



## OPEN ACCESS

## EDITED BY

Luciana Cabral Pereira Bessa,  
University of Trás-os-Montes and Alto  
Douro, Portugal

## REVIEWED BY

A. Syahid Robbani,  
Ahmad Dahlan University, Indonesia  
Rafika Rabba Farah,  
Universitas Muhammadiyah Malang, Indonesia

## \*CORRESPONDENCE

Gökhan Haldun DEMİRDÖVEN  
✉ gokhandemirdoven@subu.edu.tr

RECEIVED 26 January 2025

ACCEPTED 14 April 2025

PUBLISHED 06 May 2025

## CITATION

DEMİRDÖVEN GH (2025) CEFR updates  
(2020)-based next-gen immersive learning in  
5 steps. *Front. Educ.* 10:1567249.  
doi: 10.3389/educ.2025.1567249

## COPYRIGHT

© 2025 DEMİRDÖVEN. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License \(CC  
BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in  
other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication  
in this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# CEFR updates (2020)-based next-gen immersive learning in 5 steps

Gökhan Haldun DEMİRDÖVEN \*

Rectorate/Turkish Language Coordination, Sakarya University of Applied Sciences, Serdivan, Türkiye

The Common European Framework of Reference for Languages (CEFR) can be defined as an action-oriented framework that systematically employs “can do” descriptors to structure the processes of foreign language teaching and learning. After a comprehensive literature review, it was evident that there was no descriptive content analysis study on the CEFR (2020) in terms of immersive learning technologies. Based on the existing shortcomings, the focus of this study was to identify the updates in the CEFR and the key elements that formed the connection points of immersive learning. This study, due to its scope and content, was conducted within the framework of qualitative research methodologies, employing document analysis and descriptive content analysis. Based on the research results, it is possible to assert that *the situational teaching method based on the digital/human digital twins; cognitive immersive language learning (CILL) approach rooted cognitive immersive rooms; interactive conversational agents capable of code switching; adaptive gamification approaches combining gamification techniques and educational data mining methods, accent-robots automated speech recognition systems grounded in the sociolinguistic approach, and online socialization metaverse networks woven with social virtual reality (SVR) and collaborative virtual environment (CVE) concepts built upon the phenomenon of heterotopia* stood out as the “paradigm identifiers” for the next-gen of foreign language teaching forms. As far as the research implications are concerned, the results affirmed the components and pedagogical approaches that would form the backbone for the integration of immersive technologies into CEFR-based language learning. Theoretically speaking, the CEFR-based immersive learning method was proposed for future studies in the age of AI-driven technologies.

## KEYWORDS

architectures for educational technology system, CEFR-based immersive learning, extended reality, immersive learning, simulations

## 1 Introduction

Over the last few years, the worldwide chaotic environment has generated a technological-awareness, and new innovative pedagogical approaches have been emphasized within the framework of the characteristic triggers of digitalization in educational sciences. Consequently, artificial intelligence-based virtual reality systems have captured people’s imagination, and the immersive learning method has become the new phenomenon of educational technologies. At this point, it will initially be appropriate to provide information about virtual/digital reality and artificial intelligence applications. It is simply because the anatomical structure of immersive learning contexts has been developing with the relevant infrastructure elements.

Virtual reality environments are the virtual form of a certain sample of physical world scopes, which can be perceived and interacted with our sense organs within a technology-based structured 3D simulation (Skulmowski and Rey, 2018). Within the framework of today's parameters, it is possible to classify the digital reality technologies under three sub-headings of virtual reality (VR), augmented reality (AR) and mixed reality (MR) (Lang and Müller, 2020). In addition to the relevant types of reality, the extended reality (XR) technology has been emphasized in recent years. XR, synthesizes the qualities of existing types of digital reality by blending them in the same pot. In this sense, as a unifying term, it is possible to say that it regards the digital reality contexts (VR, AR, and MR) as a subset by combining them within itself.

These new generation socialization zones have the infrastructure to activate the multiple sensory channels such as sight, touch, hearing, smell, and taste by being sensitive to the individual differences of students with different intelligence types; they also pave the way for the learning by the “doing-experiencing” method through self-activation. Therefore, it is possible to say that these environments are valuable in that they are able to form authentic atmospheres where students can interact with the target language.

On the other hand, it is possible to say that the advances in machine learning, deep learning, and natural language processing algorithms have constituted the basis of artificial intelligence technology. The artificial intelligence applications make reference to the system mechanisms designed on the functionality of displaying human-specific behaviors or performing actions that necessitate intelligence (Schooser, 2023). This system is built on the ability of a machine or device to meet higher mental functions such as problem solving, analytical thinking, cause/purpose-effect and part-whole relationship or evaluation/deduction.

Considering the rapidly developing technology, it is possible to integrate the artificial intelligence applications such as chatbots, digital twin, automated evaluation system, intelligent tutoring system etc. into the digital reality environments (Huang et al., 2022). These innovations make indistinct the line between the virtual and real, and the atmosphere turns into an immersive learning platform where the feeling of interaction can be recognized. In this context, it is possible to say that the immersive learning method was one of the new generation learning approaches.

Immersive learning is an educational method in which the real-world scenarios based on gamification elements are built on the artificial intelligence-based digital/virtual reality technologies. The learning environments structured in this framework offer digital teaching contexts patterned with fantastic fictions, where the individuals can gain experiences with the educational content. These contexts, on the other hand, immerse the individuals and create authentic scopes synthesized with possible environments and situations related to daily life.

One of the most significant problems encountered within the scope of foreign language teaching throughout the Covid 19 period was the lack or absence of authentic contexts in which students were exposed to the target language (Choi and Chung, 2021). The immersive learning technologies tend to blur the sharp line between real and virtual transmission networks. In other words, these systems create the opportunity to experience digital experiences

that are close to the original form. The learning contexts (Battal, 2022), equipped with the infrastructure compatible with these technologies are inevitable as a natural result of digital innovation, which tends to grow exponentially in terms of education reform efforts. It is simply because digitalization becomes an innate part of every field and spreads with its magic power that can transform every scope over which it has an impact. Therefore, it would not be incorrect to emphasize that the roadmap of foreign language teaching field will also move in the direction of the aforementioned “blur” trend. Clearly, the foundations of the near future have been in the process of being built on these environments. Based on the recent studies conducted (Takshara et al., 2025; see also Shruthi et al., 2025; Hwang and Lee, 2024; Takuchi and Hanks, 2024; Takshara et al., 2025; Gruber et al., 2023), it is possible to say that the framework for the field of foreign/second language teaching became tangible accordingly.

The CEFR, which calibrates the language skills that need to be acquired in foreign language teaching processes in terms of the illustrative descriptor scales, is a universal reference text for the discipline of language education for many countries beyond the borders of the European Union. The development of the CEFR by being fed on the immersive technologies will also generate a framework for the communicative competencies that the Beta generation is “anticipated to exhibit”. Given the relevant context, it is possible to assume that taking the CEFR (2020) updates into consideration in terms of immersive technologies will raise awareness regarding the new generation language learning environments.

## 2 Statement of the problem

The fact that CEFR, which has the potential to establish a framework for foreign language teaching practices in the form of a constitutional text, updates itself during the transition to the digital age is not coincidental but rather an indication of its systematic development course.

As far as the review of relevant literature is concerned, it is evident that there is no content analysis study on CEFR (2020) related to immersive learning environments. Considering this lack of study, the fact that the connection network between the CEFR and immersive learning contexts which has the power to identify the general framework of the foreign/second language education processes to be structured in the future, has not been clearly introduced, confronts us an inevitable necessity.

