

Rafael Seiz Ortiz

Universitat Politècnica de València, Spain
rseiz@idm.upv.es

THEORETICAL FRAMEWORK FOR THE PEDAGOGICAL USE OF EXTENDED REALITY IN UNIVERSITY-LEVEL ENGLISH FOR SPECIFIC PURPOSES TEACHING

Abstract

Extended reality (XR) has increasingly been adopted to support complex, task-based learning environments that foster the development of domain-specific competencies, including those required in English for Specific Purposes (ESP). Even though XR has been advocated as an adequate environment to improve language acquisition, its integration within the language teaching process mostly remains limited and exploratory. Therefore, there is a need to investigate which language learning components can be most effectively enhanced by XR technologies, and how this enhancement can take place. The objective of this study is to develop a theoretical framework to meet that aim, thus helping ESP teachers and learners to integrate XR-based technology within the language learning process at the university level. To address this research objective, the following steps have been undertaken. First, a literature review has been conducted both to compile relevant background knowledge on XR in language learning and establish pedagogical criteria. The methodology used is bibliometric analysis. Second, a semi-structured questionnaire has been administered to 22 university ESP teachers to identify teaching needs addressed by XR technologies. A grounded theory methodology has been used. Based on these steps, a theoretical framework has been developed to guide the implementation of XR technologies in ESP.

369

Key words

extended reality (XR), immersive learning, English for specific purposes (ESP), computer assisted language learning (CALL), higher education.

1. INTRODUCTION

In recent years, the rapid development of immersive digital technologies has led to a growing interest in their potential to transform many areas of human life, including language teaching and learning, as well as English for Specific Purposes (ESP) in higher education. Among these technologies, extended reality (XR) – an umbrella term that refers to a range of technologies that blend, extend, or replace the physical world with digital content, including virtual reality (VR), augmented reality (AR), and mixed reality (MR) – has emerged as a promising area in language education due to its capacity to integrate digital content with physical and social learning environments.

XR does not represent a single technology; instead, it encompasses various modalities that differ in their level of immersion, interaction, and relationship with the physical environment. These technologies are usually conceptualised along a reality–virtuality continuum that ranges from fully physical environments to fully virtual ones (Milgram et al., 1995). At one end is VR, which deploys a fully immersive, computer-generated environment in which users are visually and perceptually separated from the physical world. Interaction occurs entirely within a simulated space, typically accessed through head-mounted displays and motion-tracking devices (Bordegoni et al., 2023). VR has a high degree of immersion and is widely used for simulations, immersive training, virtual laboratories, and experiential learning environments. On the other hand, AR superimposes digital content, such as text, images or 3D objects, onto the user’s real-world environment in real time. The physical world remains primary, while virtual components enhance perception or provide contextual information (Milgram et al., 1995; Santos et al., 2016). AR-based experiences are commonly delivered through smartphones, tablets, or wearable devices (Bordegoni et al., 2023). In between VR and AR is MR, which enables digital and physical elements to coexist and interact dynamically. In MR environments, virtual objects, spatially anchored to the real world, can respond to physical objects and user actions in real time (Wienrich et al., 2021). This XR modality requires advanced sensing and spatial-mapping technologies and is particularly appropriate for complex training contexts in which real–virtual interaction is crucial. Thus, XR technologies are best understood not as discrete categories but as part of a continuum defined by the degree of immersion, the extent of real–virtual integration and the nature of user interaction (Skarbez et al., 2021). This conceptualisation highlights the flexibility of XR to serve as a framework to design experiences ranging from reality-enhancing to full immersion, depending on pedagogical or professional goals.

XR technologies enable learners to interact with language in situated, experiential, and multimodal contexts, offering new possibilities for addressing major challenges in second language acquisition (SLA), such as limited exposure to authentic communicative situations (Xu, 2010), reduced learner engagement, and difficulties in transferring language knowledge to real-world use (Ellis et al., 1994;

Schorr et al., 2024). Beyond interaction, another key concept commonly associated with XR technologies – and especially relevant to language learning – is immersion. This concept is often discussed in connection with communicative language teaching as an ideal way to put learners in contact with authentic language (Yan & Lowell, 2025; Yan et al., 2025; Zhi & Wu, 2023).

From a pedagogical point of view, XR aligns with communicative and experiential approaches to language learning (Lou, 2025). VR provides learners with fully immersive environments in which they can engage in simulated communicative events, thereby improving oral fluency, pragmatic competence, and confidence in language use, within a contextualised educational setting that mimics the real world (Akay & Kessler, 2024; Taguchi, 2022; Xie et al., 2021). AR enhances real-world contexts with situated linguistic support, enabling learners to access vocabulary, grammatical cues, or multimodal annotations in situ (Alhawsawi & Alzaid, 2025; Draxler et al., 2020; Santos et al., 2016). MR further extends these affordances by allowing digital objects to interact dynamically with the physical environment, supporting collaborative tasks and professional simulations (Filippone et al., 2025; Lyu & Li, 2022). These affordances are particularly important in ESP; consequently, XR is especially relevant in this area of language teaching and learning (Christou et al., 2025; Falcone, 2025). In linking language use to action and spatial experience, the educational value of XR-based environments lies not merely in their novelty, but in their ability to operationalise core principles of SLA (de Paula Forero Pataquiva & Klimova, 2022) and communicative language learning, such as meaningful interaction (Mackey, 2007), rich and comprehensible input (Ellis et al., 1994), noticing through multimodal feedback (Bassiri, 2011), and embodied cognition (Atkinson, 2011; Kosmas & Zaphiris, 2020). Additionally, XR technologies enable the design of low-risk and highly authentic learning scenarios through which learners can rehearse communicative practices that would otherwise be difficult to simulate in more traditional instructional settings (Kaplan-Rakowski & Gruber, 2023; Thrasher, 2022). These affordances are especially relevant in higher education and ESP, where the integration of language learning with disciplinary practice and real-world problem-solving is essential (Filippone et al., 2025; Lyu & Li, 2022).

Despite the growing body of research on XR in education, studies about its implementation in language learning are often fragmented across technologies, contexts, and pedagogical approaches (Yudintseva, 2023). Therefore, there is a need for a coherent synthesis that clarifies the pedagogical potential of XR modalities to contribute to language development. The overall aim of this article is to address this gap by analysing XR technologies from a language learning pedagogical perspective grounded in communicative competence, ESP, and task-based learning as theorised in established Task-Based Language Teaching (TBLT) frameworks (Ellis, 2003; Long, 2015; Nunan, 2004) and the Common European Framework of Reference (CEFR) action-oriented approach (Council of Europe, 2018). Specifically, we adopt Long's (2015) distinction between target tasks (real-world tasks) and pedagogic

tasks (instructional tasks), Ellis's (2003) emphasis on pragmatic language processing for meaningful outcomes, Nunan's (2004) focus on learner-driven meaning construction, and the CEFR's principle of planning backwards from real-life communicative needs. This approach aims to establish a pedagogically grounded framework for the use of XR technologies in language learning, thereby contributing to the theoretical consolidation of the field and guiding the pedagogical design of effective, research-based XR-mediated language learning environments.

In line with its objective, this study is structured around a sequence of steps. First, a literature review is conducted with the twofold aim of (1) synthesising relevant background knowledge on the effective implementation of XR in language learning, and (2) establishing pedagogical parameters and components to lay the foundations of the theoretical framework. Then, a short questionnaire is administered to 22 ESP teachers, with the aim of gathering their viewpoints and identifying strategic teaching needs that can be addressed through XR technologies in ESP. The third phase builds upon the previous steps to develop a pedagogically driven theoretical framework for implementing XR technologies in the ESP teaching and learning process. This study aims to enhance researchers' understanding of XR technologies within the context of ESP and, more practically, to assist ESP teachers in the implementation of XR-based technologies.

