E·S·P·Today
JOURNAL OF ENGLISH FOR SPECIFIC PURPOSES AT TERTIARY LEVEL

Patrizia Anesa
University of Bergamo, Italy
patrizia.anesa@unibg.it

Mariasophia Falcone
University of Bergamo, Italy
mariasophia.falcone@guest.unibg.it

Valeria Giofré
University of Bergamo, Italy
valeria.giofre@guest.unibg.it

Simone Marcutti
University of Genoa, Italy
simone.marcutti@edu.unige.it

Ellen Patat
Istanbul University, Türkiye
ellen.patat@istanbul.edu.tr

DESIGNING EXTENDED REALITY FOR ENGLISH FOR SPECIFIC PURPOSES: A CASE STUDY OF THE ESP-XR SYSTEM

Abstract

The increasing integration of extended reality (XR) into language education has highlighted the need for pedagogically grounded approaches to its use in English for Specific Purposes (ESP). This article presents a case study based on the development of an application aiming to enhance ESP competence in educational and professional contexts via augmented reality (AR) and focusing on three specialized domains (medical, business, and legal English). The study addresses the pedagogical and material-design requirements guiding the development of AR-supported ESP learning environments, including design, interaction modes, and corpus-informed material development. More specifically, it examines how ESP-specific pedagogical principles can be operationalized through AR-based tasks, how technological and interaction design choices support these principles, and how corpus-informed procedures can be used to develop ESP learning materials. The article contributes to current discussions on the role of XR in ESP by outlining a design-oriented framework that connects instructional requirements, material development, and technological implementation.

Key words

ESP-XR, AR-enhanced environment, ESP instruction, task-based learning, corpus-informed design.

1. INTRODUCTION

The rapid evolution of digital technologies has profoundly transformed the educational landscape, generating new pedagogical possibilities and reshaping how learners engage with knowledge. Among these innovations, extended reality (XR) has emerged as a promising domain for reimagining teaching and learning, particularly in the field of language education (Anesa, 2025; Fan et al., 2020; Hockly, 2019; Tegoan et al., 2021). XR affords immersive, interactive environments that transcend the limitations of conventional classroom instruction, offering learners authentic, multimodal, and situated experiences that blend the virtual with the physical. Within this spectrum, augmented reality (AR) is especially promising for language education because of its accessibility and scalability; it requires only devices such as smartphones or tablets and enables the seamless integration of digital objects into real-world contexts (Karacan & Mollamehmetoğlu, 2022).

Research on AR-supported language learning has grown considerably over the past decade, revealing multiple affordances for enhancing vocabulary acquisition, communicative competence, and learner motivation (Ibrahim et al., 2018). Such findings align with broader pedagogical frameworks that emphasize situated learning and multimodal meaning-making (Lave & Wenger, 1991; Vygotsky, 1978), according to which learning emerges through social interaction and the integration of multiple semiotic resources. From this perspective, AR can contribute to transforming classrooms into interactive, practice-based environments in which learners assume active rather than passive roles.

The affordances of AR become particularly salient within the specialized domain of English for Specific Purposes (ESP), which equips learners with linguistic, pragmatic, and discourse skills that are directly applicable to professional and disciplinary contexts, such as medicine, law, business, and engineering, *inter alia*. While ESP pedagogy has long employed case studies, simulations, and role-play to approximate authentic practice (Anthony, 2018; Belcher, 2006; Paltridge, 2012; Tymbay, 2022), such methods often lack the immediacy and interactional complexity of real-world professional encounters (Bhatia, 2008; Hyland, 2006; Taguchi, 2017). XR environments can bridge this gap by immersing learners in realistic, task-based scenarios in which communicative competence is both contextualized and operationalized. Despite this potential, existing reviews of XR in language learning and ESP (Christou et al., 2025; Huang et al., 2021) suggest a relative predominance of virtual reality (VR) applications, while AR remains comparatively underexplored, in spite of its accessibility and ecological validity.

Against this backdrop, the ESP-XR project was conceived.[1] The central outcome was the creation of the ESP-XR app, designed not merely as a technological innovation but also as a pedagogical ecosystem that integrates theory-driven design,

[1] *Enhancing ESP Competence in Educational and Professional Contexts via Augmented Reality,* PRIN 2022 (DD MUR n. 104 del 02.02.2022) funded by EU – NextGenerationEU, CUP F53D23004810006.

inclusivity, and interdisciplinarity (ESP-XR Project, 2025). The app simulates professional and academic environments through AR-enhanced tasks, enabling learners to interact with multimodal resources, engage in authentic communicative practices, and cultivate competences that extend beyond linguistic proficiency to include digital literacy, creativity, and critical thinking.

A distinctive contribution of the ESP-XR system is its emphasis on inclusivity and accessibility. By prioritizing mobile-based AR, the system aims to ensure that immersive ESP learning is not restricted to well-resourced contexts but is made available to learners across diverse socioeconomic and educational environments. This design ethos reflects broader commitments to democratizing access to education and fostering digital equity. Furthermore, the project's interdisciplinarity (combining linguistics, pedagogy, discourse analysis, psychology, and information technology) underscores its ambition to construct a research-driven, practice-sensitive model for technology-enhanced ESP pedagogy.

This study adopts a design-oriented case study approach to identify the pedagogical, linguistic, and design requirements underlying the development of an ESP-XR system as an AR-supported ESP learning environment. It is guided by the following research questions:

RQ1. Which ESP-specific pedagogical and syllabus-design principles should guide the development of an AR-supported learning environment for field-related academic contexts?

RQ1a. Which ESP pedagogical principles inform the design of AR-enhanced learning tasks?

RQ1b. Which technological and interaction-design choices are required to operationalize these pedagogical principles in an AR-supported environment?

RQ1c. How can corpus-informed procedures be used to develop and structure the linguistic materials embedded in ESP-XR?

2. THEORETICAL AND METHODOLOGICAL FRAMEWORK

2.1. The background

The integration of AR into ESP instruction raises a fundamental question about how technological affordances can be aligned with the pedagogical and disciplinary requirements of ESP learning in field-related academic contexts. While AR has been investigated in general education (Kabra et al., 2025), its application to ESP remains relatively under-theorized, particularly with regard to the interplay between instructional design, domain-specific language use, and immersive technologies. The relevant literature highlights both the potential and the limitations of AR in educational settings, pointing to the need for structured design approaches that integrate pedagogical, technological, and contextual factors (Akçayır & Akçayır, 2017; Wu et al., 2013). In particular, no single instructional blueprint exists for the development of AR-based learning environments, especially in ESP settings, where disciplinary specificity, learner needs, and technological constraints have to be dealt with simultaneously. As noted in previous research, the absence of standardized design methodologies in mobile augmented reality (MAR) applications calls for more systematic approaches to the development and evaluation of such applications (Akçayır & Akçayır, 2017; Fan et al., 2020).