## 3 Purpose of the study

CEFR and immersive learning contexts are the key parameters with significant impact in shaping the next generation of foreign language teaching processes. Based on this perspective, the focus of this study is to examine the key elements that formed the connection points between the updates of the CEFR (2020) and immersive technologies. The target aimed here is to describe a framework regarding the elements that would constitute the cornerstones of the new generation

foreign language learning forms and raise awareness in the relevant context.

## 4 Significance of the study

The pioneering innovation in integrating web technologies into educational processes lies in its ability to provide infrastructure for real-time user interaction. The relevant technologies, thanks to this feature, have captured people's imagination in the field and have had almost a magical effect on shaping the phenomenon of "learning" related to Generation Z. This shift in focus has made collaborative learning, peer teaching, and other interaction-based methods the central driving force behind the new generation learning processes, emphasizing learning through a "collective intelligence" approach with the gamification elements.

The immersive learning method, with its realization based on social interaction in virtual reality environments and its structure built on a framework of gamification mechanics, exhibits the qualities of being the new phenomenon of digitization.

It is self-evident that by integrating the digital reality technologies into the foreign language education, it is possible to create an alternative to the resistance of the Generation Z, who grew up with digital-centered social web networks, against the traditional learning methods. In this sense, investigation of the anatomical structure of the updates regarding the CEFR in the light of the immersive learning environments will be a guide in terms of designing the foreign language education atmospheres of the future and create milestones for the new generation education paradigms.

## 5 Design of the study

The study, due to its scope and content, was conducted within the framework of qualitative research methodologies, employing both document analysis and descriptive content analysis.

### 5.1 Data collection and analysis

The results obtained within the framework of this study were attained by subjecting the illustrative descriptors updated in the CEFR (2020) to the descriptive content analysis within the center of immersive learning technologies.

In line with this, firstly, the updates to the CEFR Companion Volume were analyzed within the scope of document analysis, and primarily, the codes were identified for each updated illustrative descriptor and then similar ones from the identified codes were brought together and eventually the themes were created. Within the relevant framework, five sub-titles were categorized as "mediation", "plurilingualism-pluriculturalism", "creative texts and literature", "phonology", and "online interaction". During this process, the innovations related to the "Information Exchange" and "Using Telecommunications" illustrative descriptor scales were regarded within the "Online Interaction" theme as the overarching context. "Expressing a Personal Response to Creative Texts" and

"Analysis and Criticism of Creative Texts (Including Literature)" scales were discussed under the "Creative Texts and Literature" category and "Signing Competence", on the other hand, was not included in the analysis process due to the scope and content of the study.

In order to establish the reliability of the codes and themes formed throughout the study process, three researchers performed thematic content analysis of the CEFR (2020) updates. Subsequently, the codes and themes that were previously created were compared with those that had already been formed independently by three different researchers in a separate content analysis. The comparative reliability of the encoders was calculated based on the Miles and Huberman's (1994) formula [ $\text{Consensus}/(\text{Consensus} + \text{Dissensus}) \times 100$ ], and the result was 95.4%. Figure 1 illustrates a visual presentation of the relevant process.

As is clear in Figure 1, the initial phase of the data collection process and analysis consisted of the thematic classification of CEFR (2020) updates within the scope of document analysis.

In the continuation of the study, it was aimed to reveal the pattern network between the related theme elements and immersive technologies within the descriptive content analysis by positioning each theme item within the focus of the immersive learning method. In this process, the professional opinions of three foreign language teaching experts, three instructional technology experts and two information technology experts were consulted, and it was intended to ground the study on an interdisciplinary basis based on a multi-perspective approach. Consistent with this, forums with the participation of relevant experts were organized, and all analyzes were collected in a pool, and separate discussion sessions were organized for each theme item, and the data analysis, on which a general consensus was achieved, was finalized. At this point, it was aimed to structure all the items presented within the scope of data analysis around a common opinion, and in case of lack of consensus, it was aimed to present the different analysis styles of the experts individually. Finally, it was explicit in the process that structured data analyzes emerged with unanimity rather than majority vote, and it was established that a common analysis pool was created by achieving a general consensus on all theme items. Figure 2 illustrates a visual presentation of the curriculum line of research II.

As was emphasized in Figure 2, it is possible to say that the last stage of the data collection process and analysis was shaped by subjecting the findings obtained in the first stage to a descriptive content analysis procedure.

## 6 Findings

The understanding of basic language skills is presented with an innovative perspective in the CEFR under the categories of "communicative language activities and strategies", "plurilingual and pluricultural competence", and "communicative language competences". When the modernization processes carried out in CEFR (2001) are subjected to a descriptive content analysis regarding the immersive technologies, it is possible to describe the fundamental elements of the key points as follows:



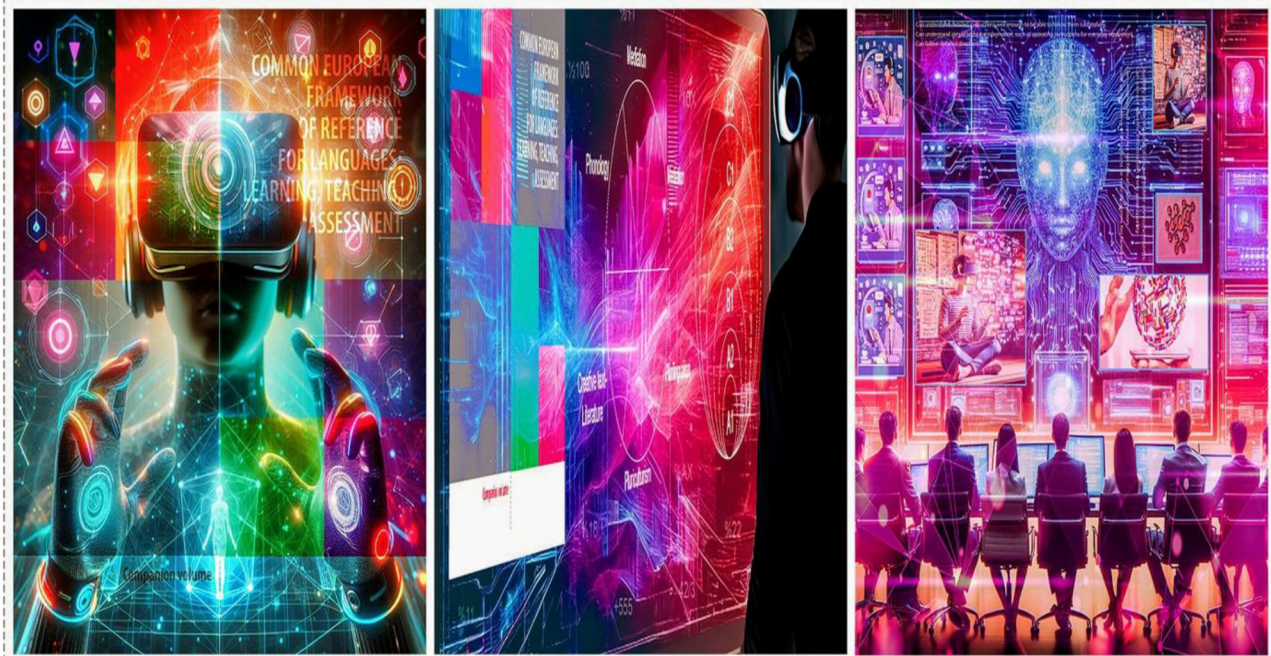


Document Analysis

Thematic Classification

Reliability of Data

FIGURE 1  
The process line of the research I.



CEFR at the  
Focus of Immersive Tech

Descriptive  
Content Analysis

Consensus  
Building Through Unanimity

FIGURE 2  
The process line of the research II.