2. PREVIOUS RESEARCH ON THE USE OF XR TECHNOLOGIES IN LANGUAGE LEARNING

2.1. Need for pedagogical frameworks and modality alignment

Several review articles (Huang et al., 2021; Luo et al., 2024; Özçelik et al., 2022; Parmaxi & Demetriou, 2020; Pegrum & Lan, 2023; Peixoto et al., 2021; Tafazoli, 2024) identify a key tension: empirical studies often report positive outcomes, yet the field remains methodologically inconsistent and theoretically fragmented, making it difficult to determine *when* XR is effective, *for whom*, and *under what pedagogical conditions*. This mismatch motivates the present literature review, which aims to consolidate knowledge on the efficient use of XR in language learning and to extract certain pedagogical criteria as a basis for a theoretical framework. Prior reviews (Luo et al., 2024; Pegrum & Lan, 2023; Tafazoli, 2024) emphasise that XR should not be treated as a uniformly "superior" medium; rather, its effectiveness depends on the alignment between XR modality, learning objectives, task design, learner characteristics, teacher mediation and implementation constraints.

The differentiation of XR technologies is grounded in pedagogy rather than purely technical distinctions: AR is better suited to situated vocabulary learning, contextual noticing, place-based tasks, or tangible interaction, whereas VR is better suited to interaction-rich communication, experiential scenarios, and embodied

rehearsal of language use (Luo et al., 2024; Parmaxi & Demetriou, 2020; Pegrum & Lan, 2023; Peixoto et al., 2021). Therefore, a core criterion for effective XR integration is the alignment between technological modality and pedagogical purpose, rather than adoption driven by novelty or availability (Pegrum & Lan, 2023; Tafazoli, 2024).

2.2. XR across language domains: Evidence and pedagogical implications

With regard to the role of XR in the development of specific linguistic skills and language domains, research suggests that vocabulary learning, especially form-meaning mapping and retention, is the most consistent learning outcome associated with XR use (Huang et al., 2021; Luo et al., 2024; Özçelik et al., 2022; Parmaxi & Demetriou, 2020; Peixoto et al., 2021). According to some studies (Dhimolea et al., 2022; Peixoto et al., 2021), XR learners may not always outperform traditional approaches immediately, but they often show stronger delayed retention, particularly when tasks imply embodied interaction and contextual grounding. Classroom-level empirical work also provides evidence for this pattern in AR-supported vocabulary learning with younger learners, indicating that visual/contextual overlays can strengthen lexical encoding, engagement, and retention when activities are well structured (Karacan & Akoğlu, 2021; see also Özçelik et al., 2022 for broader review patterns). Similarly, immersive VR supports discipline-specific terminology in Content and Language Integrated Learning (CLIL)/ESP-type contexts by embedding vocabulary in rich conceptual environments (Parra Pérez et al., 2025). Thus, one evidence-based pedagogical criterion is that XR tends to be most efficient when the learning goal benefits from multimodal contextualisation and embodied memory cues.

Regarding speaking skills, research frequently reports that XR is associated with increased willingness to communicate, reduced anxiety, and opportunities for interaction (Huang et al., 2021; Luo et al., 2024; Panagiotidis, 2021; Tafazoli, 2024). However, the evidence is also consistently conditional: speaking gains depend on whether XR activities are structured as genuine tasks, that is, activities requiring meaningful communication rather than form-focused practice (Ellis, 2003; Long, 2015; Nunan, 2004). Speaking gains also depend on interaction design, feedback, scaffolding, and learner readiness (Song et al., 2023; Zhi & Wu, 2023).

Several studies also point out that the pedagogical advantage of VR is frequently experiential and affective, since learners feel less self-conscious and more willing to participate, but these gains must be converted into actual learning through goal-oriented tasks and structured interaction (Luo et al., 2024; Panagiotidis, 2021). Following Ellis's (2003) strong version of TBLT, tasks should serve not merely to activate existing knowledge but as the primary means through which learners experience how language is used in communication. In line with the

CEFR framework, this means designing collaborative tasks whose primary focus is on achieving outcomes beyond language practice itself, thereby channelling affective engagement into meaningful language use (Council of Europe, 2018). Likewise, studies examining VR-based learning activities that incorporate real-time feedback mechanisms and adaptive input suggest that immediate contextualised feedback can support noticing and automatization, provided it is pedagogically focused and cognitively manageable (Lou, 2025; Song et al., 2023).

Writing is still an emerging area, particularly descriptive or expository writing supported by immersive environments, although it remains less consistently represented than vocabulary and speaking (Huang et al., 2021; Tafazoli, 2024). AR and VR may provide rich stimuli and contextual prompts that support written production, but the efficiency of XR for developing writing depends on instructional sequencing, task constraints and integration with feedback and revision (Luo et al., 2024; Tafazoli, 2024). Overall, XR appears to be more consistently effective for vocabulary/retention and certain interaction-driven oral learning outcomes than for text-based skills.

Finally, evidence for grammar learning and extended reading comprehension is comparatively limited, suggesting that XR is not inherently advantageous for all language domains (Huang et al., 2021; Luo et al., 2024; Pegrum & Lan, 2023; Tafazoli, 2024). This means that an efficiency-oriented framework should avoid overgeneralised claims and advocate a selective-use principle based on task-skill-modality fit.

2.3. Core mechanisms: Interaction, embodiment and multimodality

Across the reviewed corpus, XR effectiveness is associated not with immersion alone but with a cluster of pedagogical mechanisms, including situated learning, interaction, embodiment, multimodality, and cognitive-affective mediation (Huang et al., 2021; Parmaxi & Demetriou, 2020; Pegrum & Lan, 2023; Peixoto et al., 2021; Song et al., 2023; Tafazoli, 2024; Zhi & Wu, 2023).

XR frequently enhances authenticity by enabling learners to rehearse language in environments resembling everyday, professional, and intercultural situations (Luo et al., 2024; Panagiotidis, 2021; Peixoto et al., 2021; Symonenko et al., 2020; Tegoan et al., 2021). AR can similarly support place-based language activities through contextual augmentation (Parmaxi & Demetriou, 2020; Wen, 2021). These affordances align with situated and experiential learning and suggest that XR is most effective when tasks require meaningful communication. As Long (2015) emphasises, authentic TBLT bridges this gap by deriving pedagogic tasks from the target tasks learners need to accomplish in real-world contexts. Ellis (2003) similarly distinguishes tasks from exercises by their requirement for pragmatic language processing to achieve communicative outcomes.

A consistent finding is that XR should prioritise interaction over passive exposure. Immersion may increase learners' attention and engagement, but it is interaction – across learner–environment, learner–peer, and learner–task dimensions – that enables meaning negotiation, output, feedback, and reflection (Luo et al., 2024; Panagiotidis, 2021; Song et al., 2023; Tafazoli, 2024). These are fundamental concepts in SLA. Thus, immersion functions as an enabling condition, while interaction drives learning. This aligns with the CEFR's action-oriented approach, which emphasises the co-construction of meaning through interaction (Council of Europe, 2018), and with TBLT principles, where tasks require meaningful language use and collaboration (Ellis, 2003; Long, 2015; Nunan, 2004).

Immersive VR is repeatedly linked to embodiment: learners experience the feeling of being located in a setting, interacting spatially through a body/avatar. Pedagogically, this can support deeper encoding and retention, especially for vocabulary and contextualised language use (Peixoto et al., 2021; Symonenko et al., 2020; Zhi & Wu, 2023). However, embodiment must be integrated with tasks that lead the embodied activity towards linguistic goals (Pegrum & Lan, 2023; Zhi & Wu, 2023). In TBLT terms, such activities must function as genuine tasks focused on meaning, real-world relevance, and outcome-based assessment (Nunan, 2004).

XR's multimodal nature – combining visual, auditory, and interactive stimuli – can enhance dual coding and a particularly rich form of form–meaning mapping (Huang et al., 2021; Özçelik et al., 2022; Peixoto et al., 2021), but it also introduces the challenge of distraction and cognitive overload (Song et al., 2023; Tafazoli, 2024; Wen, 2021). Efficient XR design therefore requires careful orchestration of modalities to support language-learning objectives.

2.4. Cognitive and affective mediation in XR learning

Cognitive load emerges as a key constraint in XR-based learning. While XR can enrich learning environments, it may also consume working memory resources needed for linguistic processing (Peixoto et al., 2021; Song et al., 2023; Tafazoli, 2024; Wen, 2021; Zhi & Wu, 2023). Song et al. (2023) highlight cognitive load theory as a design lens: external load should be minimised, intrinsic load carefully sequenced, and processing supported through attention guidance and structured tasks.

