CERLIS
Centro di Ricerca sui Linguaggi Specialistici
Research Centre on Languages for Specific Purposes

Patrizia Anesa (ed.)

Extended reality, AI, and discursive formations

Educational and professional perspectives

With a Postface by **Stephen Amidon**

12 2025

CERLIS Series

Series Editors: Stefania M. Maci and Luisa Chierichetti

Each volume of the series is subjected to a double-blind peer-reviewing process.

CERLIS Series
Volume 12

Patrizia Anesa (ed.)

Extended reality, AI, and discursive formations: Educational and professional perspectives

With a Postface by Stephen Amidon

CELSB
Bergamo

Studio condotto nell'ambito del progetto PRIN 2022 "Enhancing ESP competence in educational and professional contexts via Augmented Reality" (2022R9K3AR_01), CUP: F53D23004810006, co-finanziato dall'Unione Europea (NextGenerationEU), Missione 4, Componente 2, Investimento 1.1. PI: Patrizia Anesa, Univeristà degli Studi di Bergamo.

CERLIS SERIES Vol. 12

CERLIS
Centro di Ricerca sui Linguaggi Specialistici
Research Centre on Languages for Specific Purposes
University of Bergamo

Extended reality, AI, and discursive formations: educational and professional perspectives
Editor: Patrizia Anesa
ISBN: 978-88-97253-11-2
Url: https://aisberg.unibg.it/handle/10446/293005
DOI: https://dx.doi.org/10.13122/978-88-97253-11-2

Published in Italy by CELSB Libreria Universitaria
Via Pignolo, 113 - 24121, Bergamo, Italy

Contents

VALERIA GIOFRÉ

Evaluating VR-Enhanced Language Learning Apps: Pedagogical Approaches and Practical Implications for ELT

Introduction

Virtual reality (VR) is a relatively new technology that has been gradually becoming more popular, especially in the last few years. Chen et al. (2022: 1) define VR as "a three-dimensional (3D) environment generated by computer technology, which can provide a context similar to visual simulation and other senses". While Pinto et al. (2021: 2) claim that "VR transports the users to another dimension (the virtual world). Instead of being just an observer, the user becomes an actor in a digital world where they can interact and receive feedback from those interactions".

VR is thus a technology that allows users to be immersed in a digital environment, and to simulate activities they would carry out in real life, e.g., VR medical simulations. It is this very aspect of VR technology that has attracted much interest and research, aiming to develop applications that would allow users to train their skills in several areas.

In this respect, English Language Teaching (ELT) has also been the target of research aiming to explore the advantages and disadvantages of using VR technology in educational settings, to assess its ability to enhance students' motivation, self-efficacy, and learning

performance, and to reduce cognitive load. The following are a sample of some of the focuses of recent research in the field: Chen et al. (2022) provide a meta-analysis of quantitative studies focusing on students' linguistic and affective gains when using VR as an educational tool. Hsu et al. (2022) examine students' learning effectiveness and the relationship between self-efficacy and behavioral intention. Xu & Ke (2014) investigate the advantages of gesture-based or motor-based learning involved in using body sensory technology. Finally, Pinto et al. (2021: 1) provide a "systematic review of empirical research aiming to understand whether the use of gaming strategies in virtual reality is beneficial for the learning of a second/foreign language or not".

All these investigations are qualitative and/or quantitative analyses of the affective and effective gains of VR in foreign language acquisition. To our knowledge, however, no investigation has been carried out focusing on the applications currently available on the market to learn a foreign language in a VR environment. This paper therefore aims to provide a review of the existing VR-Assisted Language Learning applications (from now on: VRALL apps) from a pedagogical point of view, in an attempt to assess the strengths and weaknesses of this new methodological approach.

We selected the Meta Quest 3 VR headset to conduct our review. This runs on a custom Android-based operating system designed specifically for VR devices, called Meta Quest OS. We then performed a preliminary survey of possible eligible applications using both Google and the search bar in the Meta Quest Store, the platform where users can browse, purchase, and download VR apps, games, and experiences specifically designed for Meta's Quest VR headsets. Eleven eligible apps were firstly identified, later reduced to three, after discarding those which did not meet our specifications, i.e. availability on the Meta Quest Store and a focus on language learning (thus excluding those that focused solely on soft skills). As a result, the following applications were selected: Mondly VR, Immerse, and Language Lab. Secondly, a series of parameters were chosen on which to base our evaluation. The taxonomy is presented in Section 2, followed by analyses of each app in the subsequent sections.

At this point, we would like to emphasize that the present survey consists of a preliminary evaluation of the selected VRALL

applications, according to specific parameters: as such, no quantitative or qualitative analyses were carried out to investigate either users' performance or users' perception after using the apps.

2. Methodology

In order to provide a systematic analysis of the selected applications, it was first necessary to select appropriate parameters on which to base our investigation. Since research on VR-Assisted Language Learning (VRALL) applications is still relatively limited, we opted for taxonomies that had already been used to evaluate and test mobile-assisted language learning (MALL) (see Chinnery, 2006). We were able to identify three taxonomies that fitted our purposes: the first one taken from Rosell-Aguilar (2017), the second from Kim & Kwon (2012), and the third from Richards (2006).

