5, semi-structured interview)

The theme also revealed the importance of post-VR consolidation. Instructors emphasized that the VR experience became pedagogically meaningful when it was followed by classroom reflection, vocabulary review, correction of recurring errors, discussion of alternative expressions, or peer comparison of strategies. Without such consolidation, the activity risked remaining an impressive but isolated experience. With consolidation, the simulated encounter became material for language awareness and further practice.

Orchestrating Classroom Visibility and Fragile Infrastructure

The fourth theme concerned the practical orchestration of AI-enhanced VR in real classrooms. Instructors emphasized that

implementation was not simply a matter of having enough headsets or choosing an appropriate application, though it is important to note that suitable applications remain extremely limited. Rather, the lesson depended on a fragile chain of classroom conditions: institutional Wi-Fi access, headset connectivity, laptop connection, screen-casting, display quality, classroom location, and the teacher's ability to monitor what the student inside the headset was doing. Another instructor noted that successful implementation often depended on whether "everything talks to everything else" before the lesson even began (P6, semi-structured interview). In this sense, infrastructure was not a background condition; it directly shaped whether the VR activity could function as a shared language-learning task.

The importance of visibility became especially clear when casting failed. One participant described the moment as a shift from language teaching to technical troubleshooting, noting that "sometimes the actual problem is not language; it is the connection" (P4, stimulated recall). When the casted screen disappeared, the teacher could no longer monitor the learner's interaction, and the rest of the class lost access to the task. At that point, the teacher had to step in and put the headset back on to arrange the connection again, which created a breakdown in the flow of the lesson. In this sense, the whole pedagogy could depend on "one small yet decisive technical detail" (P4, stimulated recall).

Institutional Wi-Fi added another layer of difficulty. Although universities may provide instructors with username-and-password access, connecting VR devices to institutional networks was not always straightforward. One participant explained that "the university Wi-Fi works fine on my laptop, but getting the headset connected is a completely different story" (P5, semi-structured interview). In practice, networks may require additional authentication steps, certificate details, or configuration procedures that are difficult to complete on a headset and are not always explained through clear institutional guidelines. Another instructor noted that "there was no clear guidance for connecting standalone VR devices to the campus network" (P5, semi-structured interview). When institutional Wi-Fi could not be used, teachers sometimes had to rely on mobile hotspots. This workaround created further problems when indoor mobile coverage was weak, when classrooms were located on lower floors, or when both the headset and laptop had to share the same mobile connection. As one participant observed, "if the hotspot signal drops, the casting starts lagging immediately" (P2, semi-structured interview). Reduced connection speed then weakened casting performance and disrupted the flow of the lesson.

Headset availability also shaped classroom participation. The pedagogical issue was not simply scarcity, but how the teacher organized students around a visible VR experience. One participant explained that "rotation became the real lesson plan" when more students needed to try the activity (P7, semi-structured interview). This required assigning meaningful roles to students outside the headset rather than leaving them as passive observers. Otherwise, the activity could easily become "one person using VR and everyone else waiting and getting restless" (P7, semi-structured interview).

Projection quality also affected implementation. One participant noted that "an old classroom projector decreased the level of immersion for those who are watching" (P1, semi-structured interview). When the headset view was cast to an older projector, problems with image quality, color accuracy, resolution, and screen placement made it harder for observing students to follow the virtual scene or notice interactional details. In classrooms with smart screens or more stable display systems, the activity was easier to manage because the connection was more direct and the visual output was clearer. Thus, the same AI-enhanced VR task could become more or less pedagogically usable depending on the classroom's display infrastructure.

In short, this theme shows that AI-enhanced VR integration requires both instructional and infrastructural orchestration. Teachers need to design the language task, prepare observer roles, manage the headset user, monitor the casted view, and maintain the participation of students outside the headset. At the same time, they must navigate Wi-Fi authentication, mobile hotspot limitations, classroom location, projection quality, battery life, and device readiness. The findings therefore suggest that VR implementation should not be evaluated only by the number of available headsets. A pedagogically viable AI-enhanced VR lesson depends on whether the teacher can turn a technically fragile, single-headset or limited-headset setup into a visible, shared, and language-focused classroom experience.