Learning in ESP is inherently context-dependent, involving the acquisition of specialized vocabulary, discourse practices, and communicative strategies embedded in professional and academic settings. For this reason, this study adopts a multi-theoretical framework drawing on well-established approaches to language pedagogy and learning theory. At its core lies Communicative Language Teaching (CLT), which focuses on developing communicative competence through meaningful interaction (Canale & Swain, 1980; Hymes, 1972). In ESP contexts, this implies a shift from decontextualized language teaching to the use of language in domain-specific communicative situations.

Building on CLT, Task-Based Language Teaching (TBLT) offers a systematic framework for organizing instruction around goal-oriented tasks that mimic real-world communicative practices (Ellis, 2003; Ellis et al., 2020). This is especially relevant in ESP, where tasks such as consultations, negotiations, or problem-solving activities reflect the communicative demands of specific disciplinary domains. TBLT foregrounds purposeful language use and thus supports the integration of linguistic, pragmatic and strategic competences in authentic scenarios (Richards & Rodgers, 2001).

These pedagogical principles are further supported by Situated Learning Theory, which considers learning as a process embedded in social and professional contexts (Lave & Wenger, 1991). In ESP, this perspective highlights the importance of exposing learners to realistic communicative environments where language is used as part of domain-specific practices. AR technologies provide specific affordances in this respect, enabling the simulation of context-rich scenarios in which learners interact with both linguistic and environmental elements.

Moreover, the framework is informed by Multimodal Learning Theory, which argues that meaning is made by integrating verbal, visual, and spatial resources (Mayer, 2009). This is particularly true in professional communication, where the discourse is inherently multimodal. At the same time, the design of AR-supported learning environments should take into account Cognitive Load Theory, as the introduction of multiple modalities should not overload learners, but rather facilitate their processing and comprehension (Sweller, 1988).

With regard to design, research on AR in education has identified a set of recurring affordances and challenges. On the one hand, AR enables three-dimensional visualization, interactive engagement with content, and context-aware learning, which can facilitate the understanding of complex and abstract ideas (Fan et al., 2020). On the other hand, its implementation presents some issues associated with accessibility, technological limitations, and instructional complexity, when multimodal input is not adequately organized (Akçayır & Akçayır, 2017).

These considerations are especially important in ESP contexts, where instructional design must balance authenticity, usability, and pedagogical clarity. In particular, the development of linguistic materials poses specific challenges. While corpora and frequency lists provide valuable resources, research has shown that domain-specific authenticity and contextual relevance are essential for ensuring that learning materials reflect actual professional practices (Pallavicini & Anesa, 2026).

Considering all these theoretical approaches, there is an obvious necessity to develop a theory-based method of designing AR-supported ESP learning environments. Rather than focusing solely on technological innovation, effective integration requires the identification of pedagogical and syllabus-design principles that can guide the development of meaningful, contextually grounded learning experiences in field-related academic contexts (Akçayır & Akçayır, 2017; Fan et al., 2020).

It is precisely this requirement that prompts the present study, which aims to identify ESP-specific pedagogical and syllabus-design principles that should guide the development of an AR-supported learning environment. The ESP-XR application is examined as a design-oriented case study illustrating how these principles can be operationalized in practice.

2.2. ESP-XR as a design-oriented case study

This paper adopts a design-oriented case study approach to examine the development of ESP-XR, an AR-supported learning environment for ESP in field-related academic contexts. This single case allows for the exploration of how pedagogical and syllabus-design principles can inform the creation of AR-based learning environments (RQ1).

The ESP-XR project, which led to the creation of the ESP-XR app, addresses a specific gap in ESP instruction. While ESP pedagogy has long emphasized the importance of contextualized and task-based learning, conventional instructional formats often struggle to capture the complexity of real-world professional communication. Specifically, learners may have limited opportunities to engage in situated, multimodal, and interaction-rich scenarios that reflect the communicative demands of their target domains. The ESP-XR app is designed and developed to address this gap targeting students in higher education who are preparing to operate in professional fields such as business, law, or healthcare.

Beyond its technological dimension, ESP-XR is grounded in a conceptual framework that integrates relational pragmatics, multimodality, and constructivist approaches to learning (Hyland, 2022). From the perspective of relational pragmatics, effective communication in professional contexts depends not only on linguistic accuracy but also on pragmatic appropriateness, cooperation, and the negotiation of meaning (Taguchi, 2017). Within the ESP-XR app, these principles are implemented through AR-mediated interaction, in which pragmatic awareness plays a central role. Moreover, the design is also informed by multimodality and multiliteracies frameworks (Cope & Kalantzis, 2009), where it is implied that meaning-making emerges through the integration of linguistic, visual, auditory, and spatial resources. In professional settings, communication involves the coordination of multiple semiotic modes. By leveraging AR environments, the ESP-XR app engages learners with communicative strategies in contextually grounded scenarios. Finally, ESP-XR draws on constructivist and situated learning perspectives, particularly Vygotsky's (1978) view of knowledge as socially mediated and Lave and Wenger's (1991) notion of learning as participation in communities of practice. Through problem-based and collaborative tasks embedded in AR scenarios, the ESP-XR app is designed to support the co-construction of knowledge.

The ESP-XR project targets university-level and adult learners with an upper-intermediate proficiency in English (CEFR B2), particularly those preparing for professional or international careers. It may also be relevant for individuals who need to operate in English within specialized domains such as business, law, and healthcare, and who seek to develop greater lexical and pragmatic accuracy and confidence in context-specific communication.

3. THE ESP-XR APPLICATION: SYSTEM DESIGN AND IMPLEMENTATION

The design and implementation of the ESP-XR system were guided by the aim of creating an immersive, interactive, and pedagogically effective environment to support ESP communicative competence. The project specifically draws on principles of user-centered design (UCD), that is understood as an iterative design approach grounded in the active involvement of end users throughout the development process (Abras et al., 2004; Norman, 1988), and interaction design (IxD), which focuses on how users engage with the interface elements and interactive features (Rogers et al., 2011). The aim is to balance educational objectives, usability, technical robustness, and future extensibility. In this context, these approaches were applied to support the development of pedagogically meaningful interactions and learning tasks. In the ESP-XR project, the end users were ESP learners, specifically university students engaged in ESP contexts.

3.1. System design

The ESP-XR platform is structured around a client–server architecture: a mobile application (client) used by learners and a web-based system (server) where content and learning materials are created, stored, and managed. This separation allows the app to deliver interactive activities, while the underlying system supports the organization and updating of materials. Currently running on Android tablets, the app is responsible for delivering the interactive learning experience, including the visualization of 3D environments, the integration of augmented elements into real-world settings, the animation of avatars, and the presentation of audio and dialogue-based activities. It also supports user interaction with tasks and exercises in real time. The web-based system, in turn, supports the creation and management of learning content. It allows instructors and developers to design and organize dialogues, define learning scenarios, associate visual elements with specific tasks, and manage different user roles within the platform. The system is designed to keep learning content separate from the application itself. This allows instructors and editors to update scenarios, dialogues, and media materials without modifying the mobile application itself, thereby supporting the flexibility and ongoing development of ESP content.