1. It is possible to say that the most comprehensive first change regarding the CEFR (2020) updates emerged within the scope of “mediation” activities. The most important indicator of this particular situation is the inclusion of a total of 20 new exemplary descriptive scales within the scope of mediation activities (15) and strategies (5) in the Companion Volume. The mediation activities were associated with the context of mediating a written and spoken text in the first version of the CEFR, while in the latest version, the Companion Volume was expanded by associating with illustrative descriptors and illustrative descriptors scales competence to cover the fields of communication, learning, society, and culture as well as the relevant context. The main rationale for the related expansion policy is based on the basic philosophy of the content and language integrated learning (CLIL) approach.

The most important reference point for making content and language contexts related to each other regarding education is that the CLIL applications have the ability to offer learning atmospheres close to the authentic environments in native language acquisition (Kuteneva et al., 2021). The facts related to real life transpire when the components of being interdisciplinary, respecting the different perspectives and exhibiting a unique state of complexity become homogeneous under the same denominator. The fact that learning processes form a whole from the subset elements with similar transparency creates authentic environment frameworks. These frameworks are a reflection of the learning action, which includes the qualities that are synthesized with the possible environment and situations related to daily life.

In language education practices carried out in authentic environments, the learning process initially takes place subliminally, but as the process gets underway, it evolves into a conscious form by making sense of the language (Thompson and McKinley, 2018). The point that comes to the fore here is that foreign language education proceeds on the effort of constructing a meaning spontaneously and realistically. There is a similar systematic in the illustrative descriptors scales presented in the CEFR (2020) regarding the scope of mediation. Mediation means being able to redesign a similar or close meaning by forming the outputs indicating a certain event, phenomenon, or emotion, thought and situation with different language structures within the framework of the realization reasons of the communication. The main goal is to create fields and contexts for communication or learning, to cooperate to shape new worlds of meaning and to encourage other people along the relevant line, and to transfer new information within social, pedagogical, cultural, linguistic, and professional fields through the authentic channels.

When CLIL activities are considered in terms of “immersive learning”, it is possible state that digital twin (DT) technology provides the infrastructure to create a learning environment identical to the real-life contexts. DT technology refers to the interaction, communication and collaboration that transpire within the physical and cyber environment planes. It is also possible to contemplate this type of technology (Zhuang et al., 2017) as a simulation system based on mapping in which real-world elements are reconstructed in cyberspace. Strictly speaking, it is

possible to structure virtual identical objects and environments with DT technology (Jaensch et al., 2018). There exist two types of models of DT systems: three-dimensional and five-dimensional. While the three-dimensional DT systems are built on three denominators: cyberspace, physical environment and connection, in five-dimensional DT systems, the service and data components also become operative (Tao et al., 2019). Therefore, when there are new updates available in the physical equivalent, these innovations are simultaneously reflected in the five-dimensional DTs and a dynamic simulation process becomes apparent.

Another perspective that can mold the future of DT technology is the concept of “human digital twin” (HDT). This particular concept refers to the situation in which various tasks and responsibilities undertaken by people are implemented through their digital twins (Gräßler and Pöhler, 2017). Essentially, what is aimed here is to improve the HDTs’ ability to make autonomous decisions by modeling human skills through DT technology. The point to be highlighted here (Ansari et al., 2018) emerges at the point of coping with uncertain human behavior. Therefore, it is possible to say that the solution suggestions for the relevant problem situation can be generated by working on algorithms focused on tolerating the uncertainty. It is possible to take into consideration the “Intelligent agent” technology amidst these solution suggestions. Intelligent agent is a type of software developed to interact with people based on artificial intelligence techniques (Titova and Temuryan, 2024). It is commonly recognized that the DT systems supported by intelligent agent technology exhibit high performance in terms of analysis and decision-making capabilities (Sun et al., 2024). As far as this relevant context is concerned, it is possible to anticipate that the intelligent agent based HDTs will have a stronger mechanism for uncertain human behavior.

Given the characteristic structural features of DTs, it is possible to say that this type of technology become more visible with its potential to structure the digital equivalents of physical world contexts that have turned into authentic learning environments (Han et al., 2022). This type of technology eliminates time and spatial limitations and creates atmospheres in which the individual drifts along. Moreover, virtual forms can be shaped instead of high-cost content that can be encountered in authentic environments of daily life, thanks to DT technology. In other words, it is possible to say that immersive DT technology comes into play in the common denominators where the contexts that are not suitable for experiencing real-world atmospheres for economic reasons cannot be shaped with the modern technology.

In the new generation DT technology, where virtuality and reality coexist together beyond intertwining, the user has the opportunity to experience all kinds of real-life experiences in a one-to-one virtual representation of physical elements, by interacting with both HDTs and other users. As far as the relevant framework is concerned, it is possible to say that the DT immersive technology can offer authentic learning environments to the discipline of foreign language teaching in shaping the environments that can surround the implementation of mediation activities. The point to note here is that the information to be conveyed is presented on a task-based basis in a way that corresponds to real-life situations in the form of the smallest decomposable particle (*quanta*) without



being simplified or divided into smaller parts, as in the illustrative descriptor language competencies.

2. The innovation with the most cognitive characteristics regarding the CEFR (2020) updates was implemented within the scope of “plurilingual and pluricultural competence”. In today’s language education approaches, being able to develop a plurilingual and pluricultural proficiency rather than having a multilingual and multicultural infrastructure is considered more valuable. As far as CEFR (2020) is concerned, it is explicit that the scales related to the relevant context were expanded within the framework of the competencies of “*being able to blend different languages when necessary*”, “*being able to switch flexibly between languages*” and “*being able to exhibit intercultural competence skills*”. At this level, it is possible to say that the illustrative descriptor scales of “building on pluricultural repertoire”, “plurilingual comprehension” and “building on plurilingual repertoire” to the Companion Volume. The expansion in question was supported by the preference of “target language speakers” in the current illustrative descriptors instead of “native speakers” used in illustrative descriptor proficiency descriptions, especially on the B2, C1 and C2 common reference levels of the CEFR (2020) version. The relevant changes were valuable in that they clearly emphasized that the phenomenon of plurilingualism-culturalism was prioritized in the CEFR, which was designed with contemporary language competencies and that the “new world” language competencies should be addressed within the relevant context.

The phenomenon of plurilingualism represents the communicative development of a wide spectrum, from the language spoken in the family, to the language or languages that are commonly used in the social environment. These languages are intertwined with action-oriented skills that display a mosaic integrity together, rather than forming independent meaning systems within the cognitive framework. A specific skill area for a particular language may have a variable of proficiency in other languages or languages, with independent proficiency levels. Therefore, it is possible to define the scope of the “general linguistic range” of plurilingual individuals as a cumulative field of existence where different communication competencies of various languages create harmony in a similar fashion.

Linguistic and cultural competences related to a language develop to the extent that it grows out of the background of different languages and contributes to intercultural awareness CEFR (2020). In pluricultural perspectives, the cultural phenomenon acknowledges the concept of “identity” as the outcome of a complex, flexible, and dynamic combination in an attempt to adopt the perspective of the interlocutor (Kharkhurin et al., 2023). It becomes explicit here that there is a single cultural system that is shaped by providing resources from different contexts in the phenomenon of plurilingualism, as well as in the phenomenon of pluriculturalism. The difference between pluriculturalism from multiculturalism is related to its contextual nature. In the case of multiculturalism, the individual interacts with

various cultures as a social language user, while in pluriculturalism, the individual, beyond the interaction, synthesizes the mentioned cultures and creates a composition.