Similarly, the Interactive–Constructive–Active–Passive (ICAP) framework (Chi & Wylie, 2014) suggests that deeper learning depends on facilitating constructive and interactive engagement rather than overwhelming learners with unstructured XR stimulation (Wen, 2021). At a broader level, the Cognitive Affective Model of Immersive Learning (CAMIL), a framework explaining how immersive VR influences learning processes, explicitly considers cognitive load and self-regulation as mediators determining whether immersion leads to learning (Makransky & Petersen, 2021; Zhi & Wu, 2023). Collectively, these studies imply a key criterion:

efficient XR instruction is attention-designed, with clear constraints, cues, pacing, and debriefing. Following the task-based framework, instructional design should enable learners to give primary attention to meaning while processing language pragmatically, supported by scaffolding of complexity and clear specification of learning outcomes (Ellis, 2003; Long, 2015; Nunan, 2004).

From an affective and motivational standpoint, XR often increases motivation, enjoyment, confidence, and willingness to communicate, while reducing anxiety, especially in oral communication tasks (Huang et al., 2021; Luo et al., 2024; Panagiotidis, 2021; Symonenko et al., 2020; Tafazoli, 2024). However, affective benefits should not necessarily be equated with learning gains. Affect functions as a mediator increasing time-on-task and willingness to participate, which must be channelled through structured pedagogy to yield measurable learning outcomes (Huang et al., 2021; Luo et al., 2024; Peixoto et al., 2021; Tafazoli, 2024).

Recent research strengthens this mediator perspective. A study combining XR with AI and mobile learning highlights that motivation and engagement are closely linked and influenced by learners' technological readiness, suggesting that observed 'XR effects' may partly reflect 'readiness effects' (Zhang & Miao, 2025). Likewise, CLIL-oriented VR interventions emphasise self-efficacy as a key factor: learners' perceived control and competence in immersive tasks can increase sustained participation and improve learning efficiency (Parra Pérez et al., 2025). CAMIL similarly places interest, motivation, and self-efficacy as core mediators linking affordances (presence/agency) to learning outcomes (Zhi & Wu, 2023).

Therefore, a central pedagogical criterion is that efficient XR design must explicitly address motivational architecture, not by maximising novelty or excitement, but by aligning engagement with goal-oriented tasks, feedback, and reflection.

2.5. Agency, feedback, and adaptive XR environments

Beyond engagement, several studies argue that XR's highest pedagogical value arises when learners are not only immersed but also agentic, i.e., when they make choices, control the pace and shape learning objects (Parra Pérez et al., 2025; Pegrum & Lan, 2023; Wen, 2021; Zhi & Wu, 2023). Wen (2021) shows that learner-generated AR tasks can sustain deeper engagement over time than instructor-designed ones, especially when complemented with collaboration and explanation. Pegrum and Lan (2023) similarly highlight XR content creation as an underexplored but promising pathway to foster autonomy, creativity, and deeper language processing. This approach aligns with the concept of XR literacy, which involves not only navigating immersive environments but also interpreting and producing meaning within them (Pegrum & Lan, 2023; Zhang & Miao, 2025). Consequently, learner agency and XR literacy are increasingly viewed as core design variables rather than optional enrichments.

The integration of XR with AI-driven analytics and adaptive feedback is also highlighted as a key avenue to support personalised learning by adapting difficulty, tailoring input, and providing contextualised feedback in real time (Lou, 2025; Song et al., 2023; Zhang & Miao, 2025). Song et al. (2023) connect this adaptivity to input-based theories (Ellis, 2008; Krashen, 1985) and to cognitive load management through calibrated task difficulty. AI-driven feedback, in particular, offers potential for scalable personalisation in immersive communicative settings (Lou, 2025). However, its effectiveness depends on learners' XR/AI literacy, as motivation and engagement are partly shaped by their ability to interpret and act on feedback (Zhang & Miao, 2025).

Two key criteria emerge from the literature: (1) feedback must be pedagogically actionable rather than merely available; and (2) adaptivity must support learning goals and cognitive manageability rather than increase system complexity.

2.6. Implementation constraints and teacher mediation

A consistent conclusion across the literature is that XR efficiency is pedagogically mediated. Teachers must orchestrate XR use through goal setting, task framing, scaffolding, classroom management, debriefing, and curricular alignment (Luo et al., 2024; Parmaxi & Demetriou, 2020; Pegrum & Lan, 2023; Tafazoli, 2024; Tegoan et al., 2021; Wen, 2021). Without this orchestration, XR may generate engagement without learning, or even lead to distraction and frustration.

The large-scale studies reviewed identify a research gap: while learner outcomes are widely studied, teacher education and XR-specific pedagogy remain under-researched, despite being essential for effective and sustainable implementation (Luo et al., 2024; Tafazoli, 2024). Studies also highlight infrastructural and institutional readiness, access constraints, and the need to integrate XR within blended learning ecosystems rather than treating it as a standalone intervention (Pegrum & Lan, 2023; Peixoto et al., 2021; Tegoan et al., 2021). Consequently, teacher capacity becomes a central criterion, requiring not only technical familiarity but also principled instructional design for immersive environments (Luo et al., 2024; Tafazoli, 2024; Zhang & Miao, 2025).

XR efficiency is further constrained by usability and readiness. Initial exposure to XR may increase cognitive load due to interface demands, novelty effects, and unfamiliar interaction patterns, making onboarding and familiarisation essential (Dhimolea et al., 2022; Peixoto et al., 2021; Wen, 2021). Practical constraints include cost, hardware access, technical reliability, cybersickness/physical discomfort, and classroom management challenges (Pegrum & Lan, 2023; Peixoto et al., 2021; Tafazoli, 2024; Tegoan et al., 2021). Moreover, learners' XR/AI literacy influences motivation and engagement, suggesting that benefits may be uneven unless such

literacy is explicitly and intentionally supported (Symonenko et al., 2020; Zhang & Miao, 2025).

Overall, efficient XR pedagogy is clearly influenced by technological issues and requires a dual commitment: (1) designing tasks that are cognitively manageable and accessible, and (2) explicitly developing the literacies needed to learn effectively in immersive environments (Pegrum & Lan, 2023; Tafazoli, 2024; Zhang & Miao, 2025).

2.7. Evaluating XR efficiency: Towards an integrated framework

A crucial methodological implication is that XR efficiency should not be evaluated solely through immediate post-tests. Evidence indicates that its benefits may emerge most clearly in delayed retention and in conceptual/procedural learning rather than short-term recall (Dhimolea et al., 2022; Peixoto et al., 2021; Zhi & Wu, 2023).

Process-oriented approaches further argue that engagement and the quality of cognitive activity are central indicators of effectiveness. Wen (2021), using ICAP, links learning quality to interactive and constructive engagement, showing how process measures can explain variation in XR outcomes. Similarly, CAMIL explicitly integrates both process variables (interest, self-efficacy, cognitive load, self-regulation) and outcome types (factual/conceptual/procedural knowledge and transfer), advocating for relational analysis of these constructs rather than isolated measures (Zhi & Wu, 2023).

Accordingly, an efficiency-oriented XR framework for language learning should include three core evaluation components: (a) engagement and affect as mediators, (b) cognitive load and self-regulation as constraints, and (c) durable learning outcomes (retention/transfer) as targets (Luo et al., 2024; Peixoto et al., 2021; Wen, 2021; Zhi & Wu, 2023).

Some studies make a strong integrative contribution to building a theoretical framework. CAMIL provides a cognitive–affective model linking technological factors, XR affordances (presence/agency), mediators (interest, motivation, self-efficacy, embodiment, cognitive load, self-regulation), and learning outcomes (Zhi & Wu, 2023). Cognitive load theory complements this by guiding instructional design through attention management and sequencing (Song et al., 2023). At the same time, constructivist, situated, and task-based orientations converge in viewing XR as most effective when it supports experiential, interaction-driven language use (Huang et al., 2021; Parmaxi & Demetriou, 2020; Tafazoli, 2024). Finally, motivational theories (including Self-determination Theory, or SDT perspectives) highlight how autonomy, competence, and relatedness can be fostered in XR and AI-enhanced environments (Zhang & Miao, 2025; also consistent with the self-efficacy emphasis in Parra Pérez et al., 2025).