Overall, the system is designed to be flexible and adaptable, allowing it to support a range of ESP scenarios that can be expanded over time. From a pedagogical perspective, this means that learning materials and tasks can be progressively developed and adapted to different domains and levels. In this way, the technical design supports the implementation of scenario-based and interaction-driven ESP learning activities, while also contributing to more immersive and contextually grounded learning experiences (Lombard & Ditton, 1997; Slater, 2003).

3.2. Selection of AR functionality

To support this initial design, particular attention was given to the selection of the AR technology underlying the system, as this directly affects how clearly and reliably learning content is presented to users. During the initial prototyping phase, different AR solutions were tested to determine how virtual elements (e.g., objects, prompts, or dialogue cues) could be integrated into the learner's physical environment. These tests included both marker-based and markerless approaches (see Azuma, 1997; Van Krevelen & Poelman, 2010). In simple terms, marker-based systems rely on predefined visual cues (such as images or symbols) that trigger the appearance of virtual content, while the markerless system uses features of the surrounding environment to position virtual elements in space. From a pedagogical perspective, this distinction is important because it influences the stability and clarity of the learning environment, as well as the ease with which learners can engage with the tasks. In the subsequent evaluation phase, several alternative AR solutions were compared in terms of how consistently and accurately they displayed virtual content. This was essential to ensure that learners could focus on the task itself, rather than being distracted by unstable or unclear visual interactions.

3.3. Scenario construction and 3D asset management

A central component of the system is a set of immersive 3D environments representing different ESP domains (Figure 1). For the initial development phase, three areas (health, business, and law) were chosen as representative fields in which communication is highly structured, context-dependent, and closely tied to professional practices (Anthony, 2018; Paltridge, 2012). These domains were therefore used as test cases to explore how scenario-based, task-oriented learning could be implemented across different ESP settings.

Each environment is associated with a visual trigger that activates the corresponding scenario within the app. The 3D environments were developed from pre-existing assets and subsequently adapted to ensure clarity, visual coherence, and compatibility with mobile devices.

Figure 1. Legal scenario marker

The 3D environments were designed to allow individual elements to be reused and adapted across different learning scenarios. Interactive elements allow learners to select objects and access content, such as explanations, vocabulary items, or contextual information. This supports active engagement with the learning materials and helps connect language use to specific professional contexts.

3.4. Avatar-based interaction and dialogue design

At the core of the pedagogical design are interactive dialogues between two digital characters, in which users progress by choosing the most appropriate response from a set of three predefined options. This structure draws on the use of embodied virtual characters to reinforce immersion, social presence, and learning continuity within XR environments (Biocca, 1997; Kilteni et al., 2012).

Each dialogue follows a predefined sequence that guides the exchange between the speakers, allowing learners to engage with typical communicative situations. From a functional perspective, the system integrates AI Text-to-Speech, which provides different voice options and allows for controlled variation in speech delivery. It also synchronizes audio with avatar movements and events in the scene, ensuring that speech and interaction remain coherent. This approach aims to support comprehension and reinforce the connection between language, context, and interaction. In addition to dialogue sequences, conversations may include multiple-choice questions, user-triggered events, and feedback enabling learners to reflect on their choices and consolidate their understanding of domain-specific language use.

3.5. Interaction modes and user experience: The immersive glossary

The final system includes three complementary modes that address different pedagogical needs:

- Augmented mode (Figure 2): Learners access the scenario by pointing the device toward a physical marker, which activates the corresponding 3D scenario within their real environment. From a pedagogical perspective, this mode supports the contextualization of vocabulary and interactions by embedding learning tasks within realistic, scenario-based environments.

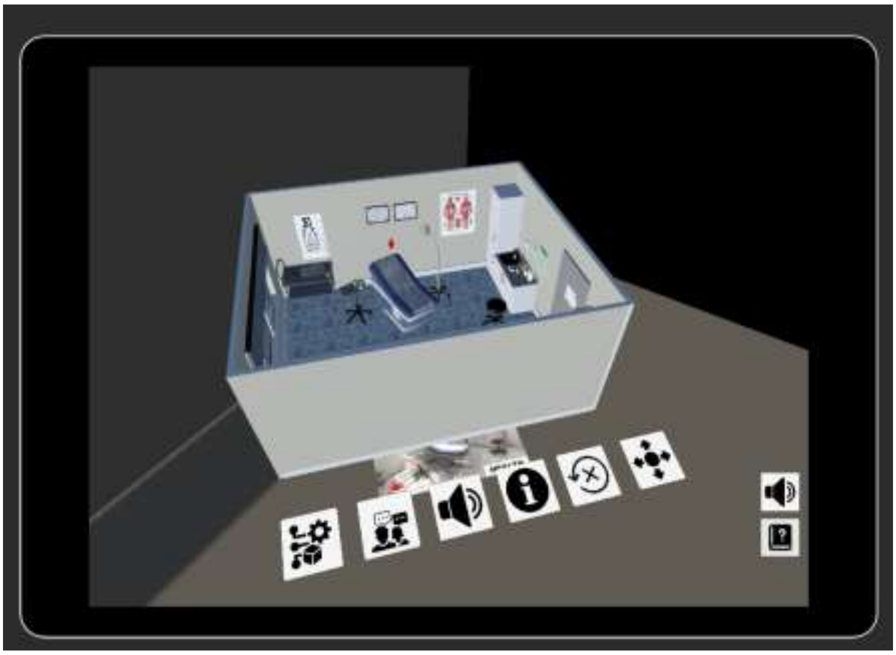

Figure 2. Example of augmented mode

- Isolated mode (Figure 3): A single object is displayed within a neutral environment, enabling users to freely manipulate it by rotating it or zooming. This mode supports focused attention on individual lexical items, allowing learners to isolate and understand domain-specific vocabulary before engaging in more complex scenario-based tasks.

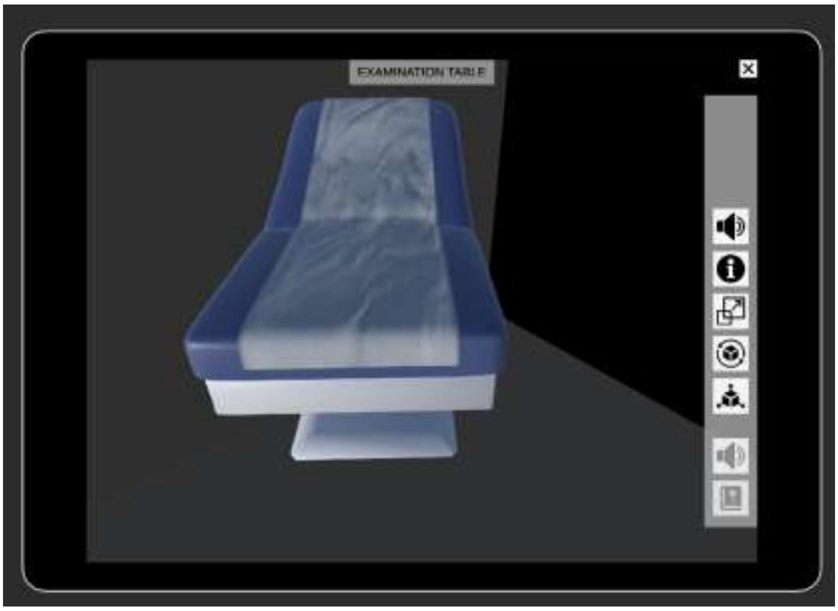

Figure 3. Example of isolated mode

- Navigation mode (Figure 4): Learners can explore a fully virtual 3D environment and move freely within the scene. This mode supports spatial understanding and familiarization with the scenarios, allowing learners to identify key objects and elements before engaging in more complex activities. It is particularly suited for preparatory tasks, as it helps learners build contextual awareness of the setting in which subsequent interactions will take place.