Perhaps the most important legacy of war, famine, natural disasters etc. is that large groups of people are relocated by changing geographies and thus this plays a triggering role in intercultural interaction. Global migration fluctuations in the post-World War II period caused societies to evolve into a bilingual structure (Bezioglu-Goktolga and Yagmur, 2022). This social structure tends to transform into plurilingual and pluricultural organic forms in today’s world, as an aspect left behind by the digital age. In fact, this transformation form is at the forefront of the elements that serve as a cornerstone related to the design of immersive learning environments. In the plurilingual community approach, the individual’s first language experience through his or her family is enriched in a perspective that expands from other languages spoken in the immediate environment to the other languages spoken in the distant environment. Communication skills, which are expected to be acquired by the new generation, the Generation Z individuals also take an angle on this issue. This situation brings about the need to be a world citizen and creates a prerequisite for the future generations of people to be educated in this way. It is a matter of debate whether most social structures around the world do not have the linguistic and cultural diversity that can meet this prerequisite, and whether those with pluri-layered social structures have the necessary understanding and tolerance relationship. This shortcoming constitutes an important obstacle for humanity to attain the form in which it is expected to evolve.

Here at this point, it is possible to assert that the XR technology comes out with its infrastructure qualities that can generate multicultural society structures and AI technology that can structure the plurilingual chatbots.

Digital reality is a type of technology that makes it possible to experiment authentic experiences based on the virtual contexts that include technology-based interaction and a sense of reality through a simulation method (Shadiev et al., 2020). Virtual reality, the first type of this technology, surrounds the individual with digital objects and spaces. The relevant state of immersion is equipped with an infrastructure in which virtual objects can be positioned in physical world contexts through the augmented reality systems. In mixed reality technology, in addition to placing virtual objects in physical environments, users have the opportunity to interact with these objects in real time. Moreover, different users who have permission to access the system can also participate in this interaction. XR (Rauschnabel et al., 2022), on the other hand, is a type of technology that synthesizes the qualities of existing types of virtual reality by blending them in the same pot. In this matter, it is possible to say that as a roof term, VR is regarded as a subset of immersive technologies by combining the feature structures that are emphasized within AR and MR. Furthermore, any existing (haptic, ambisonics, foveated rendering, etc.) or possible future creation (6G wireless, light field display, neural reality etc.) immersive technology components that can change our perception of reality are also accepted within the boundaries of XR. The digital environments dressed with multicultural elements designed with the XR technology can enable the individual to empathize with the differences of target cultures (Wu et al., 2019);

and, while, on the one hand, they pave the way for cultural diversity to be accepted as “valuable” as richness by activating the critical thinking skills, and on the other hand, they can contribute to the internalization of the phenomenon of “pluriculturalism” by triggering the individuals to imagine their possible future roles in the target cultural structures (Mills et al., 2020).

Artificial intelligence (AI) is defined as the capability of intelligent systems to perform human-specific functions such as problem-solving (Tillman and Louwerse, 2018), analytical thinking (González-Cacho and Abbas, 2022), understanding cause-and-effect relationships (Guyon et al., 2019), grasping part-whole relationships (Hinton, 2021), and making evaluations (Link et al., 2020), all based on machine learning (ML) and deep learning (DL) algorithms. Natural language processing (NLP) is an artificial intelligence technique that contributes to the ability of digital system structures to achieve the competence understand human language through specialized coding. The type of technology that is shaped at the intersection of AI components such as NLP, ML, and DL is referred to as a conversational agent (Jiang et al., 2022). Conversational agents are the software programs that perform to structure an interaction simulation between people and virtual agents within the scope of verbal or written communication skills (Laranjo et al., 2018). These dialogic interactions implemented through Smartbots generate an infrastructure for the enhancing of foreign language teaching processes in the sociocultural context, allowing students to experience an immersive experience in virtual learning environments (Zhang and Huang, 2024). As far as the experimental research in this day as age is concerned, it is evident that conversational agents can be supported with the equipment that can provide services in a plurilingual manner (Doshi and Bisht, 2021). Moreover, the voicebot technology, which has recently attracted a lot of attention, is noteworthy of attention also in terms of providing service as an individual coaching within the scope of language skills (Yu and Nazir, 2021).

As described above, “extended reality environments in the real world” can be shaped by piecing together unique elements that create new generation layers, relationships, and meanings through the new generation immersive technologies. As far as the relevant framework is concerned, it is possible to that within the scope of artificial intelligence-based immersive learning atmospheres, which are predicted to create the foreign language learning atmospheres of the future, the basis for the reconstruction of educational structures that are in the shadow of the past regarding social life can be generated. Based on the context in question, the new generation smart teaching systems, based on an action-oriented approach, can be structured by supporting elements such as extended reality learning environments patterned with plural cultural components, plurilingual chat robots and avatars and the seeds of the foreign language education environments needed by the future human generations can be sown. Categorically, the fact that language learners are considered as “social agents” who perform tasks within a certain field of action regarding certain environments under certain conditions within the scope of the CEFR supports this particular idea.

3. The most adaptable and flexible innovation within the framework of CEFR (2020) is the inclusion of new illustrative

descriptor scales (Expressing a Personal Response to Creative Texts, Analysis and Criticism of Creative Texts [Including Literature], Reading as a Leisure Activity) related to the scope of creative texts and literature. the activities to be implemented based on the action-oriented approach serve to acquire the target skill to the extent that they are patterned through creative text contexts that have not been subjected to editorial processes such as simplification or correction. it is simply because the language competencies described in illustrative descriptors are structured within the framework of real communication situations and aim to trigger the creative intelligence of the individual against the possible problem situations in life by reflecting himself/herself rather than a simulation of life. therefore, it is of great importance to support the material samples that will be used in the process with creative text products that are equally aesthetic with the texts presented in newspapers, magazines, television, blogs etc., and that the student will be drawn into them and get excited at the same time. furthermore, the phenomenon of creative text discussed within CEFR was addressed in a more comprehensive perspective rather than being associated only with literature or scenario texts, and the elements that required high imagination, such as movies, multi-modal installations and games, were also incorporated in creative text types.

The target competencies that are desired to be structured in the individual are processed into the cognitive level of the student within the scope of text analysis activities that vary according to the target audience. These activities take shape in an effort to reach a meaning based on communicative skills rather than reaching grammatical competencies. The effort to attain this meaning is carried out through the creative texts that are structured in authenticity in accordance with the immersive learning context in question, on the basis of vocabulary elements, grammatical structures etc., and elements suitable for the relevant language level. These texts have the potential to elicit reactions that activate the individual’s emotions, thoughts and imagination. This tendency to trigger a reaction is categorized under four basic subheadings in the CEFR (2020): engagement, interpretation, analysis and evaluation. Engagement refers to the state of being drawn to the appeal of a particular context or character in the context of the creative text; interpretation refers to attributing meaning to the context of the creative text; analysis refers to making analyzes from various angles regarding the context of the creative text; and evaluation refers to the state of having a critical perspective on the context of the creative text.

Immersive learning environments can deepen in an inclusiveness equivalent to the creative text contexts, patterning motivation and purpose channels for learning and creating a resultant formula with quality structures of sophistication that are attractive to the students and worth exploring. Here, the formation that can make the intersection of virtual reality environments and creative text contents dynamic, the elements of gamification, which is the phenomenon on which the immersive learning method has been structured, is noteworthy.

Gamification is a structure in which intellectual processes are triggered by the game mechanics and serve the individual’s

learning and make problem-solving skills functional in the relevant direction (Altomari et al., 2023). When we consider this structure within the scope of personal affective, intellectual and imaginary triggers, we encounter the categories of “character identification (engagement)”, “meaningful engagement (interpretation)” and “critical thinking skills (analysis and evaluation)”, which are the gamification version triggers of the creative text phenomenon.

Given this formation within the scope of personal affective, intellectual and imaginary triggers, it is possible to say that the categories of “character identification (engagement)”, “meaningful engagement (interpretation)” and “critical thinking skills (analysis and evaluation)”, which are the gamification version triggers of the creative text phenomenon, transpire. Character identification can be associated with the fact that the players incorporate the interesting features of game characters into their perception of themselves (Sierra Rativa et al., 2020); meaningful engagement with the degree to which players find the context they are in meaningful (Suh et al., 2017); and critical thinking skills with the development of metacognitive skills such as analysis, evaluation and inference through gamification mechanics (Rivas et al., 2022).