Toward Pedagogically Accountable Artificial Intelligence–Virtual Reality Adoption

The fifth theme concerned future expectations and teacher responsibility. Instructors were optimistic that AI-enhanced VR would become more sophisticated, accessible, and realistic. They imagined more natural conversations, more personalized feedback, better virtual characters, easier screen sharing, and smoother integration into classroom routines. At the same time, they were cautious about over-reliance on AI, especially when feedback appeared immediate and confident. One instructor explained that AI feedback was useful because "it responds immediately" but still needed to be checked against the student's level and the purpose of the lesson (P3, semi-structured interview).

> I would not trust the AI feedback completely by itself. It is useful because it responds immediately, but I still need to check whether the correction is right for the student's level and for the purpose of the lesson. (P3, semi-structured interview)

This quotation positions AI feedback as useful but not authoritative. The instructor valued the immediacy of automated responses, yet still saw teacher judgment as necessary for interpreting accuracy, relevance, and appropriateness. Teachers remained necessary for validating feedback, explaining pragmatic appropriateness, connecting VR activity to curriculum goals, and deciding when technology was genuinely serving language learning.

This theme also captured a broader concern about technology-led pedagogy. Instructors did not want VR to become a classroom performance in which the tool itself was the center of attention. They valued AI-enhanced VR when it helped students speak more, rehearse situated language, receive immediate support, and participate with less anxiety. They were less convinced by uses in which the activity was visually impressive but weakly connected to course outcomes.

Stimulated-recall data further showed that teacher responsibility did not disappear during AI-mediated interaction. While revisiting the final part of a session, one participant stated that the teacher's role was "not smaller" but "different" because the AI could keep the conversation going while the teacher had to decide what to do with that conversation afterward (P6, stimulated recall).

> When I watched the last part of the session, I thought my role was not smaller; it was different. The AI kept the conversation going, but I had to decide what to do with that conversation afterwards. (P6, stimulated recall)

Taken together, the findings show that instructors did not view AI-enhanced VR as a self-evidently superior technology. They saw it as a powerful but condition-dependent environment. Its promise lay in responsive interaction, low-pressure practice, classroom-based realism, and future development. Its fragility lay in infrastructure, access, monitoring, and the risk of confusing technological excitement with pedagogical value. Pedagogically accountable adoption therefore

required a clear division of labor: AI could support interaction and feedback, VR could provide situated rehearsal, and the teacher could ensure alignment, interpretation, ethics, and learning continuity.

Discussion

The findings suggest that AI-enhanced VR may represent a meaningful shift in EFL teaching because it combines immersive simulation with responsive conversational interaction. Previous research has emphasized VR's ability to create situated, engaging, and lower-pressure environments for language practice (Hua & Wang, 2023; Kaplan-Rakowski & Gruber, 2023; Parmaxi, 2023). The present study extends this discussion by showing that instructors were especially interested in what AI adds to VR: the possibility of contingent, back-and-forth communication with a computer-generated interlocutor. This finding is also consistent with recent work on conversational agents and AI-mediated language education, which suggests that AI can expand opportunities for adaptive interaction, immediate response, and repeated language practice (Godwin-Jones, 2025; Hemminki-Reijonen et al., 2025; Xiao et al., 2023). Here, however, this value was not understood as a property of the tool alone. Instructors interpreted AI-enhanced VR as pedagogically meaningful when responsive interaction was connected to language objectives, classroom sequencing, and post-task consolidation.