Figure 4. Example of navigation mode

Taken together, these three modes support different stages of the learning process by enabling learners to move from initial familiarization with the scenario, to focused work on individual elements, and to the contextualized exploration of the environment in preparation for subsequent task-based interactions.

4. MATERIALS DESIGN

The development of materials for the ESP-XR app primarily followed a corpus-informed approach. Corpus Linguistics (CL) is based on the empirical analysis of patterns and structures in systematically collected language data (McEnery et al., 2019). Corpora have long been used in language teaching to expose learners to authentic language use across different registers, genres, and professional domains (Flowerdew, 1993). In relation to ESP, learning materials must reflect the communicative practices of the community of reference (e.g., law, medicine, and business). CL can support this, as corpora normally consist of real language data that reflect a given language variety and its use. Therefore, resorting to CL for ESP purposes contributes to providing the specific conventions of their target professional contexts (Cortes, 2012). Arguably, the integration of CL into XR applications seems promising. On a theoretical level, considering the core principles of CL, such as the use of authentic input, attention to frequency and salience, and the identification of recurring patterns (McEnery et al., 2019), corpora are well suited to the design of XR learning environments. Specifically, the authenticity of input aligns with the capability of XR to create immersive simulations that incorporate corpus-based content, thereby providing a suitable setting for learning materials that reflect real-life language use. As regards frequency and salience, they may be strategically supported through visual or auditory prompts designed to highlight common words and expressions in real time. In a similar way, gap-fills or phrase-matching, as tasks integrated within XR environments, can foster pattern recognition.

Lack of authenticity, genuine stimuli, and contextual relevance are common limitations of traditional ESP teaching. Notably, these can hinder learners in their acquisition of phraseological patterns, collocations, and discourse structures, all essential in ESP pedagogy. However, the use of corpus material in the design of learning materials can address this issue. Another potential benefit concerns the approach to XR technology itself. Integrating CL and XR can help mitigate the risks associated with overly technology-driven or simplified XR implementations, which may provide engaging experiences for learners but lack solid pedagogical grounding. Rather, the well-documented pedagogical benefits of using CL for language learning provide the necessary basis for the integration of these technologies into ESP teaching. Nevertheless, at this stage of research, the combination of CL and XR still requires a methodological and conceptual framework that connects the affordances of XR with established principles of corpus-informed language learning, as discussed in recent work (Falcone, 2025). Bearing this in mind, the materials development in ESP-XR sought to integrate these perspectives in practice. As previously mentioned, the app features avatar-based conversations that simulate professional communication scenarios through interactive dialogues, complemented by multiple-choice gap exercises. To guide the development of these conversations, keywords were considered the main starting point for designing the learning materials for legal, business, and healthcare settings, as domain-specific lexis in ESP is essential both for professional communication and disciplinary knowledge. According to CL, keywords serve as central analytical tool to reveal those lexical items that are particularly characteristic of specialized corpora, (i.e., a focus corpus), when compared with a more general corpus (i.e., a reference corpus) (Đurović & Bauk, 2022; Hunston, 2002). Through this comparison, keywords are statistically identified and can be interpreted quantitatively, to reveal corpus subject matter, or qualitatively, to provide insights into the community-specific workings of a language variety (see Bondi & Scott, 2010). This dual capacity makes them a valuable entry point into discourse level aspects of specialized corpora and the communicative practices of their respective discourse community (Hunston, 2002). This applies both to more specialized keywords, i.e., terminology, and to less specialized ones, which may be used with more context-specific meanings; together, they constitute the core component vocabulary of a variety. For the initial prototype of the ESP-XR app, medical English was selected as the first domain for which avatar conversations were developed, serving as a pilot setting to test the corpus-informed material design. Accordingly, the Medical Web Corpus, freely available on Sketch Engine, was used.[2] This corpus comprises approximately 33 million words drawn from authentic medical texts sourced online (Kilgarriff et al., 2010).

To extract the keywords, enTenTen21, also available on Sketch Engine, was used as a reference corpus.[3] From the resulting single-word keyword lists, the first 1,000 highest-ranked items were chosen for further analysis. Afterwards, these keyword lists were manually inspected, and irrelevant items or those of limited pedagogical value, including function words (e.g., prepositions, conjunctions, auxiliaries) and named entities (e.g., people, institutions, places) were excluded. This meant that the final keyword list contained only 555 semantically relevant content words. During this phase, special attention was given to keywords that indicated physical objects, as these would be integrated later as interactive objects. Keywords were eventually divided into smaller subsets that would form the lexical basis for the avatar conversations. These smaller subsets were fed into an LLM to generate self-contained dialogue scripts. This specific step was motivated by the need to ensure consistency and replicability across scenarios. The initial version of the AI-generated dialogues, however, had to be substantially revised and adapted to make sure that content reflected domain specificity and pedagogical focus. Indeed, a recurrent issue observed during the development process was that AI-generated output, while initially time-efficient, often appeared overly artificial and needlessly wordy. Most importantly, this step was aimed at improving pragmatic appropriateness, and ensuring alignment with communicative context, professional tone, and the expected language proficiency level of the target learners (upper-intermediate, CEFR B2). In their final edited version, gaps in the dialogues are accompanied by three answer options for the multiple-choice exercises. Specifically, one would be the correct ESP option. The second would be a grammatically plausible but contextually less precise alternative and, finally, the third is an incorrect option that is often either semantically unrelated or grammatically inappropriate. The reasoning behind this was to make students reflect on contextual appropriateness in line with the pedagogical aims of ESP teaching. By doing this, learners would be more engaged in distinguishing and reflecting on the differences between general English and ESP, as well as domain-specific usage.