The activities to be designed on gamification mechanics will both turn the learning theme into an attractive one and bring along creative scenario texts that are compatible with the “game” logic. The important criterion here is that gamification and game-like interface applications should not be confused with each other. It is because, prioritizing learning in gamification applications is the main goal, and these applications can appeal to all age groups, while they have the infrastructure to be developed for many sectors (Krath et al., 2021). On the other hand, game-like interface designs are based on creating a certain power of influence by reviving emotion or sensitivity rather than bringing rationality to the fore (Gupta and Goyal, 2022). It is, therefore, possible to say that the relevant point of distinction will significantly affect the course of cognitive structuring mechanics (Lamb et al., 2018). Therefore, the extended reality environments can be blended with the philosophy of “gamification” instead of game-based atmospheres and can evolve into atmospheres where they turn into authentic learning contexts through the creative scenario texts. Thus, the metacognitive skills of the new Generation Z can be activated through the immersive language teaching environments harmonized with the CEFR.

4. When the CEFR (2020) is examined specifically for “linguistic competence”, it is possible to say that the most radical to the scope of illustrative descriptor scales presented were performed for the “phonological control” scale. In the CEFR (2001), there were 6 illustrative descriptors in total, with an illustrative descriptor at each level, with the same competency at c2 and c1 common reference levels within the relevant illustrative descriptor scale. In the CEFR CV (2020), on the other hand, the structure of this illustrative descriptor scale was discussed under 3 sub-headings: “Overall Phonological control”, “Sound articulation” and “Prosodic features”. 6 under the Overall Phonological Control subheading (a1-c2 common reference levels), 9 under the Sound Articulation subheading; (between a1-c2 common reference levels) and 8 (between a1-c2 common reference levels) in the subtitle

of Prosodic Features, a total of 23 illustrative descriptors was included. when the relevant changes were examined by focusing on the content structures, it was concluded that the illustrative phonological descriptors were expanded to include new and reflective thoughts on “intelligibility”, “influence of other languages spoken by the individual” and “internal features (emphasis, rhythm, intonation, etc.)”.

Linguistic competence is a system of competences that puts the coordination relations between the language components into an organized form by coordinating them within the framework of a certain functioning mechanism. Phonology deals with the oral production dimension of this systemic assembly and takes shape on the relevant plane (Soruc, 2016). The main purpose of the renewal processes for the Phonological Control illustrative descriptor scale is related to the degree of clarity and lucidity in the pronunciation of phonemes and the context of displaying the rhythm skills. In other words, it is possible to say that in the CEFR (2020), a speaker’s ability to bring certain messages to the forefront through accentuation, intonation prosodic qualities and to what extent a speaker could intonate the existing sound presence elements were highlighted. In this context, we encounter the phenomenon of “intelligibility” and it is explicit that in the latest version of the CEFR, the focus was on the competence of constructing meaning phonologically rather than idealizing the foreign/second language learner like a native speaker.

An individual’s ability to speak a foreign/second language is closely linked to his/her ability to generate production patterns of the language in question, as well as to recognize speech sounds specific to the relevant language. Considering the field studies (Hwang et al., 2025; see also Stankova et al., 2022; Senyigit and Okur, 2019), it was collectively concluded that even though people learning a new language could demonstrate acceptable skill in recognizing the sounds in a foreign language, they failed to exhibit the same performance in terms of reproducing these sounds with similar competence as target language speakers. Therefore, it is possible to say that the impact of the sound patterns components (sounds, rules, stress, intonation, etc.) in the individual’s native/first language on the pronunciation in the target language delivers with it the concept of “accent” as an inevitable phenomenon (Dong et al., 2022). Consequently, it is possible to describe accent as a concept that refers to how a particular word or group of words, depending on the individual’s native/first language, is pronounced by a foreign language speaker.

The field of computational phonetics, which has emerged as the intersection of analysis methods in the fields of computational linguistics and acoustic phonetics, with the natural language processing process being considered in relation to the information technologies deals with the echo science phenomena such as measuring the pitch unit of sound, energy value, articulation time and vibration frequency quantitative properties mathematically by developing algorithm-based analysis forms.

In today’s technology infrastructure, the speech recognition technology (SRT) has emerged at the synthesis point of various disciplines such as computational linguistics, signal processing, pattern recognition, information and communication technologies (Delić et al., 2019), and the artificial intelligence robots have



been made sensitive to the sound sensors by producing software based on the analysis of vibrations through the sound processing techniques (Lin et al., 2024). Nevertheless, with the integration of technical features of SRT and various dictation applications into the artificial intelligence systems through machine learning, automatic speech recognition (ASR) has come to exist (Lai and Chen, 2022). Therefore, it is possible to consider ASR technology as *a machine learning-based process in which verbal communication data is analyzed*.

The traditional ASR applications are composed of functionalizing independently trained acoustic, pronunciation and language models to form a whole (Chiu et al., 2018). The fact that speech accents create phonetic and linguistic diversity give rise to traditional ASR systems to generalize poorly, and this problem generate significant difficulties for the usefulness of ASR applications (Koencke et al., 2020) and lead to a serious intolerance problem (Choe et al., 2022). With the integration of the deep learning method into the automatic speech recognition (ASR) technology, the infrastructure composition of new generation ASR systems based on end-to-end modeling has changed direction and “accent-robust ASR” applications have commenced to become delineated. The stakeholders who have been trained independently in traditional ASR systems are trained together in the end-to-end ASR modeling (Prasad and Jyothi, 2020). In other words, it is possible to say that they are blended in the same pot. The components trained on the common axis refer to a single grapheme system. Therefore, the need for a pronunciation dictionary is essentially eliminated. Consequently, the difficulties posed by accent diversity regarding ASR technology cease to be a problem.

By adapting the accent-robust ASR applications to the extended reality platforms, the phonological variation can be easily included within the foreign language education processes and access to the CEFR-compliant activities structuring the infrastructure can be achieved by focusing on the semantic reference directly without causing problems with the accent of the individual. Moreover, it is also possible to perform mutual interaction between the human beings and artificial intelligence by creating sound files through PRAAT etc., software. As a result of the integration of these software applications with the digital reality technologies, the possibility of instant interaction between the avatars and the user stands out as one of the most remarkable infrastructure features of the immersive learning environments. The fact that the related type of interaction can be materialized in oral and written forms through artificial intelligence applications makes the usability of reality technologies in foreign language education even more attractive.

5. When the CEFR (2020) is examined in the light of today’s technological developments, it is noticeable that the “online interaction skills” were considered within the scope of language competencies as the most innovative update. these competencies were introduced in the illustrative descriptor scale and presented to the users under the categories of “Online Conversation and discussion” and “goal-Oriented Online Transactions and collaboration”. Both scales consisted of online competencies covering signaling, written and

verbal interaction skills within digital platforms. while the Online Conversation and Discussion scale focused on the digital competencies of being able to perform synchronous-asynchronous interaction in online environments, interacting continuously with one or more interlocutors, generating posts in online environments and commenting on the existing posts etc., the Goal-Oriented Online Transactions and Collaboration scale, on the other hand, addressed the digital interaction skills such as purchasing a certain good or service through online environments, collaborating for professional purposes regarding a business or project, and dealing with issues focused on communication problems.

The easy internet access since the beginning of the 21st century has been characterized by the use of online technologies. The online technologies, on the other hand, has brought about the digital platform phenomenon and thus, the digital living spaces related to social life have begun to transpire. Principally the quarantine process following Covid 19 created a trend toward virtual platforms, and the process of humanity’s integration into immersive atmospheres has accelerated in a revolutionary manner. Under these circumstances, there has been a transition from the traditional communication skills to the online communication skills, and new generation interaction competencies have begun to take shape accordingly. It is possible to consider the online interaction illustrative descriptor scales presented in the CEFR (2020) as a concrete indicator of this situation. Considering the online interaction illustrative descriptor scales based on immersive learning, it is possible to say that social virtual reality (SVR) and collaborative virtual environment (CVE) concepts significantly transpire.