2.8. Pedagogical criteria and design principles for XR implementation

After synthesising the reviewed corpus, it can be stated that the research literature strongly supports a number of convergent pedagogical criteria for efficient XR implementation:

- a. *Predominance of pedagogy and selective adoption of technology*: XR should be used only when it provides clear added value aligned with learning objectives and skill demands (Luo et al., 2024; Parmaxi & Demetriou, 2020; Pegrum & Lan, 2023; Tafazoli, 2024).
- b. *Alignment of technological modality and pedagogical goal*: AR, VR, and MR facilitate different forms of contextualisation, presence, and interaction; efficiency depends on matching modality to pedagogical purpose (Luo et al., 2024; Özçelik et al., 2022; Pegrum & Lan, 2023).
- c. *Situated and authentic task design*: XR is most effective when language is embedded in meaningful contexts and delivered through real-world-oriented tasks, consistent with TBLT and the CEFR action-oriented approach (Council of Europe, 2018; Ellis, 2003; Long, 2015; Nunan, 2004; Parra Pérez et al., 2025; Peixoto et al., 2021; Symonenko et al., 2020; Tegoan et al., 2021).
- d. *Interaction as a central mechanism*: Passive immersion is insufficient and tasks must require action, communication, collaboration, and feedback loops (Luo et al., 2024; Panagiotidis, 2021; Song et al., 2023).
- e. *Embodiment with linguistic focus*: Embodied activities can enhance language learning provided they are pedagogically structured and aligned with language targets (Pegrum & Lan, 2023; Peixoto et al., 2021; Zhi & Wu, 2023).
- f. *Regulation of cognitive load and attention*: Efficient XR designs reduce extraneous load, sequence complexity, and guide attention to linguistically meaningful features (Song et al., 2023; Tafazoli, 2024; Wen, 2021; Zhi & Wu, 2023).
- g. *Affect as mediator*: Some components of affect (motivation, presence-related engagement, reduced anxiety, and self-efficacy) enable participation but must be harnessed through structured pedagogical tasks (Luo et al., 2024; Panagiotidis, 2021; Parra Pérez et al., 2025; Zhang & Miao, 2025).
- h. *Teacher mediation and XR-specific pedagogy*: Teachers must orchestrate XR integration through scaffolding, alignment, debriefing, and assessment; teacher education is a basic requirement (Luo et al., 2024; Parmaxi & Demetriou, 2020; Tafazoli, 2024; Wen, 2021).
- i. *Learner readiness, XR literacy, and usability*: Onboarding, accessible design, and explicit literacy development are fundamental for effective XR use (Pegrum & Lan, 2023; Peixoto et al., 2021; Symonenko et al., 2020; Zhang & Miao, 2025).
- j. *Feedback, adaptivity, and AI integration where pedagogically justified*: AI-enhanced XR can improve efficiency through personalised, actionable feedback,

- provided it supports learning goals and does not increase cognitive load (Lou, 2025; Song et al., 2023; Zhang & Miao, 2025).
- k. *Evaluation for retention and transfer, not only immediate performance*: Efficiency of XR use in language learning should include durable learning outcomes and process measures capturing engagement, load, and regulation (Dhimolea et al., 2022; Peixoto et al., 2021; Wen, 2021; Zhi & Wu, 2023).
 - l. *Contextual feasibility and sustainability*: Cost, access, usability, reliability, and institutional support determine scalability and practical implementation (Pegrum & Lan, 2023; Peixoto et al., 2021; Tafazoli, 2024; Tegoan et al., 2021).

From the review, several pedagogical design principles emerge (see Table 1).

PEDAGOGICAL DESIGN PRINCIPLE	SUPPORTING REFERENCES
<i>Adopt a pedagogical, non-technodeterministic approach</i>	Parmaxi & Demetriou (2020); Pegrum & Lan (2023); Tafazoli (2024); Zhi & Wu (2023)
<i>Differentiate XR modalities according to pedagogical purpose</i>	Özçelik et al. (2022); Pegrum & Lan (2023); Tafazoli (2024)
<i>Use XR selectively and skill-sensitively</i>	Luo et al. (2024); Özçelik et al. (2022); Parmaxi & Demetriou (2020); Peixoto et al. (2021); Tafazoli (2024)
<i>Embed language learning in authentic, situated contexts</i>	Panagiotidis (2021); Parra Pérez et al. (2025); Pegrum & Lan (2023); Symonenko et al. (2020); Tegoan et al. (2021)
<i>Prioritise interaction and action over passive immersion</i>	Ellis (2003); Luo et al. (2024); Panagiotidis (2021); Song et al. (2023); Zhi & Wu (2023)
<i>Exploit embodiment and experiential learning</i>	Pegrum & Lan (2023); Peixoto et al. (2021); Symonenko et al. (2020); Zhi & Wu (2023)
<i>Manage cognitive load explicitly</i>	Peixoto et al. (2021); Song et al. (2023); Tafazoli (2024); Wen (2021); Zhi & Wu (2023)
<i>Leverage affective engagement as a mediating condition</i>	Luo et al. (2024); Panagiotidis (2021); Symonenko et al. (2020); Tafazoli (2024); Zhang & Miao (2025)
<i>Design XR through task-based and constructivist pedagogy</i>	Ellis (2003); Long (2015); Nunan (2004); Parmaxi & Demetriou (2020); Song et al. (2023); Tafazoli (2024); Wen (2021)
<i>Promote learner agency and participation</i>	Parra Pérez et al. (2025); Wen (2021); Zhang & Miao (2025); Zhi & Wu (2023)
<i>Include learner-generated and creative XR activities</i>	Pegrum & Lan (2023); Wen (2021)
<i>Provide immediate, contextualised feedback</i>	Luo et al. (2024); Song et al. (2023); Zhang & Miao (2025)
<i>Support adaptive and personalised learning</i>	Lou (2025); Song et al. (2023); Zhang & Miao (2025)

PEDAGOGICAL DESIGN PRINCIPLE	SUPPORTING REFERENCES
<i>Account for learner readiness and XR literacy</i>	Symonenko et al. (2020); Tafazoli (2024); Wen (2021); Zhang & Miao (2025)
<i>Ensure teacher mediation and orchestration</i>	Luo et al. (2024); Parmaxi & Demetriou (2020); Tafazoli (2024); Tegoan et al. (2021); Wen (2021)
<i>Develop XR pedagogical competence in teachers</i>	Luo et al. (2024); Tafazoli (2024); Zhang & Miao (2025)
<i>Integrate XR within blended and curricular ecosystems</i>	Parra Pérez et al. (2025); Pegrum & Lan (2023); Peixoto et al. (2021); Tegoan et al. (2021)
<i>Evaluate learning through process- and outcome-based measures</i>	Luo et al. (2024); Peixoto et al. (2021); Wen (2021); Zhi & Wu (2023)
<i>Prioritise durable learning and transfer</i>	Long (2015); Parra Pérez et al. (2025); Peixoto et al. (2021); Zhi & Wu (2023)
<i>Acknowledge contextual and accessibility constraints</i>	Pegrum & Lan (2023); Peixoto et al. (2021); Tafazoli (2024); Tegoan et al. (2021)

Table 1. Pedagogical design principles for XR in language learning

This research literature review therefore establishes a theory-based foundation to construct a theoretical framework for XR use in language learning. The review has shown that XR is an efficient language-learning tool: a modality-sensitive, task-centred, cognitively manageable, motivationally mediated, and teacher-orchestrated learning environment.

3. ANALYSIS OF TEACHERS' PERCEPTIONS OF XR IN LANGUAGE TEACHING

The next stage of the study incorporates language teachers' views on the implementation of XR technologies in language learning and teaching. Language teachers' perceptions will shed light on how XR is or may be used by instructors and will also help to identify factors that teachers consider crucial.

The methodology chosen to integrate the language teacher's perceptions was the administration of a short questionnaire (see the Appendix) to 22 practising English teachers, including predominantly ESP university teachers. The individual characteristics of the participants are outlined in Table 2.

VARIABLE	CATEGORY	N	%
Gender	Female	13	59.1
	Male	9	40.9

Age range	26-35	2	9.1
	36-45	5	22.7
	46-55	7	31.8
	56-65	8	36.4
Teaching experience	4-10 years	3	13.6
	11-15 years	3	13.6
	More than 15 years	16	72.7
Teaching context	University	13	59.1
	Secondary education	4	18.2
	University & secondary education	2	9.1
	University & professional development	1	4.5
	University (other languages)	1	4.5
	Primary education	1	4.5

Table 2. Participant characteristics

The questionnaire, entitled “Extended Reality in English Language Teaching”, was designed and administered via Google Forms. It comprised three sections: (1) participants’ demographic information; (2) their use of VR, AR, and MR in educational practice; and (3) open-ended questions eliciting their views on XR in language teaching and learning.