These taxonomies were then contrasted in order to eliminate any possible overlapping criteria, i.e. identical parameters which were named differently, despite describing the same, or similar, feature. We subsequently integrated the criteria with one further parameter relevant to our purposes, namely Extended Reality (XR)[2] usability. The resulting framework, which is shown in Table 1, has been divided into three categories: user interface/technology, pedagogy, and user experience[3].

User interface (UI). As far as the first category is concerned, we adhered to the criteria developed by Rosell-Aguilar (2017: 253), which appeared relevant enough to evaluate the apps' technical features. As mentioned above, Rosell-Aguilar's questions have been integrated with one parameter specifically focusing on XR:

[2] Extended Reality (XR) is used as an umbrella term to encompass Virtual Reality, Mixed Reality and Augmented Reality, all of which share the common feature of blending digital content with the real world to enhance user experiences.

[3] Rosell-Aguilar's categories were reduced from four to three by integrating and adapting some of the language-learning criteria proposed in his work into a broader category "Pedagogy".

- what type of XR integration does the app support? Augmented Reality (AR); Virtual Reality (VR); Mixed Reality (MR)? Anesa (2025) defines each type as follows: “VR is a multidimensional immersive space, AR overlays virtual content on real world, and MR merges the two perspectives”, whereas the notion of XR “broadly covers these approaches and also possible future ones in that it is projected towards the acquisition of any potential real virtual and hybrid environments”.

Pedagogy. As regards the pedagogical framework underpinning the selected apps, we decided to base our analysis on Richards’s survey on CLT (Communicative Language Teaching) (see Richards, 2006). We then compared Richards’s (2006) analysis of instruction typology with the pedagogical parameters provided by Kim & Kwon (2012). The resulting framework was divided as follows:

- Methodological approaches, in which both traditional and recent approaches were taken into account (audio-lingual and structural-situational, skills-based, notional-functional, content-based etc.), as well as the learning strategy used (deductive, inductive, mnemonic etc.) and the structuring of the activities (whether appropriate scaffolding and feedback are present etc.)

- Language and linguistic focus, in which the specific skills, or linguistic aspects focused on by each app, were considered (speaking, listening, reading, writing, vocabulary and grammar), as well as the level of the language (beginners, intermediate, or advanced).

- Also, the pedagogical frameworks VRALL apps are based on: constructivist learning, experiential learning, etc.

User experience (UX). To evaluate this category, we also decided to consider the Rosell-Aguilar (2017: 253) parameters, which allowed for an assessment of the apps' user-friendliness during use. At this point, a clarification is in order: this investigation was not intended to assess UX factors such as self-efficacy, motivation, or other affective gains, as our main purpose was to provide an evaluation of VRALL apps from a technical and pedagogical perspective.

A thorough analysis of VRALL apps' UX could not overlook an evaluation of the sense of presence and degree of immersive experience provided by the apps. We therefore decided to integrate Rosell-Aguilar's parameters with two additional categories: immersion and sense of presence. The criteria to evaluate these parameters were based on studies by Slater & Wilbur (1997) ("degree of immersion") and Schubert et al. (2001) ("sense of presence").

In short, immersion was assessed by examining the extent to which head-mounted displays (HMDs) – in our specific case, Meta Quest 3 – were capable of delivering an "inclusive, extensive, surrounding, and vivid illusion of reality to the senses of the human participant" (Slater & Wilbur 1997).

Sense of presence, on the other hand, is defined as a "psychological state in which even though part or all of an individual's current experience is generated by and/or filtered through human-made technology, part or all of the individual's perception fails to accurately acknowledge the role of the technology in the experience" (Riva, 2009: 159). Sense of presence was assessed here using Schubert et al. (2001)'s 10 factors: Spatial Presence, Quality of Immersion, Involvement, Drama, Interface Awareness, Exploration of Virtual Environment (VE), Predictability & Interaction, and Realness. Unlike Schubert et al. (2001), who provided a quantitative analysis of these factors, we opted for a qualitative analysis of the above-mentioned, as our investigation was not supported by a user-oriented experiment involving multiple participants. Instead, we conducted a self-observation study, where qualitative observations were documented through reflective notes.

User interface	Pedagogy	User experience
Interface: cluttered / uncluttered	Methodological approaches: Audio-lingual, Structural-situational, Skills-based, Functional, Notional, Content/Topic-based, Task-based, Text/genre-based, Competency-based	Degree of immersion
Navigation: intuitive / chaotic	Focus: Listening, Reading, Speaking, Writing, Grammar, Vocabulary, Pronunciation, Varieties of English, Culture,	Sense of presence
Instructions: yes / no	Maximizes exposure to target language: yes / no	Interaction: yes / no
Stability: yes /no	Target Learners: Children / Teenagers / Adults	Interactivity: active / passive
Gamification: yes / no	Interest: General / ESP	Sharing content: yes / no
Support: yes / no	Proficiency level: Beginners, Intermediate, Advanced	Badging: yes / no
Offline work: yes / no		Price: free / lite version[4] / full paid version
XR type: VR, AR, MR		Registration: yes / no
		Advertising: yes / no

Table 1. Framework for VRALL evaluation (adapted from: Kim & Kwon, 2012: 39; Rosell-Aguilar, 2017: 253)

[4] A lite version of an app is a simplified version of the original app, which generally offers a more basic user experience.