VR technology, as an emerging alternative to the real world, is able to create an infrastructure for people to meet in a virtual environment and interact with each other’s virtual representatives and present the phenomenon of socialization in the form of immersive experiences (Heidicker et al., 2017). These types of VR interfaces are referred to as the SVR interfaces since they can enable online social interaction (Moustafa and Steed, 2018). The SVR allows users to create avatars that represent themselves in immersive environments. The body movements of the users are simultaneously mirrored in the body movements of the avatar (Tang et al., 2023). Therefore, through body tracking, physical touch and holding actions are transferred onto the VR environments and an easier transition ground based on bodily experiences is formed to initiate social connection (Tanenbaum et al., 2020). Nevertheless, the avatars that generate an interface between the user and their digital identity provide participants with the flexibility of selective self-presentation and anonymity, thus assisting the shy and introverted people or marginalized participants to socialize. Consequently, it is possible to emphasize that the interactions via SVR offer a new way of mediating and supporting interpersonal relationships (Freeman and Acena, 2021). When we calibrate the VR technologies in terms of collaboration, it would be appropriate to highlight the CVE concept. This concept principally refers to the immersive contexts designed to organize collaborative activities (Poppe et al., 2017). These contexts generate a digital landscape based on interaction where multiple participants

can perform various tasks collaboratively (Khojasteh and Won, 2021) and provide the opportunity to “doing it there and then” by creating a feeling of “existing” within the work zone (Massey et al., 2024). The team members stationed in different geographies have the opportunity to access the same digital living space and thus can collaborate smoothly (Khalid et al., 2021).

The VR concepts, when considered within the framework of modern language teaching approaches, offer great support to the users in overcoming the prejudices and developing empathy by supporting the intercultural and transcultural learning activities (Eisenmann and Steinbock, 2024). This support based on social interaction also enables collaborative language learning and generates a sample of an unlimited immersive platform for group studies by eliminating physical restrictions (Zhao et al., 2024).

## 7 Discussion, conclusion, and recommendations

CEFR is a system structure built on competence descriptions. This system is based on a mechanism where the common reference levels are supported by illustrative descriptors. The main backbone of the system consists of illustrative descriptors scales. The functioning of the mechanism, on the other hand, operates with a subset relationship. Each of the common reference level is a subset of the illustrative descriptors scales. The illustrative descriptors' competence descriptions also constitute the unit elements of these subsets. Therefore, the sum of the unit elements of a particular cluster portrays the competencies associated with the relevant common reference level within the scope of the illustrative descriptor scale. When the intersection clusters for each scale are combined, the general framework for the relevant common reference level transpires. With the latest updates, a transparent form has been attributed to the transition between the common reference levels. The descriptive qualities of a higher language level primarily begin to manifest themselves at the plus levels of a lower language level. The same is true for the reverse case as well. The relevant transparency is due to the fact that the CEFR has a more flexible framework. The CEFR-related fiction can be renewed according to the conditions of the age and the descriptors of competence can be shaped in line with the needs of the modern life. The changes presented in the CEFR (2020) can be regarded as a concrete indication that a modernization update has been implemented for the next generation interaction competencies. As a matter of fact, considering the current conditions where traditional language teaching components are being replaced by innovative ones (Chun et al., 2016), we can regard the CEFR updates not as a coincidence, but as essential guiding parameters for digital language education.

As far as the emerging findings are concerned, it is evident that the most comprehensive change regarding the CEFR (2020) updates was constructed within the framework of “mediation activities” based on the CLIL approach. When the mediation activities were considered within the scope of the immersive learning method, it is possible to say that the DT technology clearly transpired. DT is an exceptional type of technology in that it offers the opportunity to acquire original learning experiences by simulating identical contexts that can be encountered in

real-world environments (Madni et al., 2019). Digital copies of physical elements such as language laboratories or authentic learning environments can be generated with the DT technology. Real life scenes can be transformed into language learning scenarios through these virtual replicas, and hence the immersive learning environments can be deemed as authentic learning contexts (Damaševičius and Zailskaite-Jakšte, 2024). In fact, by using the extended reality and computer image technologies, it is possible to create a teaching model based on digital twin technology within the framework of the situational teaching method and transform the language education processes into a three-dimensional form within the scope of a spatial learning experience and sense of immersion. Furthermore, thanks to the ability of artificial intelligence-based HDT technology to create various scenarios and contexts through decision-making and automatic applications (She et al., 2023), it is possible to predict that the new generation immersive learning environments can be constructed and hence revolutionary pioneering innovations in foreign language education can be generated.

Given the relevant findings, the most cognitive characteristic change regarding the CEFR (2020) updates stands out as “*plurilingualism-pluriculturalism competencies*”. The plurilingualism-pluriculturalism competencies point to the mental structure and cultural texture in which the individual is expected to evolve and refer to the next generation of humanity being shaped in terms of language-culture awareness rather than language-culture difference, mosaicking diversity within itself. When the plurilingualism-pluriculturalism competencies is considered within the scope of the immersive learning method, it is possible to say that the Cognitive Immersive Language Learning (CILL) approach based on the combination of XR and the AI technologies clearly transpires. The most important innovation that this approach brings to the field of foreign language education emerges as the Cognitive Immersive Room (CIR). It is possible to consider CIR as interactive immersive learning environments designed at the human scale, where multiple people can simultaneously experience a dynamic language learning process through multimodal input (Divekar et al., 2019). These environments have the capacity to offer a more functional learning atmosphere beyond the traditional classroom atmospheres for authentic language and culture education (Chabot et al., 2020). In the XR atmospheres, the individuals who experience a “*feeling of being within it*” traditional usage environments specific to the historical texture of the target language can internalize the phenomenon of cultural absorption regarding different languages (Divekar et al., 2022) and thus has the opportunity to develop pluriculturalism competencies. Nevertheless, when the CLIL is considered within the scope of plurilingual competences, it would be appropriate to emphasize the AI powered conversational agent technology. The conversational agents have the potential to contribute to plurilingual education by creating interaction-based language learning opportunities (Gruzdeva et al., 2024), but when the applications are analyzed, it is clear that they can function and serve for different languages provided that the language settings are changed. Therefore, it is possible to say that these applications do not have the ability to automatically detect the input language and can only interact with the selected language (Doshi and Bisht, 2021). An important aspect of plurilingual competence is the ability

to switch between more than one language during interaction, in other words, the ability to code-switch (Gotti, 2015). Therefore, processing and recognition functions of interaction data involving code switching can be considered as a prerequisite for intelligent automated systems that can interact with plurilingual individuals (Singh et al., 2024). Therefore, it can safely be emphasized that the combination of language identification algorithms and context-based heuristics approach can be considered as a solution proposal to improve the output quality of interactive conversational AI Agents structured based on code switching (Kann, 2022).