This interpretation should also be understood in relation to the interconnected nature of AI-enhanced VR. Instructors did not experience VR, conversational AI, task design, classroom infrastructure, and teacher mediation as separate variables; rather, they described them as interdependent elements of the same classroom activity. This aligns with an orchestration view of technology-rich classrooms, where pedagogical value depends on how teachers coordinate tools, tasks, constraints, and learner participation in real time (Dillenbourg, 2013). Thus, the perceived pedagogical affordances reported here are best understood as emerging from the combined arrangement of immersive simulation, responsive AI interaction, task alignment, and teacher orchestration, rather than from any single component in isolation.

The theme of lowered social cost aligns with research on foreign-language classroom anxiety and willingness to communicate, as well as with VR research suggesting that simulated environments may reduce performance pressure during language practice (Horwitz et al., 1986; Kaplan-Rakowski & Gruber, 2023; MacIntyre et al., 1998). However, the present findings emerging from perception-based data show that reduced anxiety was not the only perceived benefit. Instructors connected AI-enhanced VR to attendance, attention, curiosity, and willingness to participate. This indicates that AI-enhanced VR may operate simultaneously as an affective scaffold, a motivational tool, and a communicative rehearsal environment. Yet, instructors were careful not to equate engagement with learning. This point is important for EFL contexts where learners may have limited opportunities for authentic English interaction outside the classroom: the simulated environment may increase willingness to speak and behaviorally engage, but teacher mediation remains necessary to transform participation into language-focused outcomes.

The findings also complicate simple affordance-based accounts of educational technology. Instructors did not reject VR because of technical constraints, but they repeatedly showed that infrastructure reorganized pedagogy. Institutional Wi-Fi authentication, mobile hotspot instability, casting reliability, projection quality, classroom location, battery life, and headset orchestration shaped whether the VR experience became visible, monitorable, and usable as a shared classroom activity. In this sense, classroom visibility was not merely a technical matter. It determined who could observe the interaction, how the teacher could monitor the learner inside the headset, how feedback could be shared, and whether the activity remained a collective language-learning event rather than an isolated individual experience. These results are consistent with broader work on technology integration, which shows that successful implementation depends on teacher knowledge, contextual fit, institutional culture, and pedagogical design rather than on the tool alone (Ertmer & Ottenbreit-Leftwich, 2010; Mishra & Koehler, 2006; Tseng et al., 2022). They also extend earlier work on 3D virtual learning environments in Turkish foreign language teacher education by showing that the current challenge is no longer only familiarizing teachers with virtual worlds, but enabling them to orchestrate technically constrained and visually fragile classrooms (Can & Şimşek, 2015).

A key implication concerns teacher education. If AI-enhanced VR is introduced only as a technical innovation, instructors may learn how to operate devices but not how to design meaningful language-learning sequences around them. This finding resonates with Can and Mangır's (2024) discussion of AI integration in language classrooms, where the perceived usefulness of AI is tied to instructor preparation, practical implementation support, and critical pedagogical guidance. In the present study, this issue is intensified because AI is embedded within a VR environment that requires teachers to manage hardware, interaction, feedback, visibility, classroom flow, and post-activity consolidation at the same time. Professional development should therefore include task alignment, pre-task preparation, in-VR interaction goals, observation roles for non-headset students, post-VR consolidation, feedback verification, and fallback planning. This is particularly important in single-headset or limited-headset classrooms, where the aim should not necessarily be to place multiple learners inside VR synchronously. A limited-headset model can be pedagogically productive when screen visibility, observer roles, peer feedback, role assignment, and whole-class reflection are built into the design.

The findings further suggest that AI feedback should be treated as pedagogically useful but not authoritative. Artificial intelligence–supported agents can increase the volume of practice and provide immediate responses, yet instructors remain responsible for evaluating accuracy, appropriateness, and alignment with course goals. This reinforces the argument that AI-enhanced VR should not be framed as a replacement for teachers. Instead, it should be understood as a mediated environment in which teacher judgment becomes more, not less, important. The teacher's expertise shifts toward orchestration, validation, and critical integration.