3. Mondly VR

In this section, we will analyze Mondly VR, a VRALL application developed by Pearson plc, primarily aimed at vocabulary acquisition and presented by its developers as "the first language learning experience with chatbot and speech recognition in the world" (Mondly, n.d.). Mondly VR offers fully immersive virtual environments (VE), where users engage in quasi-authentic interactions with AI-generated interlocutors. As briefly mentioned in the introductory section, VR environments foster both linguistic and affective gains because they provide authentic learning environments in which "users are able to explore and interact with a simulated environment using multiple senses" (Hsu et al., 2022: 1620). In this respect, Mondly VR largely meets expectations, as it involves "real-time simulation of an environment that people can explore and interact with through multiple senses" (Vesisenaho et al., 2019), so that users in VR environments "are made to believe that what they see and feel is real" (Inoue, 2007 as cited in Hsu et al., 2022).

3.1 User interface

The app features a visually appealing interface that immerses users in a home environment – a modern two-story house overlooking what appears to be a campus or resort on an island. This setting aspect is particularly interesting as it includes additional features, such as a "metaversity" which, although not yet developed, suggests potential future expansions.

The app also includes a gamification mode, defined as "the use of game design elements and game mechanics in non-game contexts" (Domínguez et al., 2013). This mode allows users to teleport to an additional environment within the home interface, called Game arcade.

In this environment, learners can engage in an object-grabbing game – involving grabbing stickers that depict objects called out by a voice – whose objective is to help users practice vocabulary acquired in the app's activities, as well as to develop their listening skills. The inclusion of a gamification mode marks a significant step forward in the app's development for two key reasons:

1. Research has demonstrated that integrating gaming strategies with VR technologies is associated with positive outcomes in foreign language learning (Kim et al., 2018; Pinto et al., 2021), as it fosters motivation and engagement (Nah et al., 2014; Su & Cheng, 2015) and self-efficacy (Sitzmann, 2011).

2. Users are required to make movements and gestures to 'beat' the game and climb the international ranking. Research has demonstrated that incorporating "gestures or motor activities via body sensory technologies, may foster and retain concentration and learning engagement" (Xu & Ke, 2014: 713).

Albeit still limited, the Game arcade environment shows promise, as there are features that suggest potential for future expansions.

Mondly VR, however, also presents some features that could be improved. In particular, the interactive element, i.e. the way users interact with VR items, might benefit from further development. In most scenarios, users are denied the possibility to manipulate objects within the VRE: in other words, they cannot pick up, drop, rotate, zoom in / zoom out objects or tap on them to acquire additional information. Only in a few cases, is it possible for users to interact with the VR items: the environments included in the second set of activities, called "Immersive vocabulary", allow the user to tap on the objects and receive feedback. For example, in the task entitled 'Space' users can tap / click on the planets and satellites and receive feedback, so, for example, by clicking on Mars, additional information concerning the composition of its atmosphere will be superimposed, or by clicking on the Moon, users will get to see the dark side and so on. Although limited, these interactions provide the amount of authenticity and physical interaction that are so important in learning (see Xu & Ke, 2014).

Regarding the other interface parameters, Mondly VR largely fulfills all the technical requirements a good app should meet according to Rosell-Aguilar (2017). More specifically: the interface is clear and uncluttered, and intuitive to navigate. The app never crashed during testing. Instructions on how to use the app and carry out the activities are provided, although only for beginners or intermediate learners. Users starting as advanced learners may encounter difficulties with the initial set of activities, i.e. "Immersive conversation", as there is no tutorial to guide them through the app. This represents another feature that needs adjustment.

3.2 Pedagogy

Overview. The purpose of the app is to help learners acquire new vocabulary and collocations, as well as a limited set of fixed expressions or phraseology, that learners might need to use in some common situations, such as travelling, working, or meeting new people. The app is built around three different sets of activities, each of which adheres to a different methodological approach: immersive conversations, immersive vocabulary, extensive learning. However, the absence of an entry test to assess learners' initial level makes it challenging for users to determine the most suitable starting level for the activities, leaving this decision to self-evaluation. In our view, this represents a significant limitation that the developers might want to address.

Immersive conversations. The methodology underpinning the first set of activities is functional. According to Richards (2006: 11), a functional syllabus "is organized according to the functions the learner should be able to carry out in English, such as expressing likes and dislikes, offering and accepting apologies, introducing someone, and giving explanations". The purpose of the first set of tasks is indeed to make learners practice 'survival English', i.e. common phraseology that it is necessary to master in selected speaking situations (e.g. at the restaurant, shopping, making appointments, etc.).

Despite being very useful to practice conversational skills, Mondly VR presents a few shortcomings. First of all, the selection of real-life communication scenarios is too limited. Secondly, interactions with the AI-generated avatar lack the flexibility to recognize utterances that deviate significantly from the pre-programmed ones, which may pose a problem as the system does not encourage creative thinking and spontaneous production. For instance, advanced learners are likely to use a wider range of expressions than the preset ones, which are not recognized by the AI interlocutor. This could lead to frustration. This limitation is closely connected to another issue: the activities and tasks labelled as advanced English do not fully meet the expected level of complexity, as the app offers the same type of activities across all proficiency levels. Furthermore, learners' progress is assessed solely through recall tasks, which evaluate the retention of fixed expressions learned at the beginner or intermediate level. While this approach may offer some advantages in terms of progressive learning and scaffolding, over time the activities can become repetitive, leading to a potential loss of interest among learners.