The competencies related to the scope of *creative texts and literature* are noteworthy as the changes with the “most adaptable” characteristic within the framework of CEFR (2020) updates. It is simply because CEFR (2020) defined the phenomenon of “creative text” as films, plays, recitals, etc. rather than a perspective limited only to literature. It considered the variables from a broad perspective and thus provided a flexible application basis for illustrative descriptors related to creative text and literature. As far as the relevant illustrative descriptor group based on the immersive learning method are concerned, it is possible to say that we encountered the adaptive gamification approach. It is possible to consider adaptive gamification as a learning approach based on the synthesis of adaptive e-learning components and gamification mechanics (Hassan et al., 2021). The adaptive e-learning was built on the principles of conveying the right content to the right person in the most appropriate time frame and with the most appropriate presentation format, by considering the individual learning processes in relation to artificial intelligence. Learning algorithms make intelligent predictions about individual characteristics and needs by processing the data elements such as student profiles and performance etc. within educational data mining (EDM) methods (Diren and Horzum, 2024). Considered as a type of creative text, it was evident that the gamified learning environments fostered the creative thinking and problem-solving skills (Mee Mee et al., 2020); identification with the player character promoted intrinsic motivation and class participation (Thanyawatpokin and Vollmer, 2022); and meaningful gamified learning activities positively triggered the interest in the context (Liu et al., 2024) and aroused reactions in students with their power to activate the emotions, thoughts and imagination of students involved in the foreign language education process. Given that the gamification elements in question have the effect of creating a “sparking a fire” that varies from person to person, it is possible to say that adaptive gamification approaches, in which the gamification technique and the EDM are considered together, act as a compass in establishing the trigger mechanisms suitable for the person.

It is possible to consider the *Phonological competencies* as the “most radical” change regarding the CEFR (2020) updates. It was evident that emphasizing the concepts of accent and accuracy instead of intelligibility in teaching pronunciation failed to meet the needs of the individuals and consequently, the “Phonological Control” illustrative descriptor scale was completely redeveloped from scratch in the Companion Volume. Considering the phonological competencies within the scope of the immersive learning method, it is possible to say that the idea of integrating the ASR systems based on sociolinguistic approach into the extended reality environments clearly transpires. Accent studies

on Language 2 (L2) focus on the effects of language 1 (L1)’s prosodic features such as phonetics, stress, pronunciation and speech patterns (Ghorbani and Hansen, 2024). Therefore, either the ASR applications built on the “multi-accent” acoustic model that can process all accents in the training set or ASR models that can perform automatic accent identification are required (Najafian and Russel, 2020). The ASR applications developed from a sociolinguistic perspective explore the human-computer interaction within a sociocontextual framework, allowing individuals with different language backgrounds to implement a dynamic interaction process with the conversation agents (Lee and Jeon, 2023). Here, the “accent-robust” technology can be regarded as a pragmatic solution in terms of providing personalized feedback and allowing the selection of customized training materials. By utilizing the advanced natural language processing technologies, the artificial intelligence chat robots (Zhang and Lin, 2022) that can analyze the linguistic, cultural and identity backgrounds of the speaker can be combined with the accent-robust technology, and a contribution can be offered to the foreign language education processes in the immersive learning environments within the scope of *phonological competencies*.

It is feasible to describe the online interaction competencies as the most innovative update regarding the new generation of socialization or collaborative learning areas regarding the CEFR (2020) updates. Therefore, it is possible to say that a real parallel universe based on VR technologies has begun to surround us. This universe, which is referred to as the Metaverse, can be deemed as a concept that offers immersive experiences through thematically interrelated digital platforms (Seidel et al., 2022). This concept includes many activities such as creating new generation societies as well as being a part of them (Zhao et al., 2022) and accessing a higher level of interaction and immersion with the launch of 6G technology (Alsamhi et al., 2024). It stimulates interest with its potential to design the interaction paradigms. Therefore, it is possible to emphasize the idea of building a metaverse universe based on online interaction components, where the elements of different languages and cultures form a whole, based on the concept of “heterotopia”. Heterotopia refers to the phenomenon of creating “worlds within worlds” by combining the unique elements that can generate new layers, relationships and meanings (Foucault, 1986). This approach can be extended in an attempt to mold spatial pattern structures for various activities that can be participated in through embodiment. Consequently, creating metaverse universes woven with the SVR and CVE concepts in terms of heterotopia can be contemplated as the cornerstone of new generation interaction paradigms.

Consequently, even though immersive technologies have been developing with the power to shape the educational paradigms of the future, it would be pertinent to emphasize to the potential difficulties of the immersive learning method. Therefore, it is possible to say that factors such as “accessibility”, “content development”, “costs”, “privacy concerns” (Fernandes et al., 2023; see also Selvakumar and Sivakumar, 2023; Kuhail et al., 2022); “the degree of freedom of its users”, “regulatory concerns”, “security of the personal data”, “addictions and mental health issues” (Camilleri and Camilleri, 2023), “curriculum design” and “stakeholder expectations” (Brawn, 2024; Li et al., 2016) and “technical problems” (Serrano-Ausejo and Marell-Olsson, 2024)



come to transpire as issues that should be considered in the development of CEFR-based immersive learning. As far as the relevant context is concerned, it is possible to conclude that implementing research and development (R&D) practices based on these difficulties has transpired as a prerequisite in order for these technologies to achieve their full potential.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## Author contributions

GHD: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

## References

- Alsamhi, M. H., Hawbani, A., Kumar, S., and Hamood, S. (2024). Multisensory metaverse-6G: a new paradigm of commerce and education. *IEEE Access* 12, 75657–75677. doi: 10.1109/ACCESS.2024.3392838
- Altomari, L., Altomari, N., and Iazzolino, G. (2023). Gamification and soft skills assessment in the development of a serious game: design and feasibility pilot study. *JMIR Serious Games* 11:e45436. doi: 10.2196/45436
- Ansari, F., Erol, S., and Sihni, W. (2018). Rethinking human-machine learning in industry 4.0: how does the paradigm shift treat the role of human learning? *Proc. Manuf.* 23, 117–122. doi: 10.1016/j.promfg.2018.04.003
- Battal, M. (2022). “Learning methodology of the future: contextual learning in immersive virtual environments,” in *2nd International Conference on Digital Business Management and Economics (Books of Abstract)* (Konya: NEÜ Pub), 275–276. Available online at: <https://icdbme2022.tarsus.edu.tr/Files/ckFiles/icdbme2022-tarsus-edu-tr/Abstracts/ISL046.pdf>
- Bezcioglu-Goktolga, I., and Yagmur, K. (2022). Intergenerational differences in family language policy of Turkish families in the Netherlands. *J. Multi. Multicult. Dev.* 43, 891–906. doi: 10.1080/01434632.2022.2036746
- Bonner, E., Lege, R., and Frazier, E. (2023). Teaching CLIL courses entirely in virtual reality: educator experiences. *CALICO* 40, 45–67. doi: 10.1558/cj.22676
- Brawn, J. R. (2024). Creating immersive language learning environments for young learners. *AJHSSR*, 8/5, 165–169.
- Camilleri, M. A., and Camilleri, A. C. (2023). “Metaverse education: opportunities and challenges for immersive learning in virtual environments” in *The 4th Asia Conference on Computers and Communications (ACCC 2023)* (New York, NY: ACM Digital Library). doi: 10.2139/ssrn.4521895
- CEFR (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Strasbourg: Cambridge University Press. Available online at: <https://rm.coe.int/16802fc1bf>
- CEFRCV (2020). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume*. Strasbourg: Council of Europe Publishing. Available online at: <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>
- Chabot, S., Drozdal, J., Peveler, M., Zhou, Y., Su, H., Braasch, J., et al. (2020). “A collaborative, immersive language learning environment using augmented panoramic imagery,” in *6th International Conference of the Immersive Learning Research Network*. (San Luis Obispo, CA; New Jersey, NJ: IEEE), 225–229. doi: 10.23919/iLRN47897.2020.9155140
- Chiu, C.-C., Sainath, T. N., Wu, Y., Prabhavalkar, R., Nguyen, P., Chen, Z., et al. (2018). “State-of-the-art speech recognition with sequence-to-sequence models,” in *International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2018* (New Jersey, NJ: IEEE), 4774–4778. doi: 10.1109/ICASSP4228.2018
- Choe, J., Chen, Y., Chan, M. P. Y., Li, A., Gao, X., Holliday, N., et al. (2022). “Language-specific effects on automatic speech recognition errors for world Englishes,” in *Proceedings of the 29th International Conference on Computational Linguistics* (New York, NY: International Committee on Computational Linguistics), 7177–7186. Available online at: <https://aclanthology.org/2022.coling-1.628/>
- Choi, L., and Chung, S. (2021). Navigating online language teaching in uncertain times: challenges and strategies of EFL educators in creating a sustainable technology-mediated language learning environment. *Sustainability* 13:7664. doi: 10.3390/su13147664
- Chun, D. M., Smith, B., and Kern, R. G. (2016). Technology in language use, language teaching, and language learning. *Modern Lang. J.* 100, 64–80. doi: 10.1111/modl.12302
- Damaševičius, R., and Zailskaitė-Jakšė, L. (2024). “Digital twin technology: necessity of the future in education and beyond,” in *Automated Secure Computing for Next-Generation Systems*, ed. A. K. Tyagi (Beverly, MA: Wiley), 1–22. doi: 10.1002/97811394213948
- Delić, V., Perić, Z., Sečujski, M., Jakovljević, N., Nikolić, J., Mišković, D., et al. (2019). Speech technology progress based on new machine learning paradigm. *Comput. Intell. Neurosci.* 1:4368036. doi: 10.1155/2019/4368036
- Diren, D., and Horzum, D. M. B. (2024). Educational data mining with decision tree and rule induction: a case study of SAU ILITAM. *PUJE* 61, 94–120. doi: 10.9779/paufed.1085483
- Divekar, R. R., Drozdal, J., Chabot, S., Zhou, Y., Su, H., Chen, Y., et al. (2022). Foreign language acquisition via artificial intelligence and extended