The qualitative descriptive design of the study also shapes how these findings should be interpreted. Rather than claiming generalizable effectiveness or measuring learning gains (Sandelowski, 2000), the study provides a practice-oriented account of how instructors make sense of AI-enhanced VR under real classroom conditions. This is important because the value of emerging educational technologies is often discussed through affordances, outcomes, or future potential, while less attention is given to teachers' situated decisions during implementation. The present findings therefore contribute by showing how instructors evaluate AI-enhanced VR through the practical questions of visibility, access, task alignment, feedback reliability, and classroom manageability.

Finally, the study positions AI-enhanced VR as part of a transitional moment in language education. Future classrooms may integrate VR, AI agents, mixed reality, and contextual overlays in increasingly seamless ways. However, the current findings indicate that the immediate challenge is more practical and pedagogical: making AI-enhanced VR usable, accountable, and meaningfully aligned with EFL learning objectives under ordinary institutional conditions.

Conclusion and Recommendations

This study explored how EFL instructors perceived and implemented AI-enhanced VR activities, the challenges they

encountered, and the future possibilities they associated with this emerging technology. The findings indicate that instructors valued AI-enhanced VR because they perceived it as enabling learners to engage in responsive conversations with virtual agents, increasing engagement and attendance, lowering speaking pressure, and creating opportunities for realistic communicative rehearsal inside the classroom. However, the study also showed that implementation was constrained by infrastructure, headset access, screen-casting, battery life, classroom monitoring, and the need to validate AI-generated feedback.

The central conclusion is that AI-enhanced VR should not be treated as a self-contained solution for EFL teaching. Its pedagogical value depends on careful instructional design, institutional support, and sustained teacher professional development. Instructors need support in operating VR hardware, designing language-focused tasks, managing visibility-oriented single-headset or limited-headset classrooms, integrating AI feedback, and consolidating learning after immersive activities.

Teacher education and professional development programs should prepare instructors to evaluate AI-enhanced VR critically, design communicative tasks around clear language objectives, manage visibility-oriented single-headset or limited-headset classrooms, verify AI-generated feedback, and connect immersive activities to broader curricular goals. Institutions considering AI-enhanced VR adoption should also provide reliable internet access, clear Wi-Fi authentication guidelines for VR devices, adequate headset maintenance, stable screen-casting infrastructure, smart-screen or high-quality projection systems, technical support, and realistic time allocation for setup, troubleshooting, and post-activity consolidation.

Limitations and Future Research

This study has a context-specific sample and relies primarily on instructor accounts rather than longitudinal classroom observation, behavior change, or student learning outcomes. The findings are highly related to the context of participants. The findings should therefore be interpreted as a qualitative account of instructor experience rather than as evidence of generalizable effectiveness. In addition, the study design cannot determine whether the perceived benefits reported by instructors resulted from the VR setting itself, the AI-powered conversational agent operating within it, or the combination of both. This can be a point of exploration in future research. Studies in the future should examine student perspectives, compare different AI-enhanced VR task designs, investigate learning outcomes over time, and explore how teachers develop expertise in orchestrating AI-mediated immersive environments. As mixed-reality and object-aware AI systems mature, future studies may also examine how VR, AR, and AI converge in classroom practice without assuming that all immersive technologies are equally available or equally pedagogically developed at the present moment.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author.

Artificial Intelligence Usage Statement: This study used large language model tools to support preliminary open-code suggestions during qualitative data analysis. These tools were used only to identify recurring ideas, patterns, and possible thematic clusters in anonymized textual data. All coding decisions, category development, theme construction, interpretation, and reporting were reviewed, revised, and finalized by the authors. The authors take full responsibility for the content of the manuscript.

Ethics Committee Approval: Ethical committee approval was received from the Ethics Committee of Istanbul Aydın University (Approval No.: E-45379966-100-162512; Date: August 1, 2025).

Informed Consent: Written informed consent was obtained from all participating EFL instructors.