Immersive vocabulary. This set of activities may broadly be referred to as a content-topic-based approach. Richards (2006: 28) defines 'content' as follows: "the information or subject matter that we learn or communicate through language rather than the language used to convey it". In the app, learners encounter a variety of topics – ranging from space and animals, to fruit and vegetables – and are given the opportunity to learn new words and expressions related to the selected topic. This is achieved by drawing on mnemonic strategies, such as repetition, recall, and consecutive translations of short sentences. Theoretically, this could represent an effective way of introducing English vocabulary to learners. However, in our view, this section requires further developing for the following reasons:

1. Limited range of topics.
2. The 'subject matter' is provided in the source language.
3. Only the 'Space' activity appears to be fully developed as an immersive content-based task, where the new vocabulary related to Space is introduced through simple but engaging anecdotes.

> In contrast, the activities for the other topics (Animals, Fruit, and Vegetables) do not much differ from those included in the Extensive learning section (described below). Additionally, they lack supplementary information or content, making the topics feel more like excuses to utilize different immersive environments, such as a zoo and a grocery store than ones with a defined pedagogical purpose.

In our view, this activity has the potential to be very effective, provided that the developers clarify its learning purposes. If the aim is to introduce new vocabulary, there is no justification for keeping this section separate from the Extensive learning segment, as simply providing a different type of VR environment does not entail the creation of a distinct section. If, on the contrary, the developers' intention aligns with Krahnke's (1987: 65) perspective of "teaching [of] content or information in the language being learned with little or no direct or explicit effort to teaching the language itself separately from the content being taught", then the content must be presented in the target language.

Extensive learning. This set of activities represents the most developed section of the entire app, featuring 55 authentic situations and topics, whose phraseology the learner is ultimately expected to master. Each exercise is structured as follows: words, mainly nouns and verbs or multi-word expressions, are first introduced and tested; subsequently, relevant phrases that are deemed useful for the corresponding authentic situations are presented. The methodology underpinning this set of activities is functional, and the exercises rely on mnemonic strategies, such as repetition and recall, similar to those found in the Immersive conversation section. The two afore-mentioned sections in fact differ mainly in the immersive experience they offer: Immersive conversations allows users to engage in quasi-authentic environments, e.g. shops, hotel reception, hotel rooms, where they interact with AI-generated interlocutors, while 'Extensive learning' offers a more traditional classroom environment with teacher-student interactions.

All sections include ongoing assessments of the vocabulary and phraseology learned, which helps learners retain new words more effectively. However, incorporating an entry test to assess learners'

initial level, along with a final assessment after a certain number of tasks would also be advisable. These additions would enable learners to determine whether to begin activities as beginners, intermediate or advanced, as well as monitor their progress.

3.3 User experience

In terms of UX, Mondly VR provides a fully immersive experience, as described by Slater & Wilbur (1997): "Immersion […] describes the extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding, and vivid illusion of reality to the senses of a human participant". The environments are designed to be as realistic as possible, giving users the sensation of being immersed in a real-world setting, while also interacting with real people. For example, the first scenario in the 'Immersive experience' section takes place on a train, where learners feel as though they are actually seated on a passenger seat with other passengers sitting all around. Additionally, the sense of presence feels highly authentic: users may indeed feel prompted to look out of the window or even to lean forward.

Of the 10 factors defining presence according to Schubert et al. (2001), all are fulfilled except perhaps the sense of realness, or, in terms of Slater & Wilbur (1997), "the suspension of disbelief", referring to the degree to which the virtual world seems real to users. Although users can experience a high level of involvement, which is further enhanced by the spatial audio system, the cartoonish design of the environments makes it still challenging to perceive them as fully realistic, 'place-like' worlds. This sense of immersion is also disrupted by occasional glitches in the Extensive learning section, when the AI avatar cannot process user input (which may in turn be caused by audio system issues).

As regards the other parameters, Mondly VR is primarily designed as an app for single users, thus not allowing for much interaction with other students, such as sharing material or content on social media. However, as mentioned previously, users can still practice their interactive skills in exchanges with AI interlocutors. Regarding pricing, the app is not free and is sold at a full price of €14.99. At the

time of our purchase, there was a 33% discount, allowing us to buy it for €9.99. The app contains no pop-up ads or in-app purchase offers, and no lite version is available. Registration is not required as users are automatically identified through their Meta account.

Our overall impression is that Mondly VR has considerable potential but requires further development, especially concerning the immersive conversation section. We would also recommend that developers differentiate the content offered at each level of language (beginners, intermediate, advanced), as, the only current difference between levels is the number of questions included in the ongoing assessment.

4. Immerse

In this section, we analyze Immerse, a VRALL application developed by the company of the same name, which is primarily designed to enhance speaking skills. The application is described as a "social, virtual reality language platform"[5] and promoted as more successful than its competitors at reducing speaking anxiety and more effective at developing fluency and confidence across all skills (Immerse, n.d.). This claim probably stems from the layer of anonymity avatars offer within VRE, potentially reducing the feeling of anxiety some users can experience when practicing their speaking in real-life scenarios. Consequently, in spite of all the challenges that new technology is said to bring about, the cognitive load may actually diminish.