[2] https://www.sketchengine.eu/english-medical-corpus/
[3] https://www.sketchengine.eu/ententen-english-corpus/

5. ADDITIONAL MATERIAL AND CONTENT ORGANIZATION

The learning material in ESP-XR was developed following two complementary pedagogical approaches: the ESP approach, which provided the organizing principle for the app's content (Anthony, 2018), and a TBLT approach, which guided the methodology through which learners would acquire that content (Ellis et al., 2020). As outlined in ESP research, particularly with regard to establishing suitable learning objectives and designing genre- and context-based syllabi (Anthony, 2018; Paltridge, 2012), course material should be aligned with clearly defined learning objectives and an appropriate syllabus, informed by a prior needs analysis. In ESP-XR, these objectives focused on communicative competence in professional context, including domain-specific vocabulary, interactional management, and contextually appropriate language use. Since ESP-XR is conceived to support the aforementioned communicative competence, a situational syllabus was adopted, allowing the content to focus on key workplace scenarios, such as doctor–patient interactions or business meetings, while prioritizing the functional language typical of these settings (Anthony, 2018). While the situational syllabus provides the overall organizational structure, it is combined with a task-based methodology, as discussed in the theoretical framework, which hinges on experiential and practice-oriented learning. As a result, each unit follows the classic three-stage TBLT cycle, consisting of a pre-task, a task cycle (comprising the activities embedded in the app), and a post-task (Willis, 1996).

The design of ESP-XR incorporates three-stage task cycle, supported by additional pre- and post-task activities and companion resources hosted on the project website (ESP-XR Project, 2025), which extend and reinforce the in-app learning experience. The structure of the learning sequence can be summarized as follows:

- Pre-task: Lead-in activities designed to activate learners' prior knowledge and elicit relevant vocabulary and communicative scenarios. These activities aim to introduce the topic and prepare learners for the task by drawing on visual prompts, personal experience, or guided discussion (Ahmadian & García Mayo, 2017; East, 2021; Richards, 2006).
- Task cycle: Object-identification tasks and interactive dialogues with embedded quizzes delivered within the app environment.
- Post-task: Activities focused on reflection and production, including brief task reports and role-play exercises in which learners practice the functional language of the target professional setting, consolidating both linguistic and pragmatic competence (Ahmadian & García Mayo, 2017; East, 2021; Richards, 2006).

Companion resources can be used either to support the three-stage learning cycle or as additional material to expand lesson focus through form-focused or meaning-focused tasks. Previous research has reported contrasting results regarding the role of form-focused activities, such as grammar-based input, within task-based instruction (Ellis et al., 2018; Mochizuki & Ortega, 2008; Van de Guchte et al., 2019). Accordingly, the selection and use of these materials may vary depending on the instructional approach adopted. Within the ESP-XR app, two types of instructional resources are provided: (1) a glossary and (2) materials focusing on language use, including grammar and functional expressions. Both types of material are closely aligned with the task cycle, as the targeted keywords, functional language, and grammar structures correspond to those used in the object-identification task and interactive dialogues within the app. Examples of companion material for selected modules include:

1. Medical English – Unit 1: At the General Practitioner's Office
   - Glossary: digestive system, digestive diseases, basic medical procedures and instruments, basic diagnostic procedures
   - Use of English: patients–doctors terminology; describing the radiation of pain
2. Business English – Unit 2: Finance
   - Glossary: key financial terms
   - Use of English: expressions of certainty and probability in business contexts; financial statement structures
3. Legal English – Unit 3: Contract Law
   - Glossary: legal documents, repair and renovation issues, breach of contract, small claims court
   - Use of English: verb patterns (verb + object + infinitive/gerund); language of contracts

Drawing from the established ESP research emphasizing the role of context and disciplinary specificity in language learning (Anthony, 2018; Paltridge, 2012), the glossary and language support materials were integrated with cultural information boxes focusing on culturally specific aspects of each unit to account for intercultural variation in medical, business, and legal practices. Examples include "The UK health system" (Medical English module – Unit 1), "Corporate culture" (Business English module – Unit 1), and "Criminal law" (Legal English module – Unit 1).

Research has highlighted the importance of integrating intercultural competence (IC) into ESP or English-medium instruction settings, focusing on methodology (Aguilar, 2018), empirical studies on learner performance in intercultural setting (Andrienko et al., 2021), and students' attitudes toward intercultural learning (Candel-Mora, 2015; Fielden Burns & Rico García, 2022). The cultural information boxes in the companion serve as starting points for promoting cross-cultural awareness and supporting classroom discussions. They encourage students to compare differences and similarities across medical, business and legal practices, and to negotiate appropriate communication strategies in intercultural professional settings.

All additional materials are hosted (in pdf format) on the project website, and can be freely accessed and downloaded by users. The same interface is used to manage the app's interactive conversations, which are prepared in a structured spreadsheet format. This format follows a predefined organization that allows the system to interpret dialogue sequence and question prompts, while also incorporating basic metadata such as scenario type and description. The interface also supports the selection of visual assets associated with each scenario, enabling their integration into the AR environment.

This system supports accessible, user-friendly updating as well as straightforward customization, allowing instructors to format and upload their own material, thus contributing to populating the app's backend content.

6. CHALLENGES AND DESIGN CONSIDERATIONS IN AR-ENHANCED ESP INSTRUCTION

As previously discussed, AR is increasingly investigated as a medium of ESP instruction that provides immersive, contextually enriched learning environments while fostering motivation, promoting interaction, and easing learning anxiety (Huang et al., 2021; Ibrahim et al., 2018), among other benefits. Despite its potential, its implementation poses several challenges that must be critically examined to support designers and educators in developing effective learning tools. This section focuses on two key dimensions for refinement in dialogue-based AR instruction and AR-enhanced object design.

6.1. Challenges of dialogue-based instruction

ESP-XR incorporates branching dialogues with multiple response options that allow learners to actively engage in simulated communicative situations rather than simply studying linguistic forms in the abstract. This experiential interaction helps to develop learner autonomy and strategic awareness, making communicative competence a tangible, embodied practice (Kilteni et al., 2012) within the learning process. However, dialogue-based interactions present some challenges that require careful attention while designing instructional materials. One of the main concerns relates to time constraints and cognitive load that together affect learners' ability to process linguistic input while taking into account visual and interactive cues. Such multimodal engagement may result in mental fatigue (see, e.g., Jeffri & Rambli, 2021), especially in lengthy or complex scenarios. To reduce overload and optimize learner engagement, dialogues in ESP-XR were intentionally limited to brief interactions. This shortness supports focused attention and efficient learning.

While shorter exchanges do not permit extended negotiations of meaning, requests of clarification, or complex problem-solving, they provide a manageable environment for practicing key communicative skills in professional contexts, such as medical or legal consultations or business negotiations.

Another limitation arises from the scripted structure of dialogues. Predefined interactions ensure coverage of target vocabulary, technical terminology, and specific communicative functions, yet they inherently lack the unpredictability and spontaneity of real-world discourse. Students might learn to identify and select appropriate responses within controlled scenarios. Moreover, inadequate personalization of response options further compounds this issue. Another challenge is multimodal constraints. Despite the visual and interactive potential of AR, nonverbal communication (e.g., gestures, prosody, and facial expressions) is often underrepresented. Such cues are important in professional communication, where subtle signals can indicate politeness, assertiveness, or emphasis (Rzayeva, 2025). Their absence should be taken into account so as to provide learners with other means of gaining the ability to interpret pragmatic intent, recognize emotion, or adapt language to sociocultural norms.