The open-ended questions in the third section were analysed using a grounded theory methodology (Corbin, 2017) combined with inductive thematic analysis, whereby responses were coded into recurring meaning units and subsequently grouped into broader themes.

3.1. Use of XR technologies

Regarding the use of XR technologies, an overall conclusion is that these technologies are not yet mainstream in the sampled classrooms and are not yet being fully exploited by teachers. Indeed, only a small proportion of teachers (22%) report any meaningful use of XR-based technologies in their English classes. This limited uptake is reflected in the following figures: four teachers (18%) report using or having used VR, whereas 18 teachers (82%) report never having used it. Only one teacher (4%) has used AR, and no teacher has used MR in their classes. The tools mentioned in this section are consequently scarce, both for VR (*Umety sets, Google Glass, Second Life, OpenSim, Oculus Quest, AltspaceVR app and BigScreen app*) and for AR (*classvr.com*).

3.2. Teachers' views on XR technologies

The first question of the third section in the questionnaire deals with the perceived potential of XR for language teaching. The analysis shows a strong convergence around a dominant theme: XR's capacity to embed language use in immersive, authentic, context-rich environments for situated communication. This main theme is complemented by perceived benefits for engagement, practice opportunities and affective support. Thus, the most important perceived potential of XR points strongly to immersive contextualisation: the ability to recreate or simulate environments that support situated, meaningful communication. Teachers then extend this central claim through three common explanatory pathways: (1) immersion increases motivation and engagement; (2) XR expands opportunities for linguistic exposure and practice (including speaking); and (3) XR can support affective conditions that encourage participation (confidence, reduced anxiety). A small set of responses, however, highlights that XR's potential depends on implementation feasibility and pedagogical legitimacy (risk of distraction without clear learning value).

The second open-ended question asked participants what they considered to be the most important advantage of extended reality for language teaching and learning. Answers to this question converged on XR's capacity to enable contextualised, immersive communication practice, with secondary emphasis on motivation and engagement, affective support for speaking, and cultural immersion. The most prominent topic was the advantage of contextualised immersion for authentic, quasi-real communication, where language learners can engage in meaningful communication that is close to real-world language use. Responses included significant expressions such as "immersive, meaningful communication in realistic contexts," "quasi-real situations," and "real-life communicative scenarios" without physically leaving the classroom. This advantage was generally regarded as pedagogically relevant because it framed language learning as situated practice. Several responses stated that XR makes it possible to simulate communicative environments and interactions that are otherwise difficult to access, thereby increasing the authenticity of classroom tasks and supporting experiential learning. Other salient themes emerging from the responses are: (1) enhanced learner engagement and motivation; (2) affective benefits for oral communication (safe practice, reduced anxiety); (3) cultural and communicative immersion; (4) professional development and the updating of teaching practice; and (5) uncertainty and scepticism (minor but meaningful), indicating that perceived advantages may depend on greater familiarity or evidence of effectiveness.

In the next open-ended question, participants were asked to state what they believed to be the main disadvantage of XR for language teaching and learning. The most frequently reported disadvantages relate to cost and inequitable access, followed by technical and usability constraints (including physical discomfort), and pedagogical risks when XR is used without sufficient guidance or preparation. A

further topic concerns insufficient teacher training. A small number of teachers reported uncertainty. The strongest and most consistent theme throughout the responses is the high cost of XR hardware, software, and implementation, often linked to limited accessibility and widening digital divides. Other relevant topics that emerged from the analysis of the answers were: (1) technical limitations and implementation fragility; (2) usability and physical discomfort; (3) distraction and the risk of “edutainment” without pedagogical guidance; (4) teacher training and readiness as a barrier to effective use; (5) uncertainty and still-forming perceptions.

The final open-ended question invited participants to provide any additional comments on the topic. This part reveals three cross-cutting perspectives: conditional optimism, structural and pedagogical constraints, and uneven familiarity with XR. Overall, respondents tend to view XR as potentially valuable, but only when supported by adequate resources, training, and pedagogical clarity. First, several comments express conditional support for XR, emphasising that these technologies can be a “powerful tool” when used purposefully and aligned with clear learning goals, enhancing motivation and communication skills rather than replacing traditional methods. One comment highlights the potential of XR combined with AI-driven personalisation to make practice more natural and contextually rich. Second, resource-related and pedagogical constraints are a common concern. Cost, lack of funding, insufficient classroom hardware, and the need for substantial teacher training are repeatedly mentioned. Third, the comments reveal different levels of awareness and acceptance. Some participants openly reported limited understanding of what XR implies and expressed interest in learning more, while others rejected its value completely.

4. THEORETICAL FRAMEWORK FOR THE PEDAGOGICAL USE OF XR IN ESP

The proposed theoretical framework articulates a principled model for the pedagogical use of XR in ESP, grounded in both the literature review and the teachers’ perspectives. It integrates cognitive, sociocultural, affective, and pedagogical dimensions to explain how XR can enhance ESP learning when aligned with clear instructional purposes.

XR is thus conceptualised not as a technology-driven innovation per se, but as a pedagogical mediator that enables learners to engage with disciplinary language in authentic, situated, and interaction-rich environments. The pedagogical value of XR depends on its alignment with communicative goals, task design, and disciplinary specificity rather than on technological sophistication alone.

The framework comprises three principal dimensions: (1) theoretical foundations; (2) pedagogical affordances of XR for ESP; and (3) constraints and mediating conditions.

The first dimension incorporates three *theoretical foundations*. The first is “sociocultural and situated learning perspectives.” The framework draws on sociocultural theory, which conceptualises language learning as a socially mediated process embedded in meaningful activity (Lantolf, 2011; Lantolf & Thorne, 2006). XR environments enable learners to participate in simulated professional practice, thereby facilitating situated forms of language use (Lave & Wenger, 1991; Mondada & Doehler, 2005). These environments also promote interaction, negotiation of meaning, and scaffolded participation, which are central mechanisms in second language development (Canals, 2021; Donato & MacCormick, 1994; Hung & Nguyen, 2022; van Compernelle, 2015). The second theoretical foundation is “embodied cognition and experiential learning.” This highlights XR as an experiential medium allowing learners to engage cognitively and physically with virtual or augmented environments, supporting multimodal meaning-making through visual, spatial, and kinaesthetic channels (Mills et al., 2022; Pasfield-Neofitou et al., 2015; Repetto, 2014; Zhang et al., 2024). In ESP, this embodied engagement is particularly relevant for learning technical content and participating in discipline-specific interactions (Kosmas & Zaphiris, 2020; Ratcliffe & Tokarchuk, 2020; Vázquez et al., 2018). The third theoretical foundation is “cognitive load and multimedia learning principles,” since the framework also incorporates principles from cognitive load theory and multimedia learning (Çeken & Taşkın, 2022; Makransky & Petersen, 2021). While XR can enhance learning through multimodality and immersion, XR environments must be designed to reduce extraneous cognitive load and focus learners’ attention on linguistically and pedagogically relevant elements (Gonnermann-Müller, 2025; Huang et al., 2023).

The second dimension concerns four *pedagogical affordances of XR for ESP*. The first affordance is “authentic contextualisation,” since XR enables the simulation of professional settings, supporting domain-specific language use. This affordance directly supports Long’s (2015) TBLT principle that pedagogic tasks should be derived from target tasks representing what learners actually need to do in the L2. The second is “enhanced interaction and communication,” so that XR environments promote learner–learner and learner–environment interaction, fostering pragmatic competence and communicative fluency. The third is “active engagement,” as XR use is associated with increased motivation, interest, and learner confidence, particularly in oral and professional communication tasks. The fourth is “experiential skill development,” related to the ability of XR to support learning-by-doing approaches that integrate linguistic, procedural, and disciplinary knowledge. This aligns with the CEFR’s conceptualisation of learners as social agents who learn language by accomplishing purposeful tasks where the primary focus is achieving outcomes that require language use (Council of Europe, 2018). In Nunan’s (2004) terms, such experiential tasks involve learners in comprehending, manipulating, producing, or interacting in the target language while their attention is focused on expressing meaning. These four affordances align with ESP pedagogy, where relevance, authenticity, and task-based learning are of prime importance.