Therefore, the way this application is described further contributes to the ongoing debate about whether immersive VR helps reduce cognitive load or, conversely, increases it. Literature has repeatedly demonstrated that cognitive load in immersive VR is, in fact, higher (Makransky et al., 2019; Lin et al., 2021). On the contrary, Immerse developers' claim points toward the opposite direction, making this an issue still open to debate and worthy of further

[5] Description provided by Misty Wilson Ed.D., as previewed in the Meta Quest Store.

investigation. In this section, we will also attempt to address this issue, drawing on personal observations from a personal trial of the app.

4.1 User interface

Immerse's interface is clear, uncluttered, and easy to navigate. The developers have managed to create intuitive and familiar environments, so that learners who are not used to VR can quickly adjust to the technology. More specifically, once logged-in, users are sent to a home interface consisting of a virtual studio that serves as the hub of all other environments and features. From this environment, users can choose to join a tutorial in order to familiarize themselves with both the technology and the application, or to begin the available activities. The tutorial clearly details how to move around the app's environments – by sliding or teleporting – how to interact with objects, and how to use the backpack utilities (scanner, camera, laser pen, notepad, 3D pen). It also explains how to participate in lessons and earn points and 'stamps', which form the Immerse customized reward system.

The ability to interact with objects and environments is featured in most VRALL applications, as discussed in Section 3, due to the established relationship between sensory (gesture-based) learning and cognitive processing (Xu & Ke, 2014). In this respect, Immerse is no exception: users can manipulate objects – specifically by grabbing and throwing them – and virtually scan items using the virtual scanner to identify them in the target language and learn their correct pronunciation by physically performing the action. With respect to this, the inclusion of backpack utilities is an asset unique to this app: users can note down new words with the notepad and 3D pen, which is particularly innovative as no other apps share this feature. The only drawback is that the pen is difficult to use without a compatible VR stylus, limiting accessibility and inclusivity for some users.

Another characteristic that Immerse has in common with Mondly VR is the inclusion of a gamification mode. On this matter, Immerse users can engage with individual activities aiming at testing memorization of new words, such as 'whack a word', and a shooting gallery, as well as group activities, like 'guess the word' and pictionary.

As discussed in the previous section, including gaming strategies represents a significant advantage, as the gaming element has been proved to be essential in learning environments (Domínguez et al., 2013; Kim et al., 2018).

4.2 Pedagogy

Overview. As mentioned in Section 4.1, the app's purpose is to help learners improve their speaking skills and feel more confident when speaking. Learners are encouraged to interact actively with both their teachers and peers. Peer interaction, in particular, may be considered an asset unique to this app when compared to the others considered. Research has demonstrated that learners benefit from interacting with their peers (Sato & Ballinger, 2016): "there is some evidence indicating that learners feel more comfortable during peer interaction in comparison with student-teacher interaction. It can be argued that this comfort level positively affects learners' L2 processing by helping them notice and point out errors in their partners' speech and encouraging them to modify their own errors when given feedback" (5).

All classes are held live. Users can choose the learning experience they want to join from the home environment (Classes, Connect, Practice, Play, Tutorial, You). The first four sections are focused on language learning. More specifically, users who want to engage with live sessions have to select 'Classes'. All upcoming classes at all levels are listed in the user's Home Room. The methodology underpinning classes at an intermediate and advanced level is functional, i.e. an approach that focuses on "the functions the learner should be able to carry out in English, such as expressing likes and dislikes, offering and accepting apologies, introducing someone, and giving explanations" (Richards, 2006: 11). Functional syllabi have been in use since their theoretical inception in the 1970s. In the case of Immerse, the innovative element lies in the realistic scenarios students can experience, such as in the restaurant, the shopping center, the park, the bar, the doctor's office, etc. This, of course, helps enhance students' sense of presence and, consequently their "intrinsic motivation, self-efficacy, and learning in sequence" (Hsu et al., 2022: 1620). The

possibility for students to be immersed in situations that mimic real environments, where they are likely to use certain linguistic functions and expressions, constitutes the main asset of this app and is a feature that distinguishes all VRALL approaches from more traditional ones, i.e. face-to-face classes or distance learning via video conferencing apps.

Moreover, learners have the opportunity to physically interact with items (grab objects, write on boards etc.), or express their positive attitudes (i.e. give a thumbs up), which have been demonstrated to increase learners' sense of presence (Xu & Ke 2014; Slater et al. 1998), engagement, and learning. As suggested by Riva (2009: 161), "the subject is 'present' in a space if he/she can act in it". The main advantage of harnessing gestures in VRALL contexts is, first and foremost, the improvement of users' learning performances, which comes from the reduction of "the cost associated with cognitive activities by releasing resources of working memory" (Xu & Ke 2014: 730). Several reasons explain why embodiment fosters language learning and aids memory retention: one reason is multimodal processing, whereby gestures evoke multi-sensorial modalities (primarily visual and auditory), allowing for quicker memorization (Macedonia & Knösche 2011: 197; Xu & Ke 2014: 730); another reason is motor imagery, i.e. "kinetic representation of the word's semantics created through the action" (Macedonia & Knösche 2011: 196).