Peer-review: Externally peer-reviewed.

Acknowledgments: The authors thank the participating English as a foreign language instructors for sharing their experiences and reflections.

Author Contributions: Concept – C.G.K., T.C.; Design – C.G.K., T.C.; Supervision – T.C.; Resources – C.G.K., T.C.; Materials – C.G.K.; Data Collection and/or Processing – C.G.K.; Analysis and/or Interpretation – C.G.K., T.C.; Literature Search – C.G.K., T.C.; Writing Manuscript – C.G.K.; Critical Review – T.C.

Declaration of Interests: The authors declare that they have no competing interests.

Funding: The authors declare that this study received no financial support.

References

Can, T., & Mangır, C. (2024). Empowering language educators: Integrating artificial intelligence in classrooms and beyond. In Y. Özkan, H. Tuncer, & E. Bada (Eds.), *An approach to digitalization in language teaching from different perspectives: A reflection of empirical work* (pp. 45–60). Nova Science Publishers. [CrossRef]

Can, T., & Şimşek, I. (2015). The use of 3D virtual learning environments in training foreign language pre-service teachers. *Turkish Online Journal of Distance Education*, *16*(4), 114–124. [CrossRef]

Chen, J., Dai, J., Zhu, K., & Xu, L. (2022a). Effects of extended reality on language learning: A meta-analysis. *Frontiers in Psychology*, *13*, Article 1016519. [CrossRef]

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022b). Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational Technology and Society*, *25*(1), 28–47. [CrossRef]

Dikici, A., & Altaş, B. (2024). Relationship between English as a foreign language learners' use of reading strategies and their perceptions of critical reading self-efficacy. *HAYEF: Journal of Education*, *21*(3), 279–286. [CrossRef]

Dillenbourg, P. (2013). Design for classroom orchestration. *Computers and Education*, *69*, 485–492. [CrossRef]

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115. [CrossRef]

Ertmer, P. A., & Ottenbreit-Leftwich, A. T. (2010). Teacher technology change: How knowledge, confidence, beliefs, and culture intersect. *Journal of Research on Technology in Education*, *42*(3), 255–284. [CrossRef]

Godwin-Jones, R. (2025). AI and VR converge: The future of language learning in an emerging metaverse. In Y.-J. Lan, G. Y. Qi, & D. Chun (Eds.), *AI-mediated language education in the metaverse era* (pp. 221–246). Springer. [CrossRef]

Hemminki-Reijonen, U., Hassan, N. M. A. M., Huotilainen, M., Koivisto, J.-M., & Cowley, B. U. (2025). Design of generative AI-powered pedagogy for virtual reality environments in higher education. *npj Science of Learning*, *10*(1), Article 31. [CrossRef]

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal*, *70*(2), 125–132. [CrossRef]

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288. [CrossRef]

Hua, C., & Wang, J. (2023). Virtual reality-assisted language learning: A follow-up review (2018–2022). *Frontiers in Psychology*, *14*, Article 1153642. [CrossRef]

Kaplan-Rakowski, R., & Gruber, A. (2023). The impact of high-immersion virtual reality on foreign language anxiety. *Smart Learning Environments*, *10*(1), Article 46. [CrossRef]

Kıyançiçek, E., & Çoban, M. (2024). An investigation of the factors affecting teacher and student motivation in online education. *HAYEF: Journal of Education*, *21*(1), 57–65. [CrossRef]

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *Modern Language Journal*, *82*(4), 545–562. [CrossRef]

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research*, *26*(13), 1753–1760. [CrossRef]

Marougkas, A., Troussas, C., Krouska, A., & Sgouropoulou, C. (2023). Virtual reality in education: A review of learning theories, approaches and methodologies for the last decade. *Electronics*, *12*(13), Article 2832. [CrossRef]

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record: The Voice of Scholarship in Education*, *108*(6), 1017–1054. [CrossRef]

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16*(1), 1–13. [CrossRef]

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health*, *42*(5), 533–544. [CrossRef]

Parmaxi, A. (2023). Virtual reality in language learning: A systematic review and implications for research and practice. *Interactive Learning Environments*, *31*(1), 172–184. [CrossRef]

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed). SAGE.