The third dimension includes the *constraints and mediating conditions*, since the model explicitly recognises that XR effectiveness depends on several mediating conditions. The first is “pedagogical alignment,” given that XR must serve clearly defined ESP learning objectives. The second condition is “teacher competence and training,” since inadequate XR literacy among teachers was identified as a major hindrance. The third condition is “learner readiness and accessibility,” which implies that differences in learners’ digital skills, physical comfort, and access to XR technologies influence learning outcomes. The fourth condition is “institutional and technical support,” as sustainable XR integration requires infrastructure, time, and curricular support. These constraints reinforce the need to conceptualise XR as part of a broader pedagogical ecosystem rather than as an isolated instructional tool. Figure 1 illustrates the structure of the theoretical framework.

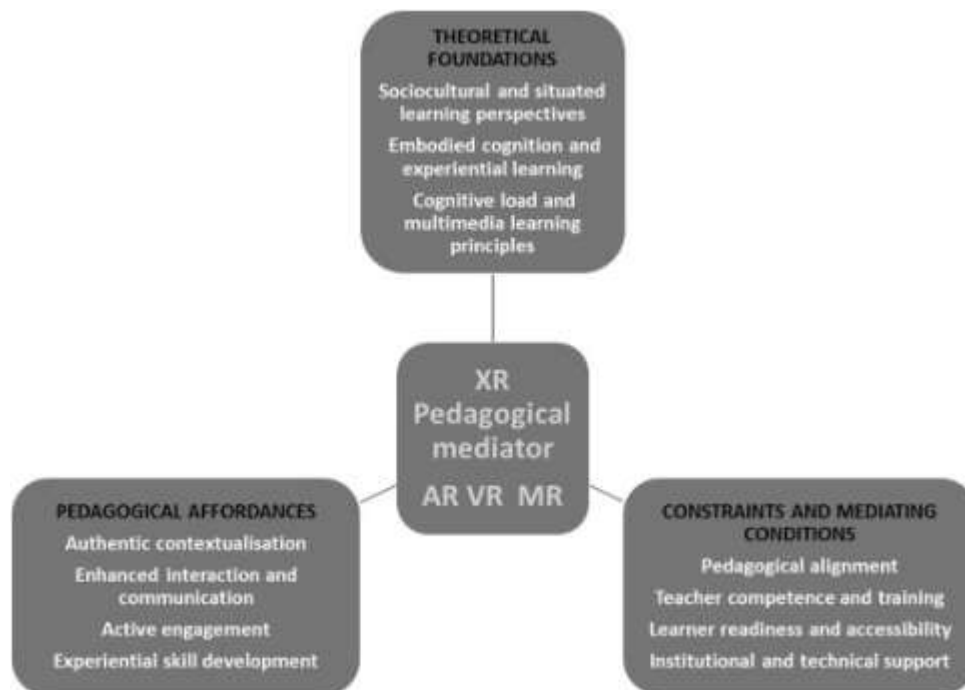


Figure 1. Theoretical framework for the implementation of XR in ESP

Therefore, this framework proposes an integrated model in which XR functions as a mediating tool connecting three components that are central to XR implementation in ESP: (1) pedagogical design (task-based, discipline-specific activities); (2) learning processes (cognitive, social, and affective engagement); and (3) learning outcomes (ESP communicative competence, professional language use, and learner confidence). Within this model, XR is most effective when embedded in structured pedagogical scenarios that combine authenticity, interaction, and reflection, guided by teachers who are pedagogically and technologically prepared. These scenarios should be organised around pedagogic tasks that are progressively

sequenced modifications of target tasks identified through needs analysis. Such tasks require learners to process language pragmatically to achieve evaluable outcomes, with interaction serving as the mechanism through which learning occurs. This task-based structure ensures that XR's technological affordances serve pedagogical goals grounded in established second language acquisition principles.

This theoretical framework provides a foundation for future empirical research on XR in ESP by clarifying the mechanisms through which XR may contribute to language learning. For practice, it underscores that successful XR integration requires pedagogical intentionality, teacher development, and alignment with ESP principles. Rather than transforming ESP pedagogy, XR extends and amplifies existing communicative and experiential approaches when used appropriately.

5. CONCLUSION

This study set out to address a persistent gap in the research on extended reality in language learning by developing a pedagogically grounded theoretical framework for its use in ESP at the university level. Drawing on a systematic synthesis of recent research and on empirical evidence gathered from ESP teachers' perceptions, the article moves beyond technology-centred accounts of XR and reframes these technologies as pedagogical mediators whose educational value depends fundamentally on instructional design, teacher mediation, and contextual conditions.

The findings confirm that XR appears to hold significant potential to enhance ESP learning, particularly by enabling authentic contextualisation, interaction-rich communication, experiential skill development, and increased learner engagement. These affordances are especially relevant in ESP contexts, where language learning is closely linked to disciplinary practices, professional scenarios, and real-world problem solving. However, the study also demonstrates that such benefits are not inherent to immersion itself. Rather, they emerge only when XR use is aligned with communicative and task-based principles, supported by structured interaction, and designed with careful attention to cognitive load, learner readiness, and affective factors.

The integration of teachers' perceptions reinforces this conditional view of XR. While teachers in this study broadly recognise the pedagogical promise of immersive technologies, particularly for contextualised communication and reduced speaking anxiety, they also highlight substantial constraints, including high costs, unequal access, technical limitations and fragility, physical discomfort, and insufficient teacher training. These concerns underscore the importance of situating XR within a broader pedagogical and institutional ecosystem, rather than treating it as a standalone or universally valid solution.

Building on these findings, the proposed theoretical framework articulates three interrelated dimensions: theoretical foundations (sociocultural learning, embodied cognition, and cognitive load principles), pedagogical affordances (authentic contextualisation, interaction, engagement, and experiential learning), and constraints and mediating conditions (pedagogical alignment, teacher competence, learner readiness, and institutional support). By integrating these dimensions, the framework determines how XR can contribute to ESP learning outcomes, such as professional communicative competence and learner confidence, when embedded in purposeful, teacher-orchestrated instructional scenarios.

From a pedagogical perspective, the framework proposed in this study offers several implications for the design of ESP learning environments. First, it suggests that XR should be integrated through task-based and discipline-specific scenarios that simulate professional contexts relevant to learners' fields of study. Second, the findings highlight the importance of teacher mediation and instructional scaffolding, indicating that XR activities should be embedded in structured learning sequences that include preparation, guided interaction within immersive environments, and reflective debriefing afterwards. Third, the framework emphasises the need to develop XR-related pedagogical competences among teachers, including the ability to design immersive tasks, manage multimodal input and align XR activities with curricular objectives. Finally, XR should be implemented as part of blended and multimodal learning ecosystems, complementing rather than replacing other pedagogical approaches.

The framework developed here also opens several avenues for further research. First, empirical studies are needed to test and refine the proposed framework in authentic classroom settings, particularly through longitudinal research designs that examine learning processes, retention, and transfer of ESP competences over time. Second, future research could investigate discipline-specific applications of XR in different ESP domains, in order to determine how immersive technologies support specialised discourse and professional interaction. Third, further work is required on teacher education and professional development, exploring how instructors acquire XR-related pedagogical competences and how institutional support structures influence successful implementation. Finally, emerging technological developments such as the integration of XR with artificial intelligence, learning analytics, and adaptive feedback systems offer promising directions for future studies seeking to personalise immersive language learning environments.

This study contributes to the theoretical consolidation of XR research in ESP by offering a coherent, pedagogy-first model that explains not only what XR can offer, but how and under what conditions it can be educationally effective. For researchers, the framework provides a foundation for more theory-driven and methodologically sound empirical studies. For practitioners, it offers guidance for the principled integration of XR into ESP curricula, emphasising that XR does not

replace established communicative approaches, but rather extends and amplifies them when used appropriately.

[Paper submitted 22 Dec 2025]

[Revised version received 9 Mar 2026]

[Revised version accepted for publication 11 Apr 2026]

Acknowledgement

I would like to acknowledge the use of the artificial intelligence tool ChatGPT 5.1 during the preparation of this manuscript, exclusively for proofreading purposes. I am fully and solely responsible for the final content and interpretation of the article.