ESP-XR provides dialogues in scripted format, ensuring structured and comprehensive input for the learners, who can interact with language models that exemplify clarity, accuracy, and a professional register. Although learner output is not fully spontaneous but guided at present, this controlled design reduces cognitive load and promotes focused attention on core communicative forms. Future updates will seek to expand the app's interactive potential by allowing users to cocreate dialogues. This is expected to foster greater learner agency and creativity, while maintaining pedagogical coherence.

6.2. Design challenges of AR-enhanced objects

Along with the challenges related to dialogue, the design and implementation of AR-enhanced objects involve additional constraints that can affect the usability and pedagogical effectiveness. Research shows that AR enhances learners' understanding of abstract concepts by providing visual representations that make complex ideas more accessible (Klopfer & Squire, 2008; Wu et al., 2013). In the ESP-XR learning environment, students begin by familiarizing themselves with hotspot-enhanced objects that are presented with visual and AI-generated audio descriptions to permit learners to explore their features and functions within the augmented space. Students then progress to interactive dialogues in which the objects are incorporated as answers to specific gaps or blanks. Embedding the objects within realistic scenarios encourages learners to draw on the knowledge acquired during the exploration phase. This approach strengthens the connection between theoretical understanding and practical use, enhancing both comprehension and long-term retention.

Effective AR design for the ESP-XR requires careful spatial planning and iterative testing to ensure objects remain appropriately sized and legible across different environments and viewpoints. Designers should prioritize clarity and instructional relevance by sequencing visual information so that essential elements are foregrounded and auxiliary cues are introduced only when needed. This approach reduces visual clutter and minimizes learners' cognitive load. Device variability (differences in hardware, resolution, and processing power) can produce inconsistent rendering and interaction behaviour across platforms. While platform-specific optimization and progressive enhancement address many issues, developers should recognize that uniform performance is not always possible in large-scale or remote deployments.

More importantly, abstract concepts present considerable design challenges in specialized ESP contexts, including medicine, business, and law. Many ideas in these domains, ranging from procedural frameworks and diagnostic reasoning to contractual obligations and regulatory principles, are inherently intangible and difficult to visualize. Attempting to render such concepts in AR can risk oversimplification, misrepresentation, or loss of nuance unless they are carefully supported by complementary textual, auditory, or contextual cues. Designers of learning materials must therefore find a balance between the needs for clarity and conceptual fidelity, ensuring that AR-enhanced objects function as effective scaffolds that guide learners' understanding rather than distorting or trivializing complex professional knowledge. Thoughtful integration of multimodal cues, such as synchronized narration, descriptive labels, or scenario-based interactions, can narrow the gap between abstract concepts and perceptual experience, making learning both engaging and accurate.

7. FINAL REMARKS

In conclusion, the ESP-XR project provides a strategically designed framework for enhancing ESP competence in educational and professional contexts through mobile-based AR. The initiative foregrounds inclusivity, accessibility, and interdisciplinary collaboration, aiming to develop a research-informed, practice-sensitive paradigm for technology-mediated ESP instruction. Grounded in multimodality and multiliteracy theories and guided by UCD and IxD principles, the ESP-XR app integrates AR technologies, modular 3D environments, and interaction design in ways that operationalize pedagogical requirements within a digital learning environment (RQ1b). Its architecture is scalable, and adaptable to future extensions, including additional scenarios and platforms, reflecting AR's growing role in fostering engaging, context-rich educational experiences (Lee, 2012; Wu et al., 2013).

The ESP-XR materials combine a corpus-driven approach with XR, using keywords from specialized corpora to develop avatar-based dialogues and task-based exercises. This integration supports authentic, context-sensitive learning and the development of domain-specific language use (RQ1c), while task design encourages learners to make contextually appropriate linguistic choices within simulated professional interactions (RQ1a). Learning is structured around a situational ESP syllabus and task-based approach, with pre-task, task, and post-task cycles integrating lexical, functional, and grammatical resources. Complementary cultural materials foster IC, and a future backend system will allow instructors to upload, customize, and deploy interactive dialogues. Despite inherent challenges in designing AR-mediated objects and facilitating authentic dialogue-based interaction, the ESP-XR synthesizes theoretical rigor and technological innovation, offering a coherent model for scenario-based, multimodal ESP learning.

User testing of the ESP-XR app was conducted within the project framework, with results reported separately (Pallavicini & Tossici, 2026). While the first phase was limited in terms of the number of users and scenarios (focusing only on the health domain, for technical reasons), the second phase of testing involved a larger cohort of students (119 participants) across business, legal, and health domains, and focused on usability, clarity, and perceived educational value in comparison with more traditional ESP learning approaches.

The second phase revealed some interesting results concerning affective gains (Chen et al., 2022; Wang et al., 2020), which can be interpreted in relation to the application's potential to foster 'learning by doing' (Xue, 2022). In AR-supported learning, emotions play a central role as AR environments and objects enhance interest and engagement, through novelty and environmental interaction, thereby supporting learners' attention, active participation in learning tasks and ultimately memorization (Akçayır & Akçayır, 2017; Alzahrani, 2020; Pallavicini, 2024).

Research has demonstrated that long-term vocabulary retention is fostered by the internal kinetic image that performing a gesture may trigger in the learners' mind (Macedonia, 2014; Wu et al., 2024; see also Giofré, 2026). In this respect, AR – with all its affordances – proves particularly effective, as users have the possibility to interact with and manipulate digital representations of objects that simulate reality and the physical world.

Preliminary results from the second phase of testing seem to point in the same direction. The application's AR environments and interactive objects require learners to actively interact with language within contextualized scenarios. In this respect, analyses regarding the emotional impact across all scenarios generally revealed positive responses (e.g., increased surprise and reduced negative emotions), indicating that the integration of AR elements in a language learning application may contribute to a more stimulating learning experience. Similarly, results from the ad hoc questionnaire on the perceived level of learning using AR were also consistent with these findings: in particular, learners reported high levels of agreement with statements related to enjoyment and tended to find the application more engaging and, to a lesser extent, more effective than traditional lecture-based instruction. Overall, these preliminary results suggest that the ESP-XR application provides an engaging and pedagogically promising environment for ESP practice, although it should be noted that these results are based on participants' self-reported perceptions.

Future research should build on these initial findings by conducting more systematic, longitudinal studies to evaluate learning outcomes, learner performance, and the impact of AR-mediated tasks on the development of ESP competences. Further investigation into learner perceptions, task effectiveness, and the integration of such tools into different instructional contexts would also be needed to fully assess the pedagogical value of the approach.