Sandelowski, M. (2000). Whatever happened to qualitative description? *Research in Nursing and Health*, *23*(4), 334–340. [CrossRef]

Sarı, Y. E. (2022). Authentic language learning in metaverse: A big step in the technological revolution or a buzzword? *HAYEF*, *19*(3), 255–259. [CrossRef]

Stahl, N. A., & King, J. R. (2020). Expanding approaches for research: Understanding and using trustworthiness in qualitative research. *Journal of Developmental Education*, *44*(1), 26–28.

Tai, R. H., Bentley, L. R., Xia, X., Sitt, J. M., Fankhauser, S. C., Chicas-Mosier, A. M., & Monteith, B. G. (2024). An examination of the use of large language models to aid analysis of textual data. *International Journal of Qualitative Methods*, *23*, Article 16094069241231168. [CrossRef]

Tai, T.-Y., Chen, H. H.-J., & Todd, G. (2022). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Computer Assisted Language Learning*, *35*(4), 892–917. [CrossRef]

Than, N., Fan, L., Law, T., Nelson, L. K., & McCall, L. (2025). Updating "The Future of Coding": Qualitative coding with generative large language models. *Sociological Methods and Research*, *54*(3), 849–888. [CrossRef]

Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative Inquiry*, *16*(10), 837–851. [CrossRef]

Tseng, J.-J., Chai, C. S., Tan, L., & Park, M. (2022). A critical review of research on technological pedagogical and content knowledge (TPACK) in language teaching. *Computer Assisted Language Learning*, *35*(4), 948–971. [CrossRef]

Xiao, F., Zhao, P., Sha, H., Yang, D., & Warschauer, M. (2023). Conversational agents in language learning. *Journal of China Computer-Assisted Language Learning*, *4*(2), 300–325. [CrossRef]

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), Article 39. [CrossRef]

Interview and Stimulated-Recall Protocols

Note on analytic openness. These protocols were used as flexible data-collection guides, not as a coding framework. They were designed to elicit instructors' accounts in their own terms. The interview and stimulated-recall prompts did not assume that AI-enhanced VR was useful, effective, motivating, problematic, realistic, or successful. Themes were generated later through inductive qualitative content analysis.

Supplementary Material 1. Semi-Structured Interview Protocol

Purpose and use.

This protocol guided individual semi-structured interviews with EFL instructors who had prior experience using, observing, or demonstrating VR-based, VR-compatible, or AI-supported language-learning activities. The aim was to understand instructors' experiences, implementation decisions, and reflections. The interviewer could change the order of prompts, omit questions already answered, and use follow-up probes where needed.

Estimated duration: approximately 30–45 minutes. Mode: face-to-face or online individual interview.

Interviewer guidance.

Use neutral, open-ended wording. Do not suggest that the activity was beneficial, innovative, motivating, difficult, or unsuccessful. Avoid evaluative reactions such as "That sounds effective" or "That must have been a problem." Use neutral probes such as "Could you tell me more about that?", "What made that moment stand out?", "Could you give an example?", and "Were there also limitations or alternative interpretations?"

Before-session checklist.

Before the interview, the interviewer confirmed study information, informed consent, permission for audio recording, voluntary participation, the right to skip questions or stop, confidentiality and anonymization procedures, and the request not to name students, colleagues, institutions, or identifiable individuals. The interviewer also checked the recording tool and noted the date, mode, participant code, and contextual information.

Opening script.

Thank you for agreeing to take part in this interview. The purpose of this study is to understand how EFL instructors experience and interpret AI-enhanced VR activities in language teaching. I am interested in your own experience, including what worked, what did not work, what was uncertain, and how you made decisions during implementation. This is not an evaluation of your teaching. There are no right or wrong answers. You may skip any question or stop at any time. Please avoid using identifiable names or institutional details. With your permission, I will audio-record the interview for accurate transcription.