References

- Akay, M., & Kessler, M. (2024). The impact of virtual reality (VR) pedagogy on L2 English learners' oral communication and pragmatic competence. *Language Learning & Technology*, 28(1), 1–11. <https://doi.org/10.64152/10125/73568>
- Alhawsawi, S., & Alzaid, F. (2025). Systematic review of virtual reality's uses on vocabulary learning in higher education. *Smart Learning Environments*, 12, Article 60. <https://doi.org/10.1186/s40561-025-00414-0>
- Atkinson, D. (2011). A sociocognitive approach to second language acquisition: How mind, body, and world work together in learning additional languages. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (143–166). Routledge.
- Bassiri, M. A. (2011). Interactional feedback and the impact of attitude and motivation on noticing L2 form. *English Language and Literature Studies*, 1(2), 61–73. <http://dx.doi.org/10.5539/ells.v1n2p61>
- Bordegoni, M., Carulli, M., & Spadoni, E. (2023). Multisensory interaction in eXtended reality. In M. Bordegoni, M. Carulli, & E. Spadoni (Eds.), *Prototyping user eXperience in eXtended reality* (49–63). Springer.
- Canals, L. (2021). Multimodality and translanguaging in negotiation of meaning. *Foreign Language Annals*, 54, 647–670. <https://doi.org/10.1111/flan.12547>
- Çeken, B., & Taşkın, N. (2022). Multimedia learning principles in different learning environments: A systematic review. *Smart Learning Environments*, 9, Article 19. <https://doi.org/10.1186/s40561-022-00200-2>
- Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243. <https://doi.org/10.1080/00461520.2014.965823>
- Christou, E., Parmaxi, A., & Christoforou, M. (2025). Implementation and application of extended reality in foreign language education for specific purposes: A systematic literature review. *Universal Access in the Information Society*, 24, 2061–2076. <https://doi.org/10.1007/s10209-025-01191-w>
- Corbin, J. (2017). Grounded theory. *The Journal of Positive Psychology*, 12(3), 301–302. <https://doi.org/10.1080/17439760.2016.1262614>
- Council of Europe. (2018). *CEFR: Common European Framework of Reference for Languages: Learning, teaching, assessment: Companion volume with new descriptors*.

- Council of Europe. <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>
- de Paula Forero Pataquiva, F., & Klimova, B. (2022). A systematic review of virtual reality in the acquisition of second language. *International Journal of Emerging Technologies in Learning (ijET)*, 17(15), 43–53. <https://doi.org/10.3991/ijet.v17i15.31781>
- Dhimolea, T. K., Kaplan-Rakowski, R., & Lin, L. (2022). A systematic review of research on high-immersion virtual reality for language learning. *TechTrends*, 66(5), 810–824. <https://doi.org/10.1007/s11528-022-00717-w>
- Donato, R., & McCormick, D. (1994). A sociocultural perspective on language learning strategies: The role of mediation. *The Modern Language Journal*, 78(4), 453–464. <https://doi.org/10.1111/j.1540-4781.1994.tb02063.x>
- Draxler, F., Labrie, A., Schmidt, A., & Chuang, L. L. (2020). Augmented reality to enable users in learning case grammar from their real-world interactions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (1–12). Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376537>
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford University Press.
- Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning*, 44(3), 449–491. <https://doi.org/10.1111/j.1467-1770.1994.tb01114.x>
- Falcone, M. (2025). Integrating corpus linguistics and extended reality for ESP learning material design. *International Journal of English Linguistics*, 15(5), 1–12. <https://doi.org/10.5539/ijel.v15n5p1>
- Filippone, A., Carlo, M. S., & Di Fuccio, R. (2025). Virtual English LAB: The impact of virtual worlds on English language learning and life skills in higher education. *Excellence and Innovation in Learning and Teaching*, 1, 23–44. <https://doi.org/10.3280/exioa1-2025oa20579>
- Gonnermann-Müller, J. (2025). *A cognitive load-based design framework for augmented reality training* [Doctoral dissertation, University of Potsdam]. <https://doi.org/10.25932/publishup-68760>
- Huang, X., Zou, D., Cheng, G., & Xie, H. (2021). A systematic review of AR and VR enhanced language learning. *Sustainability*, 13(9), Article 4639. <https://doi.org/10.3390/su13094639>
- Huang, X., Huss, J., North, L., Kirsten, W., & Boyd-Devine, A. (2023). Cognitive and motivational benefits of a theory-based immersive virtual reality design in science learning. *Computers and Education Open*, 4, Article 100124. <https://doi.org/10.1016/j.caeo.2023.100124>
- Hung, B. P., & Nguyen, L. T. (2022). Scaffolding language learning in the online classroom. In R. Sharma & D. Sharma (Eds.), *New trends and applications in Internet of things (IoT) and big data analytics* (109–122). Springer. https://doi.org/10.1007/978-3-030-99329-0_8
- Kaplan-Rakowski, R., & Gruber, A. (2023). The impact of high-immersion virtual reality on foreign language anxiety. *Smart Learning Environments*, 10(1), Article 46. <https://doi.org/10.1186/s40561-023-00263-9>
- Karacan, C. G., & Akoğlu, K. (2021). Educational augmented reality technology for language learning and teaching: A comprehensive review. *Shanlax International Journal of Education*, 9(2), 68–79. <https://doi.org/10.34293/education.v9i2.3715>

- Kosmas, P., & Zaphiris, P. (2020). Words in action: Investigating students' language acquisition and emotional performance through embodied learning. *Innovation in Language Learning and Teaching*, 14(4), 317–332. <https://doi.org/10.1080/17501229.2019.1607355>
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
- Lantolf, J. P. (2011). The sociocultural approach to second language acquisition. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (24–47). Routledge.
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Wiley Blackwell.
- Lou, Y. (2025). The impact of virtual reality environments on English language acquisition: Innovative immersive learning technologies for communication skills development. *Journal of Computational Methods in Sciences and Engineering*, 25(5), 4616–4632. <https://doi.org/10.1177/14727978251337950>
- Luo, S., Zou, D., & Kohnke, L. (2024). A systematic review of research on xReality (XR) in the English classroom: Trends, research areas, benefits, and challenges. *Computers & Education: X Reality*, 4, Article 100049. <https://doi.org/10.1016/j.cexr.2023.100049>
- Lyu, W., & Li, L. (2022). Blended learning of ESP based on SPOC and VR: A case study of maritime English. In *Proceedings of the 2022 6th International Conference on Education and E-Learning (ICEEL 2022)* (117–123). Association for Computing Machinery. <https://doi.org/10.1145/3578837.3578854>
- Mackey, A. (2007). Interaction as practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (85–110). Cambridge University Press.
- Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, 33, 937–958. <https://doi.org/10.1007/s10648-020-09586-2>
- Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum. In H. Das (Ed.), *Telemanipulator and Telepresence Technologies* (282–292). SPIE. Digital Library. <https://doi.org/10.1117/12.197321>
- Mills, K. A., Scholes, L., & Brown, A. (2022). Virtual reality and embodiment in multimodal meaning making. *Written Communication*, 39(3), 335–369. <https://doi.org/10.1177/07410883221083517>
- Mondada, L., & Doehler, S. (2005). Second language acquisition as situated practice: Task accomplishment in the French second language classroom. *Canadian Modern Language Review*, 61(4), 461–490. <https://doi.org/10.3138/cmlr.61.4.461>
- Nunan, D. (2004). *Task-based language teaching*. Cambridge University Press.
- Özçelik, N. P., Ekşi, G., & Baturay, M. H. (2022). Augmented reality (AR) in language learning: A principled review of 2017–2021. *Participatory Educational Research*, 9(4), 131–152. <https://doi.org/10.17275/per.22.83.9.4>