Going back to the description of 'Classes', in a typical lesson, activities are carefully scaffolded following a three-stage sequence, which in turn aligns with the P-P-P (Presentation-Practice-Production) model developed within the framework of the Communicative Language Teaching Approach. In the first stage ('Presentation'), the teacher, who is referred to as the 'guide' within the app, introduces the lesson's objective by providing essential phrases and vocabulary that learners would need to use in specific situations, such as making recommendations at a restaurant or discussing pros and cons of a new home. During the second stage ('Practice)', students are invited to use the new structures in a controlled environment. Depending on their level, these activities may take the form of a game, such as phrase-reordering, or pair-discussions as interaction, cooperation and negotiation are paramount in this stage. Finally, in the last stage

('Production'), learners engage in active interaction, by producing an original conversation, thus fulfilling Richards's (2006) 'Production' stage, where they use "the new structures in different contexts, often using their own content or information, in order to develop fluency with the new pattern" (8). For example, learners may be asked to simulate a conversation between customer and waiter at a restaurant, or discuss the advantages of cooking at home vs eating out. Throughout this process, the teacher intervenes as little as possible, encouraging students to practice the language, negotiate meaning, and learn from mistakes. In this way, students are encouraged to work independently, thus helping them experience agency, i.e. to experience the feeling of being in control or be in charge of their own choices and their own learning path (Taguchi 2022). Each class is 40-minutes long and a maximum of 12 students may join the group – a manageable group size that maximizes each learner's opportunity to speak and practice the language (one of the most frequently requested features, as seen in our own teaching experience).

As discussed at length, the app's primary focus is on speaking, with some emphasis on pronunciation, as teachers provide corrective feedback during activities. Exposure to the target language is consistently high, ensuring a productive language immersion experience. Finally, learners can monitor their progress using a pie chart displayed on the home environment floor. This chart visually represents the percentage of achievement in each macro-function (e.g. home and community, opinions and ideas, getting things done, expressing feelings, managing communication, health and safety, giving facts, socializing, life and living). Additionally, the same tracking tool shows the number of points required to earn a 'stamp', adding a gamified layer to the learning experience.

4.3 User experience

As regards UX, Immerse offers a fully immersive experience, as users can engage with people, interact with objects, and move through realistic environments. The sense of presence is even more pronounced when compared to Mondly VR, as users are immersed in virtual

scenarios with real people from all over the globe, which perhaps helps sustain a stronger sense of presence, countering any detachment that cartoonish designs of avatars and environments might otherwise cause. This feature highlights a key advantage of XR technologies: people who may not have the resources to travel or attend in-person courses can now learn foreign languages through means other than that of self-study books. In this sense, Immerse, along with other XR educational applications, serves as a valuable tool that promotes accessibility and inclusivity.

Interaction and interactivity are the backbone of Immerse, without which the app would lose much of its pedagogical purpose. Additionally, Immerse allows users to accumulate points whenever they successfully attend classes, events, and practice conversations with AI chatbots. Once they have enough points, learners are awarded a 'stamp', symbolizing their progress. Additional features that enhance the app's value include a help section and a 14-day trial, after which users are required to subscribe or choose a subscription plan. Registration is also required to use the app.

Our overall impression of Immerse is that it is a well-designed and thoughtfully developed application for language learning. The only drawback we were able to detect – which is also shared by Mondly VR – is the sense of motion sickness that may occur when using the sliding mode to navigate through the different environments. Unfortunately, this seems to be a relatively common issue with immersive VR platforms (see Chen et al. 2022). However, users can mitigate this issue by switching the navigation mode to teleporting.

5. Language Lab

In this final section, we will analyze Language Lab, a VRALL application developed by the Language Lab company, based in the US, primarily aimed at vocabulary acquisition. This application heavily relies on multisensory stimuli for language acquisition (i.e. visual

identification and gesturing), and reiteration and gaming for memorization.

5.1 User interface

The Language Lab interface is clear, uncluttered, and easy to navigate. However, in spite of its intuitive architecture, the app lacks a help section or tutorial to guide users on how to move around the environments and interact with objects, a feature that may be beneficial for inexperienced users. Unlike the previous applications, Language Lab offers three different modes of navigation:

- Smooth movement: comparable to the fluid motion in Immerse and Mondly VR.
- Teleporting: allowing users to jump to different areas.
- Tunnel vision: similar to smooth movement, but with reduced peripheral vision, possibly aimed at mitigating dizziness common in full 360° motion.

The app also features a dedicated game section, though the selection is currently limited to three options: Ultimate ascent, Hot shots, Sabers. These games reinforce vocabulary recently learned by requiring users to identify items named by a virtual instructor. In each game, learners practice retrieval by either shooting or striking the correct items, which reflects the pedagogical strategy of retrieval practice. An internet connection is required to use the app. However, performance issues occasionally occur, including crashes or freezes after periods of inactivity, which may be due to its developmental stage.

5.2 Pedagogy

Language Lab does not adhere to a specific pedagogical framework, as its main objective is vocabulary acquisition. The app primarily employs retrieval practice, a strategy designed to reinforce learners' ability to

memorize and retain new words, and understand simple phrases and sentences, such as “put the ball between the boxes” or “give the flower to the man”.

With respect to its architecture, Language Lab is divided into four sections: Vocab, Lessons, Open worlds / Events, and Games. The Open worlds / Events section, which involves interacting with peers, will be discussed in further detail in the next part, whereas the Games section has already been reviewed in Section 5.1.