Interview domains and prompts.

- Teaching context: Could you briefly describe your teaching context and your experience as an EFL instructor? Optional probes: course type, student group, language level, usual speaking or role-play activities.
- Prior experience: Could you describe your previous experience with VR-based, VR-compatible, or AI-supported language-learning activities? Optional probes: first encounter, classroom use, demonstration, training, observation, platform, headset, application, or classroom setup.
- Activity description: Please describe one AI-enhanced VR or VR-compatible AI-supported activity that you used, observed, or helped implement. Optional probes: activity goal, student task, VR role, AI-supported interlocutor or feedback role, grouping, lesson sequence, pre-task and post-task steps.
- Instructional intention: What were you trying to achieve instructionally through this activity? Optional probes: language focus, speaking/listening/vocabulary/pronunciation/pragmatic goal, course objective, and relevance for students.
- AI and VR roles: How did you interpret the role of AI and the role of VR in the activity? Optional probes: what the AI did, how students responded, reliability of AI responses or feedback, what VR added or did not add, and whether the task felt like a simulation, conversation, game, practice task, or something else.
- Participation and interaction: What did you notice about students' participation, speaking, communication, or behavior during the activity? Optional probes: active or hesitant participation, students outside the headset, turn-taking, waiting time, language produced, hesitation, repair, clarification, continuation, comfort, confusion, distraction, or overexcitement.
- Implementation conditions: What practical or technical conditions shaped the activity? Optional probes: Wi-Fi, authentication, mobile hotspot, headset connection, laptop connection, casting, audio, projection, display, battery, device readiness, classroom layout, device access, and monitoring inside the VR environment.
- Feedback and follow-up: How was feedback handled during or after the activity? Optional probes: AI feedback, teacher feedback, peer feedback, how feedback was checked, error treatment, post-task discussion, vocabulary or expression review, reflection, and follow-up tasks.
- Evaluation and future use: Looking back, how would you evaluate this activity as part of EFL teaching? Optional probes: useful, limited, uncertain, or unexpected aspects; what to keep, remove, or redesign; needed training, technical support, institutional guidelines, infrastructure, and future research needs.

Closing.

Is there anything about your experience with AI-enhanced VR in EFL teaching that we have not discussed but you think is important? The interviewer then thanked the participant, reminded them of anonymization procedures, and noted possible points for member checking.

Post-interview field notes. Participant code; date; mode; duration; technology or platform mentioned; type of activity described; language-learning focus; implementation conditions; key issues raised; possible follow-up points; researcher reflections; assumptions or biases to monitor; deviations from the protocol.

Supplementary Material 2. Stimulated-Recall Protocol

Purpose and use.

This protocol guided stimulated-recall sessions with EFL instructors who had participated in the main interview and had access to a short lesson-recording excerpt or demonstration excerpt. The aim was to elicit situated reflections on teacher decision-making, classroom monitoring, learner interaction, technical conditions, and post-task interpretation while revisiting a specific moment. The excerpt was used as a reflection prompt, not as evidence for evaluating teaching performance.

Estimated duration: approximately 12–20 minutes. Stimulus material: one short classroom excerpt, teacher-led demonstration, screen-cast, projected VR view, or recording showing monitoring, student interaction, or post-task discussion. Identifying details were removed or avoided where possible.

Researcher guidance.

Invite the participant to describe and interpret the excerpt in their own terms. Avoid suggesting that the moment was successful,

unsuccessful, motivating, realistic, innovative, or problematic. The participant may pause, replay, or comment on any moment. Use neutral prompts such as “What do you notice at this moment?”, “How did you interpret what was happening?”, “What were you monitoring?”, and “What would you do differently, if anything?”

Before-session checklist.