- Panagiotidis, P. (2021). Augmented and mixed reality in language learning. *European Journal of Education*, 4(2), 28–44. https://revistia.com/files/articles/ejed_v4_i2_21/Panagiotidis.pdf
- Parmaxi, A., & Demetriou, A. A. (2020). Augmented reality in language learning: A state-of-the-art review of 2014–2019. *Journal of Computer Assisted Learning*, 36(6), 861–875. <https://doi.org/10.1111/jcal.12486>
- Parra Pérez, D. A. P., Cely, J. P. C., Torres, M. C. C., & Moreno, M. E. U. (2025). A quantitative study on the effects of virtual reality on students' engagement and self-efficacy for learning. *Journal of Educators Online*, 22(3), 187–200. <https://doi.org/10.9743/JEO.2025.22.3.17>
- Pasfield-Neofitou, S. E., Huang, H., & Grant, S. (2015). Lost in second life: Virtual embodiment and language learning via multimodal communication. *Educational Technology Research and Development*, 63(5), 709–726. <https://doi.org/10.1007/s11423-015-9384-7>
- Pegrum, M., & Lan, Y.-J. (2023). Extended reality (XR) in language learning: Developments and directions. *Language Learning & Technology*, 27(3), 1–5. <https://doi.org/10.64152/10125/73528>
- Peixoto, B., Pinto, R., Melo, M., Cabral, L., & Bessa, M. (2021). Immersive virtual reality for foreign language education: A PRISMA systematic review. *IEEE Access*, 9, 48952–48962. <https://doi.org/10.1109/ACCESS.2021.3068858>
- Ratcliffe, J., & Tokarchuk, L. (2020). Evidence for embodied cognition in immersive virtual environments using a second language learning environment. In *2020 IEEE Conference on Games (CoG)* (471–478). IEEE. <https://doi.org/10.1109/COG47356.2020.9231752>
- Repetto, C. (2014). The use of virtual reality for language investigation and learning. *Frontiers in Psychology*, 5, Article 1280. <https://doi.org/10.3389/fpsyg.2014.01280>
- Santos, M. E. C., Lübke, A. i. W., Taketomi, T., Yamamoto, G., Rodrigo, M. M. T., Sandor, C., & Hirokazu, K. (2016). Augmented reality as multimedia: The case for situated vocabulary learning. *Research and Practice in Technology Enhanced Learning*, 11(1), Article 4. <https://doi.org/10.1186/s41039-016-0028-2>
- Schorr, I., Plecher, D. A., Eichhorn, C., & Klinker, G. (2024). Foreign language learning using augmented reality environments: A systematic review. *Frontiers in Virtual Reality*, 5, Article 1288824. <https://doi.org/10.3389/frvir.2024.1288824>
- Skarbez, R., Smith, M., & Whitton, M. C. (2021). Revisiting Milgram and Kishino's reality-virtuality continuum. *Frontiers in Virtual Reality*, 2, Article 647997. <https://doi.org/10.3389/frvir.2021.647997>
- Song, C., Shin, S.-Y., & Shin, K.-S. (2023). Optimizing foreign language learning in virtual reality: A comprehensive theoretical framework based on constructivism and cognitive load theory (VR-CCL). *Applied Sciences*, 13(23), Article 12557. <https://doi.org/10.3390/app132312557>
- Symonenko, S. V., Zaitseva, N. V., Osadchyi, V. V., Osadcha, K. P., & Shmeltser, E. O. (2020). Virtual reality in foreign language training at higher educational institutions. In A. E. Kiv & M. P. Shyshkina (Eds.), *CEUR Workshop Proceedings* (37–49). CEUR-WS.org. <http://ceur-ws.org/Vol-2547/paper03.pdf>
- Tafazoli, D. (2024). Extended reality in computer-assisted language learning. In H. P. Bui & E. Namaziandost (Eds.), *Innovations in technologies for language teaching and learning* (17–34). Springer.
- Taguchi, N. (2022). Immersive virtual reality for pragmatics task development. *TESOL Quarterly*, 56(1), 308–335. <https://doi.org/10.1002/tesq.3070>

- Tegoan, N., Wibowo, S., & Grandhi, S. (2021). Application of the extended reality technology for teaching new languages: A systematic review. *Applied Sciences*, 11(23), Article 11360. <https://doi.org/10.3390/app112311360>
- Thrasher, T. (2022). The impact of virtual reality on L2 French learners' language anxiety and oral comprehensibility. *CALICO Journal*, 39(2), 219–238. <https://doi.org/10.1558/cj.42198>
- van Compernelle, R. A. (2015). *Interaction and second language development: A Vygotskian perspective*. John Benjamins Publishing Company.
- Vázquez, C., Xia, L., Aikawa, T., & Maes, P. (2018). Words in motion: Kinesthetic language learning in virtual reality. In *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)* (272–276). IEEE. <https://doi.org/10.1109/ICALT.2018.00069>
- Wen, Y. (2021). Augmented reality enhanced cognitive engagement: Designing classroom-based collaborative learning activities for young language learners. *Education Technology Research and Development*, 69, 843–860. <https://doi.org/10.1007/s11423-020-09893-z>
- Wienrich, C., Komma, P., Vogt, S., & Latoschik, M. E. (2021). Spatial presence in mixed realities: Considerations about the concept, measures, design, and experiments. *Frontiers in Virtual Reality*, 2, Article 694315. <https://doi.org/10.3389/frvir.2021.694315>
- Xie, Y., Chen, Y., & Ryder, L. H. (2021). Effects of using mobile-based virtual reality on Chinese L2 students' oral proficiency. *Computer Assisted Language Learning*, 34(3), 225–245. <https://doi.org/10.1080/09588221.2019.1604551>
- Xu, G. (2010). Understanding the role of interaction from linguistic, affective, and social perspectives. *TESL Canada Journal*, 27(2), 68–88. <https://doi.org/10.18806/tesl.v27i2.1049>
- Yan, W., & Lowell, V. L. (2025). The evolution of virtual reality in foreign language education: From text-based MUDs to AI-enhanced immersive environments. *TechTrends*, 69, 853–858. <https://doi.org/10.1007/s11528-025-01086-w>
- Yan, W., Li, B., & Lowell, V. L. (2025). Integrating artificial intelligence and extended reality in language education: A systematic literature review (2017–2024). *Education Sciences*, 15(8), Article 1066. <https://doi.org/10.3390/educsci15081066>
- Yudintseva, A. (2023). Virtual reality affordances for oral communication in English as a second language classroom: A literature review. *Computers and Education: X Reality*, 2, Article 100018. <https://doi.org/10.1016/j.cexr.2023.100018>
- Zhang, Y., & Miao, Z. (2025). Enhancing EFL learners' engagement and motivation through immersive technologies: The role of artificial intelligence, augmented reality, virtual reality, and mobile applications. *European Journal of Education*, 60(2), Article e70128. <https://doi.org/10.1111/ejed.70128>
- Zhang, R., Zou, D., & Cheng, G. (2024). Concepts, affordances, and theoretical frameworks of mixed reality enhanced language learning. *Interactive Learning Environments*, 32(7), 3624–3637. <https://doi.org/10.1080/10494820.2023.2187421>
- Zhi, Y., & Wu, L. (2023). Extended reality in language learning: A cognitive affective model of immersive learning perspective. *Frontiers in Psychology*, 14, Article 1109025. <https://doi.org/10.3389/fpsyg.2023.1109025>

RAFAEL SEIZ ORTIZ holds a PhD in Applied Linguistics from the Universitat Politècnica de València. He has been a lecturer in the Department of Applied Linguistics at this university since 1995. He teaches English for Specific Purposes (Technical English for Engineering) at the School of Aerospace Engineering and Industrial Design. His research focuses on Computer-Assisted Language Learning (CALL), particularly on the pedagogical use and evaluation of educational technology resources in language learning. ORCID: 0000-0002-4368-4233

Appendix

The questionnaire

EXTENDED REALITY IN ENGLISH LANGUAGE TEACHING

SECTION 1

1. What is your gender?
2. What is your age range?
3. What is your teaching experience?
4. Where do you teach English?

SECTION 2

1. Are you using / have you used Virtual Reality (VR) in your classes?
2. If so, what Virtual Reality (VR) tools are you using / have you used?
3. Are you using / have you used Augmented Reality (AR) in your classes?
4. If so, what Augmented Reality (AR) tools are you using / have you used?
5. Are you using / have you used Mixed Reality (MR) in your classes?
6. If so, what Mixed Reality (MR) tools are you using / have you used?

SECTION 3

1. In your view, what is the potential of Extended Reality in Language Teaching and Learning?
2. In your opinion, what is the most important ADVANTAGE of Extended Reality for Language Teaching and Learning?
3. In your opinion, what is the most important DISADVANTAGE of Extended Reality for Language Teaching and Learning?
4. Please write any additional comments on the use of Extended Reality in language teaching and learning.