[Paper submitted 20 Dec 2025]
[Revised version received 20 Mar 2026]
[Revised version accepted for publication 12 May 2026]

Acknowledgements

Although this paper was co-authored, responsibility for the individual sections is as follows: Patrizia Anesa: Section 1; Mariasophia Falcone: Section 4; Valeria Giofré: Section 5; Simone Marcutti: Section 3; Ellen Patat: Sections 2, 6 and 7. All authors have read and approved the final manuscript. We would like to acknowledge the use of ChatGPT during the preparation of this manuscript for the purpose of proofreading. We are solely responsible for the final content and interpretation.

References

Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-centered design. In W. Bainbridge (Ed.), *Encyclopedia of human-computer interaction* (445–456). Sage Publications.

Aguilar, M. (2018). Integrating intercultural competence in ESP and EMI: From theory to practice. *ESP Today*, *6*(1), 25–43. https://doi.org/10.18485/esptoday.2018.6.1.2

Ahmadian, M., & García Mayo, M. D. P. (Eds.). (2017). *Recent perspectives on task-based language learning and teaching*. Walter de Gruyter Inc.

Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, *20*, 1–11. https://doi.org/10.1016/j.edurev.2016.11.002

Alzahrani, N. M. (2020). Augmented reality: A systematic review of its benefits and challenges in e-learning contexts. *Applied Sciences*, *10*(16), Article 5660. https://doi.org/10.3390/app10165660

Andrienko, T., Genin, V., & Kozubska, I. (2021). Developing intercultural business competence via team learning in post-pandemic era. *Advanced Education*, *8*(18), 53–69. https://doi.org/10.20535/2410-8286.214627

Anesa, P. (2025). Extended reality in language learning and interpreting: From solutionism to a social constructivist bricolage. *International Journal of Language Studies*, *19*(1), 1–26. https://hdl.handle.net/10446/297849

Anthony, L. (2018). *Introducing English for specific purposes*. Routledge; Taylor & Francis Group.

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, *6*(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

Belcher, D. D. (2006). English for specific purposes: Teaching to perceived needs and imagined futures in worlds of work, study, and everyday life. *TESOL Quarterly*, *40*(1), 133–156. https://doi.org/10.2307/40264514

Bhatia, V. K. (2008). Genre analysis, ESP and professional practice. *English for Specific Purposes*, *27*(2), 161–174. https://doi.org/10.1016/j.esp.2007.07.005

Biocca, F. (1997). The cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of Computer-Mediated Communication*, *3*(2), Article JCMC324. https://doi.org/10.1111/j.1083-6101.1997.tb00070.x

Bondi, M., & Scott, M. (Eds.). (2010). *Keyness in texts*. John Benjamins Publishing Company. https://doi.org/10.1075/scl.41

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*(1), 1–47. https://doi.org/10.1093/applin/I.1.1

Candel-Mora, M. Á. (2015). Attitudes towards intercultural communicative competence of English for specific purposes students. *Procedia: Social and Behavioral Sciences*, *178*, 26–31. https://doi.org/10.1016/j.sbspro.2015.03.141

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

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

Cope, B., & Kalantzis, M. (2009). "Multiliteracies": New literacies, new learning. *Pedagogies: An International Journal*, *4*(3), 164–195. https://doi.org/10.1080/15544800903076044

Cortes, V. (2012). Corpora in the teaching of language for specific purposes. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (1201–1207). Wiley-Blackwell. https://doi.org/10.1002/9781405198431.wbeal0227

Đurović, Z., & Bauk, S. (2022). Testing corpus linguistics methods in ESP vocabulary teaching. *Journal of Teaching English for Specific and Academic Purposes*, *10*(2), 319–337. https://doi.org/10.22190/JTESAP2202319D

East, M. (2021). *Foundational principles of task-based language teaching*. Taylor & Francis Group. https://doi.org/10.4324/9781003039709

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Ellis, R., Li, S., & Zhu, Y. (2018). The effects of pre-task explicit instruction on the performance of a focused task. *System*, *80*, 38–47. https://doi.org/10.1016/j.system.2018.10.004

Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2020). *Task-based language teaching: Theory and practice*. Cambridge University Press.

ESP-XR Project. (2025). *Project overview and aims* (PRIN 2022 No. 104). PNRR/EU NextGenerationEU. https://www.esp-xr.eu

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

Fan, M., Antle, A. N., & Warren, J. L. (2020). Augmented reality for early language learning: A systematic review of augmented reality application design, instructional strategies, and evaluation outcomes. *Journal of Educational Computing Research*, *58*(6), 1059–1100. https://doi.org/10.1177/0735633120927489

Fielden Burns, L. V., & Rico García, M. (2022). Intercultural and linguistic competences for engineering ESP classes: A didactic framework proposal through problem-based learning. *International Journal of Applied Linguistics*, *32*(1), 3–24. https://doi.org/10.1111/ijal.12370

Flowerdew, J. (1993). Concordancing as a tool in course design. *System*, *21*(2), 231–244. https://doi.org/10.1016/0346-251X(93)90044-H

Giofré, V. (2026). Designing material for ESP-XR: An augmented reality-based mobile application for English for specific purposes learning. *International Journal of Language Studies*, *20*(1), 1–24. https://doi.org/10.5281/zenodo.18409171

Hockly, N. (2019). Augmented reality. *ELT Journal*, *73*(3), 328–334. https://doi.org/10.1093/elt/ccz020

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

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.

Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Routledge.

Hyland, K. (2022). English for specific purposes: What is it and where is it taking us? *ESP Today*, *10*(2), 202–220. https://doi.org/10.18485/esptoday.2022.10.2.1

Hymes, D. H. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected readings* (269–293). Penguin.

Ibrahim, A., Huynh, B., Downey, J., Höllerer, T., Chun, D., & O'Donovan, J. (2018). ARbis Pictus: A study of language learning with augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, *24*(11), 2867–2874. https://doi.org/10.1109/TVCG.2018.2868568

Jeffri, N. F. S., & Rambli, D. R. A. (2021). A review of augmented reality systems and their effects on mental workload and task performance. *Heliyon: A Cell Press Journal*, *7*(3), Article e06277. https://doi.org/10.1016/j.heliyon.2021.e06277

Kabra, K., Mane, H., Falke, A., Daitkar, S., & Barik, S. (2025). Augmented reality in education. *International Research Journal on Advanced Engineering and Management*, *3*(05), 1744–1748. https://doi.org/10.47392/IRJAEM.2025.0276

Karacan, C. G., & Mollamehmetoğlu, E. (2022, October 20). Augmented reality in language education: dARe to know. *The FLTMAG*. https://doi.org/10.69732/TMYR6702

Kilgarriff, A., Reddy, S., Pomikálek, J., & Avinesh, P. V. S. (2010). A corpus factory for many languages. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, & D. Tapias (Eds.), *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)* (904–910). European Language Resources Association.

Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, *21*(4), 373–387. https://doi.org/10.1162/pres_a_00124

Klopfer, E., & Squire, K. (2008). Environmental detectives: The development of an augmented reality platform for environmental simulations. *Education Technology Research and Development*, *56*(2), 203–228. https://doi.org/10.1007/s11423-007-9037-6

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Lee, K. (2012). Augmented reality in education and training. *TechTrends*, *56*(2), 13–21. https://doi.org/10.1007/s11528-012-0559-3

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, *3*(2), Article JCMC321. https://doi.org/10.1111/j.1083-6101.1997.tb00072.x

Macedonia, M. (2014). Bringing back the body into the mind: Gestures enhance word learning in foreign language. *Frontiers in Psychology*, *5*, Article 1467. https://doi.org/10.3389/fpsyg.2014.01467

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

McEnery, T., Brezina, V., Gablasova, D., & Banerjee, J. (2019). Corpus linguistics, learner corpora, and SLA: Employing technology to analyze language use. *Annual Review of Applied Linguistics*, *39*, 74–92. https://doi.org/10.1017/S0267190519000096

Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research: LTR*, *12*(1), 11–37. https://doi.org/10.1177/1362168807084492

Norman, D. (1988). *The psychology of everyday things*. Basic Books.

Pallavicini, F. (2024). *Augmented reality (AR) for English for specific purposes (ESP): Evaluation and testing manuals* (Report No. 2). Università degli Studi di Milano-Bicocca.

Pallavicini, F., & Anesa, P. (2026). A narrative review on augmented reality in education. *Education Sciences*, *16*(2), Article 261. https://doi.org/10.3390/educsci16020261

Pallavicini, F., & Tossici, G. (2026). *Augmented reality (AR) for English for specific purposes (ESP): Testing 2 results*. Università degli Studi di Milano-Bicocca.

Paltridge, B. (2012). Genre and English for specific purposes. In B. Paltridge & S. Starfield (Eds.), *The handbook of English for specific purposes* (347–366). John Wiley & Sons. https://doi.org/10.1002/9781118339855.ch18

Richards, J. C. (2006). *Communicative language teaching today*. Cambridge University Press.

Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching*. Cambridge University Press.

Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction design: Beyond human-computer interaction*. John Wiley & Sons.

Rzayeva, E. (2025). The impact of gestures and facial expressions in language acquisition. *Acta Globalis Humanitatis et Linguarum*, *2*(2), 203–214. https://doi.org/10.69760/aghel.025002110

Slater, M. (2003). A note on presence terminology. *Presence Connect*, *3*(3), 1–5.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Taguchi, N. (2017). *Context, individual differences, and pragmatic competence*. Multilingual Matters.

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

Tymbay, A. (2022). Balancing authentic content and "hot topic" discussions in an ESP classroom. *ESP Today*, *10*(2), 310–328. https://doi.org/10.18485/esptoday.2022.10.2.6

Van de Guchte, M., Rijlaarsdam, G., Braaksma, M., & Bimmel, P. (2019). Focus on language versus content in the pre-task: Effects of guided peer-video model observations on task performance. *Language Teaching Research*, *23*(3), 310–329. https://doi.org/10.1177/1362168817735543

Van Krevelen, D. W. F., & Poelman, R. (2010). A survey of augmented reality technologies, applications and limitations. *International Journal of Virtual Reality*, *9*(2), 1–20. https://doi.org/10.20870/ijvr.2010.9.2.2767

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, C.-P., Lan, Y.-J., Tseng, W.-T., Lin, Y.-T. R., & Gupta, K. C.-L. (2020). On the effects of 3D virtual worlds in language learning: A meta-analysis. *Computer Assisted Language Learning*, *33*, 891–915. https://doi.org/10.1080/09588221.2019.1598444

Willis, J. (1996). *A framework for task-based learning*. Addison Wesley.

Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, *62*, 41–49. https://doi.org/10.1016/j.compedu.2012.10.024

Wu, J.-F., Dong, J., Wu, Y., & Chang, Y. P. (2024). Shopping through mobile augmented reality: The impacts of AR embedding and embodiment attributes on consumer-based brand equity. *Information & Management*, *61*(6), Article 103999. https://doi.org/10.1016/j.im.2024.103999

Xue, S. (2022). A conceptual model for integrating affordances of mobile technologies into task-based language teaching. *Interactive Learning Environments*, *30*(6), 1131–1144. https://doi.org/10.1080/10494820.2019.1711132

PATRIZIA ANESA is Associate Professor in English Language, Linguistics and Translation at the University of Bergamo, Italy, where she also coordinates the MA programme "Lingue Moderne per la Comunicazione e la Cooperazione Internazionale." She holds a Ph.D. in English Studies, with a specialization in professional communication. She is a member of CERLIS and Associate Editor of the IDEA project (International Dialects of English Archive). Her research focuses on ESP and intercultural communication. She is currently investigating Extended Reality in ESP learning and leads the PRIN project "Enhancing ESP competence in professional and educational contexts via augmented reality," 2023-2026. ORCID: 0000-0002-0795-745X

MARIASOPHIA FALCONE is an Adjunct Professor of English Language at the University of Bergamo, Italy, where she previously held a Postdoctoral Research position working on Extended Reality applications in ESP teaching. She holds a Ph.D. in English Language and Linguistics from the University of Modena and Reggio Emilia. Her research interests are corpus linguistics and discourse analysis, with a specific focus on digital contexts, sustainability, and the use of XR technologies in language learning and teaching. ORCID: 0000-0002-0589-0474

VALERIA GIOFRÉ is an independent researcher and language instructor at the University of Bergamo, Italy. She earned her Ph.D. in Linguistic Sciences, with a specialization in historical linguistics. Her research has primarily focused on Old English morphosyntax and grammaticalization phenomena. With a background in English language teaching across schools and tertiary education settings, she has also held a post-doctoral research position at the University of Bergamo, where she explored the integration of XR technologies into ESP education. Her research aims to enhance language acquisition and pedagogical practices, while contributing to the advancement of innovative teaching methodologies. ORCID: 0009-0006-6232-837X

SIMONE MARCUTTI is a Research Fellow at the University of Genoa, Italy and Founder of SimarLab. He holds a Ph.D. in Digital Humanities, with a dissertation titled "Immersive systems for social contexts: Designing and developing new types of experiences and interactions." His work lies at the intersection of immersive technologies, interaction design, and digital communication. His expertise includes VR, AR, MR, IoT, and natural user interfaces, applied to both research and real-world projects in web, mobile, and interactive systems. He has collaborated on and contributed to several relevant research and development projects in the fields of immersive technologies and human-computer interaction. ORCID: 0000-0001-8406-2302

ELLEN PATAT is a researcher specializing in multilingualism, language teaching, and comparative literature. She has held research positions at the University of Bergamo (2024-2025) and at Sapienza University of Rome (2023–2024). She has taught English language, Business English, and ESP at the Università degli Studi di Milano, Università degli Studi di Milano-Bicocca, and Università dell'Insubria (Italy), as well as EAP and Italian as a Foreign Language at Bahçeşehir University (Istanbul, Turkey). Her main research interests include EAP and ESP, language teaching, modern and contemporary travel literature, and soundscapes. ORCID: 0000-0003-0994-8415