In the Vocab section learners are prompted to grab and move items belonging to five semantic fields as instructed by the virtual teacher (e.g. “put the blender on the table”). Target items are marked with an exclamation point, an approach that offers both benefits and limitations. On the one hand, beginners are guided through the identification and selection process, thus minimizing frustration. On the other, however, this highlighting could hinder cognitive engagement if used excessively; ideally, this feature would appear only after several failed attempts, encouraging learners to exert more cognitive effort initially.

The Lessons section involves additional actions. Each lesson presents vocabulary from various semantic fields (kitchen and kitchen utensils, clothes, numbers, food, days of the week). Typically, new words are introduced by the virtual instructor who then prompts the learner to perform actions, involving one of the tools/items previously introduced, in a structured sequence such as: “Fan > light > I turn on the light > I turn on the fan”. Feedback is provided through brief sound cues that indicate whether the response is correct, though no further linguistic explanations are given – a feature developers might consider enhancing in future updates. Both the Vocab and Lesson sections incorporate movements and gestures, thus retaining the motor element featured in most VRALL applications, which has proven effective for language learning (see previous sections).

Overall, Language Lab is tailored for beginners or near-beginners, evidenced by the CEFR A1/A2-level vocabulary and

expressions used[6]. This focus may be the reason why section labels are consistently translated into the selected Source Language. Unlike Mondly VR, however, spoken instructions are given exclusively in the Target Language (TL, English, in our case), thus allowing learners to get used to the TL from the outset.

5.3 User experience

As regards UX – particularly the interactive element – Language Lab strikes a balance between Mondly VR, which lacks any interactive elements, and Immerse, whose foundation is in fact teacher-student interaction and peer interaction. Language Lab offers a single mode of interaction through the Open world / Event section. Here, users have the possibility to create their own events, such as "a dinner with friends", specifying the time and place the event will occur (the place being one of the open worlds available). This feature encourages peer-instruction, a student-centered approach where "Students had to become comfortable with listening to their peers in group work or pair work tasks, rather than relying on the teacher for a model" (Richards 2006). Arguably, the event planner should ideally be proficient in the language, even though this is not a prerequisite. Unlike Immerse, Language Lab does not involve any real instructors, which might question the effectiveness of the app's interactive features, as events lack supervision or guided support. The degree of interactivity offered aligns with that of the other apps discussed, although it tends to become repetitive over time. Instructions from virtual instructors are consistently repeated, which may aid memorization and retention but could also lead to user fatigue.

Language Lab offers an immersive experience comparable to that of Mondly VR. However, in some cases, the sense of presence may feel disrupted, especially when considering behavioral phenomena, one of the two core features of presence according to Slater & Wilbur (1997),

[6] To establish the target group of learners, we compared the vocabulary and phraseology used in the app with the General English syllabi used by the Language Center at the University of Bergamo.

which is defined as "the extent to which individuals behave in a VE similar to the way they would behave in analogous circumstances in everyday reality". In particular, the repetition of several unnatural actions that users are required to perform, such as "Put the cutting board in the bowl" or "Put the watch on the bed" strongly counteracts the suspension of disbelief that a fully immersive experience aims to provide.

As regards the other parameters, Language Lab does not allow outcomes to be shared on social media, and its badging system remains quite basic: users can enter the "Top login streaks" ranking when they access the app for at least two days in a row. Regarding pricing, Language Lab is completely free, with no pop-up ads or in-app purchases.

Overall, while Language Lab, like Mondly VR, has potential – particularly for vocabulary acquisition – it may benefit from some improvements, especially in the learning sections (Vocab and Lessons); in particular, it would be advisable to include exercises that teach vocabulary through realistic tasks, that reflect real-life communication scenarios.

6. Conclusions

This paper aimed to review the main VR applications currently available for ELT. For this purpose, existing taxonomies used to test and evaluate mobile-assisted language learning applications were compared and combined. The final taxonomy was then refined by incorporating parameters describing the type of pedagogical approach adopted, along with parameters defining specific VR features, such as immersion and sense of presence.

The analysis revealed that, among the three apps considered, only one – Immerse – seems to have been developed with a specific pedagogical approach in mind, namely a functional approach. Indeed, all activities primarily aim to develop users' speaking skills in specific

scenarios, while grammar and vocabulary teaching are corollary to this goal. Furthermore, Immerse enables users to interact live with people across the globe, thus adding to the authenticity of all simulated scenarios. In contrast, Mondly VR and Language Lab focus primarily on vocabulary acquisition. In particular, Mondly VR includes a series of activities based on different pedagogical frameworks, i.e. content-based and functional. However, some tasks intended to follow a functional framework do not always achieve their purpose; in some user-avatar interactions, the chatbots seem unable to generate appropriate responses when users deviate too far from the pre-set phraseology or pronunciation.

Overall, all three applications met the expectations regarding the degree of immersion and sense of presence, i.e. the two core characteristics of any proper VR-based app. Each app, to varying extents, allows users to immerse themselves in realistic environments, while at the same time encouraging them to engage and interact with objects and avatars, thus also fulfilling another core educational principle, i.e. embodiment. An interesting direction for future research in this sense could involve English for Specific Purposes (ESP) teaching, as no VR applications have yet been developed specifically for this purpose, to our knowledge. ESP, in our opinion, represents a promising area in the market, especially in tertiary education and professional contexts, where the demand for specialized language knowledge and field-specific communicative skills has been exponentially increasing – for instance in medical, business, and legal contexts.