Before the session, the researcher confirmed prior consent, permission to use the selected excerpt for stimulated recall, voluntary participation, the right to skip questions or stop, anonymization procedures, and the request not to name identifiable people or institutions. The researcher checked that playback, audio, video, screen sharing, and recording tools were working, and noted the participant code, date, excerpt type, excerpt duration, and session mode.

Opening script.

Thank you for taking part in this stimulated-recall session. We will look at a short excerpt from an AI-enhanced VR or VR-compatible language-learning activity. The purpose is not to evaluate your teaching or the students’ performance. The aim is to understand what you noticed, what you were trying to do, what decisions you made, and how you interpret this moment now. You may pause or replay the excerpt at any point. There are no right or wrong answers. Please avoid identifiable names or institutional details. With your permission, I will audio-record our discussion for accurate transcription.

Viewing procedure.

The researcher briefly identified the excerpt without interpreting it. The participant watched the excerpt once, unless they chose to pause. The excerpt could then be replayed, paused, or discussed in segments. The researcher used neutral prompts to explore teacher intention, learner interaction, AI/VR role, classroom conditions, monitoring, and follow-up. The participant was invited to identify moments that were clear, unclear, expected, unexpected, easy, or difficult to interpret.

Recall areas and prompts.

- Initial description: Please watch the excerpt and describe what you notice happening here. Optional probes: what is happening, who is involved, what the learner is doing, what the teacher is doing, what is happening outside the headset, and what stands out now.
- Instructional purpose: How would you describe the purpose of this moment in the lesson or demonstration? Optional probes: language-learning focus, communication, fluency, vocabulary, pronunciation, pragmatic or task-completion goal, and connection to the larger activity sequence.
- Teacher decision-making: What were you trying to do as the teacher or facilitator at this point? Optional probes: monitoring, waiting, deciding whether to intervene, focusing on language, technology, classroom flow, student response, or another issue.
- Learner interaction: What do you notice about the learner’s interaction and language use? Optional probes: response type, hesitation, continuation, repetition, repair, clarification, language used, expected or unexpected responses, and how participation was interpreted at the time.
- AI and VR components: What role did the VR environment and AI-supported component play in this excerpt? Optional probes: virtual setting, visual or spatial features, AI response, prompt, correction, suggestion, feedback, quality of response, and moments requiring teacher judgment.
- Visibility and monitoring: How were you able to monitor what was happening inside the VR environment? Optional probes: headset view, audio, casting, projector, smart display, delay, interruption, loss of visibility, and how students outside the headset followed the activity.
- Technical and institutional conditions: What technical or infrastructural factors affected this moment, if any? Optional probes: internet connection, institutional Wi-Fi, login, authentication, hotspot use, connection stability, casting, projection, smart screen, audio, battery, headset readiness, classroom location, and layout.
- Response to breakdowns: If something did not work as expected, how did you respond? Optional probes: pausing, restarting, switching tasks, putting on the headset, changing grouping, continuing, student reactions, and effects on the language-learning focus.
- Follow-up and interpretation: What happened after this moment, and how would you interpret it now? Optional probes: class discussion, language review, AI feedback, student reflection, follow-up task, teacher role, student language use, classroom conditions, and whether the moment is interpreted differently now.
- Redesign: If you repeated this activity or a similar activity, what would you keep, change, or prepare differently? Optional probes: task design, timing, sequencing, student roles, headset management, casting, display, AI feedback, preparation, and institutional support.

Closing.

Is there anything else about this excerpt that is important for understanding AI-enhanced VR in EFL teaching? The researcher then thanked the participant, reminded them of anonymization procedures, and noted possible points for member checking.

Post-session field notes.

Participant code; date; mode; duration; excerpt type and length; activity context; language-learning focus; AI-supported component; VR or display setup; technical conditions; teacher decisions; learner actions; students outside the headset; follow-up; notable reflections; researcher observations; assumptions or biases to monitor; deviations from the protocol.