In conclusion, Virtual Reality represents an innovative frontier of ELT that appears capable of addressing the core challenge of teaching authentic communicative skills. However, it is essential to treat virtual reality as a supplement to, rather than a substitute for, traditional teaching and learning methods. A blended-learning approach, combining VR's immersive and interactive advantages with conventional instruction can indeed lead to more effective outcomes.

References

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.

Chen, B., Wang, Y., & Wang, L. 2022. The Effects of Virtual Reality-Assisted Language Learning: A Meta-Analysis. *Sustainability*, *14*(6), 3147. https://doi.org/10.3390/su14063147.

Chinnery, G. M. 2006. *Going to the MALL: Mobile Assisted Language Learning*. http://hdl.handle.net/10125/44040.

Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J.-J. 2013. Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*, 63, 380–392. https://doi.org/10.1016/j.compedu.2012.12.020.

Hsu, C.-C., Chen, Y.-L., Lin, C.-Y., & Lien, W. 2022. Cognitive development, self-efficacy, and wearable technology use in a virtual reality language learning environment: A structural equation modeling analysis. *Current Psychology*, 41/3, 1618–1632. https://doi.org/10.1007/s12144-021-02252-y.

Kim, S., Song, K., Lockee, B., & Burton, J. 2018. *Gamification in Learning and Education*. Springer International Publishing. https://doi.org/10.1007/978-3-319-47283-6.

Kim, H.Y., & Kwon, Y.H. (2012). Exploring Smartphone Applications for Effective Mobile-Assisted Language Learning. *Multimedia-Assisted Language Learning*, 15/1, 31-57. https://doi.org/10.15702/MALL.2012.15.1.31.

Krahnke, K. 1987. *Approaches to Syllabus design for Foreign Language Teaching.* Washington, DC: Center for Applied Linguistics.

Lin, T. J., & Lan, Y. J. 2015. Language Learning in Virtual Reality Environments: Past, Present, and Future. *Educational & Society*, 18/4, 486–497.

Macedonia, M., & Knösche, T. R. 2011. Body in Mind: How Gestures Empower Foreign Language Learning. *Mind, Brain, and*

Education, 5/4, 196–211. https://doi.org/10.1111/j.1751-228X.2011.01129.x.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. 2019. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 60, 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007.

Nah, F. F.-H., Zeng, Q., Telaprolu, V. R., Ayyappa, A. P., & Eschenbrenner, B. 2014. Gamification of Education: A Review of Literature. In F. F.-H. Nah (Ed.), *HCI in Business*, 8527, 401–409). Springer International Publishing. https://doi.org/10.1007/978-3-319-07293-7_39.

Pinto, R. D., Peixoto, B., Melo, M., Cabral, L., & Bessa, M. 2021. Foreign Language Learning Gamification Using Virtual Reality—A Systematic Review of Empirical Research. *Education Sciences*, 11/5, 222. https://doi.org/10.3390/educsci11050222.

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

Riva, G. (2009). Is presence a technology issue? Some insights from cognitive sciences. *Virtual Reality*, 13/3, 159–169. https://doi.org/10.1007/s10055-009-0121-6.

Rosell-Aguilar, F. 2017. State of the App: A Taxonomy and Framework for Evaluating Language Learning Mobile Applications. *CALICO Journal*, 34/2. https://doi.org/10.1558/cj.27623.

Sato, M., & Ballinger, S. G. (Eds.). 2016. *Peer interaction and second language learning: pedagogical potential and research agenda*. Amsterdam: John Benjamins Publishing Company.

Schubert, T., Friedmann, F., & Regenbrecht, H. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments*, 10/3, 266–281. https://doi.org/10.1162/105474601300343603.

Sitzmann, T. 2011. A Meta-Analytic Examination of the Instructional Effectiveness of Computer-Based Simulation Games. *Personnel Psychology*, 64/2, 489–528. https://doi.org/10.1111/j.1744-6570.2011.01190.x.

Slater, M., Steed, A., McCarthy, J., & Maringelli, F. 1998. The Influence of Body Movement on Subjective Presence in Virtual

Environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40/3, 469–477. https://doi.org/10.1518/001872098779591368.

Slater, M., & Wilbur, S. 1997. A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence*, 6/6, 603–616.

Su, C., & Cheng, C. 2015. A mobile gamification learning system for improving the learning motivation and achievements. *Journal of Computer Assisted Learning*, 31/3, 268–286. https://doi.org/10.1111/jcal.12088.

Vesisenaho, M., Juntunen, M., Häkkinen, P., Pöysä-Tarhonen, J., Fagerlund, J., Miakush, I., & Parviainen, T. 2019. Virtual Reality in Education: Focus on the Role of Emotions and Physiological Reactivity. *Journal for Virtual Worlds Research*, 12/1. https://doi.org/10.4101/jvwr.v12i1.7329.

Xu, X., & Ke, F. 2014. From psychomotor to 'motorpsycho': learning through gestures with body sensory technologies. *Educational Technology Research and Development*, 62/6, 711–741. https://doi.org/10.1007/s11423-014-9351.