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that Gen AI was used in the creation of this manuscript. The raw form of the figures was developed using Copilot (2024), while the numerical data, text and visual additions were created by author using AutoCAD (2024) and Adobe, Photoshop (2024).

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- reality: design and evaluation. *Comp. Assisted Lang. Learn.* 35, 2332–2360. doi: 10.1080/09588221.2021.1879162
- Divekar, R. R., Peveler, M., Rouhani, R., Zhao, R., Kephart, J. O., Allen, D., et al. (2019). “CIRA: an architecture for building configurable immersive smart-rooms,” in *Intelligent Systems and Applications. IntelliSys 2018. Advances in Intelligent Systems and Computing*, eds. K. K. Arai and S. R. Bhatia (Springer), 76–95. doi: 10.1007/978-3-030-01057-7\_7
- Dong, Y., Gui, T., Zhang, M., and Zong, Y. (2022). Language confidence and acquisition: Perception of accent affects oral speaking skill in second language acquisition. *Adv. Soc. Sci. Educ. Hum. Res.* 638, 1123–1126. doi: 10.2991/assehr.k.220110.210
- Doshi, G., and Bisht, A. (2021). Implementation of multilingual chatbot. *IJRASET*, 9/7, 1367–1373. doi: 10.22214/ijraset.2021.36012
- Eisenmann, M., and Steinbock, J. (2024). “Global citizenship education in social virtual reality for future English teachers,” in *10th International Conference on Higher Education Advances (HEAd24)* (Valencia: Editorial Universitat Politècnica de València), 83–90. doi: 10.4995/HEAd24.2024.17165
- Fernandes, F. A., Rodrigues, C. S. C., Teixeira, E. N., and Werner, C. M. L. (2023). Immersive learning frameworks: a systematic literature review. *IEEE Trans. Learn. Technol.* 16/5, 736–747. doi: 10.1109/TLT.2023.3242553
- Foucault, M. (1986). Of other spaces. *Diacritics* 16, 22–27. doi: 10.2307/464648
- Freeman, G., and Acena, D. (2021). “Hugging from a distance: Building interpersonal relationships in social virtual reality,” in *Proceedings of the 2021 ACM International Conference on Interactive Media Experiences* (New York, NY: Association for Computing Machinery), 84–95. doi: 10.1145/3452918.3458805
- Ghorbani, S., and Hansen, J. H. L. (2024). Advanced accent/dialect identification and accentness assessment with multi-embedding models and automatic speech recognition. *J. Acoust. Soc. Am.* 155, 3848–3860. doi: 10.1121/10.0026235
- González-Cacho, T., and Abbas, A. (2022). Impact of interactivity and active collaborative learning on students’ critical thinking. *Higher Educ.* 17, 254–261. doi: 10.1109/RITA.2022.3191286
- Gotti, M. (2015). Code-switching and plurilingualism in English-medium education for academic and professional purposes. *Lang. Learn. Higher Educ.* 5/1, 83–103. doi: 10.1515/cercl-2015-0005
- Gräßler, L., and Pöhler, A. (2017). “Produktentstehung im Zeitalter von Industrie 4.0,” in *Handbuch Gestaltung digitaler und vernetzter Arbeitswelten*, eds. G. W. Maier, G. Engels and E. Steffen (Berlin: Springer), 15, 1–21. doi: 10.1007/978-3-662-52903-4\_23-1
- Gruber, A., Canto, S., and Jauregi-Ondarra, K. (2023). Exploring the use of social virtual reality for virtual exchange. *ReCALL*, 35, 258–273. doi: 10.1017/S0958344023000125
- Gruzdeva, M. L., Frolova, N. K., Smirnova, Z. V., Tsybalov, S. D., and Garin, A. P. (2024). “The use of artificial intelligence in teaching foreign languages,” in *Ecological Footprint of the Modern Economy and the Ways to Reduce It. Advances in Science, Technology and Innovation*, eds. B. S. Sergi, E. G. Popkova, A. A. Ostrovskaya, A. A. Chursin, and Y. V. Ragulina, Y. V. (Cham: Springer), 261–265. doi: 10.1007/978-3-031-49711-7\_44
- Gupta, P., and Goyal, P. (2022). Is game-based pedagogy just a fad? A self-determination theory approach to gamification in higher education. *Int. J. Educ. Manage.* 36, 341–356. doi: 10.1108/IJEM-04-2021-0126
- Guyon, I., Goudet, O., and Kalainathan, D. (2019). “Evaluation methods of cause-effect pairs,” in *Cause Effect Pairs in Machine Learning, The Springer Series on Challenges in Machine Learning*, eds. I. Guyon, A. Statnikov and B. Batu (Springer, Cham), 27–99. doi: 10.1007/978-3-030-21810-2\_2
- Han, X., Yu, H., You, W., Huang, C., Tan, B., Zhou, X., et al. (2022). Intelligent campus system design based on digital twin. *Electronics* 11/21, 1–20. doi: 10.3390/electronics11213437
- Hassan, M. A., Habiba, U., Majeed, F., and Shoaib, M. (2021). Adaptive gamification in e-learning based on students’ learning styles. *Inter. Learn. Environ.* 29, 545–565. doi: 10.1080/10494820.2019.1588745
- Heidicker, P., Langbehn, E., and Steinicke, F. (2017). “Influence of avatar appearance on presence in social VR,” in *2017 IEEE Symposium on 3D User Interfaces (3DUI)* (New Jersey, NJ: IEEE), 233–234. doi: 10.1109/3DUI.2017.7893357
- Hinton, G. E. (2021). How to represent part-whole hierarchies in a neural network. *Neural Comput.* 35, 413–452. doi: 10.1162/neco\_a\_01557
- Huang, W., Hew, K. F., and Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *J. Comp. Assist. Learn.* 38, 237–257. doi: 10.1111/jcal.12610
- Hwang, G.-J., Fathi, J., and Rahimi, M. (2025). Fostering EFL learners’ speaking skills and flow experience with video-dubbing tasks: a flow theory perspective. *J. Comp. Assist. Learn.* 41/e13120, 1–18. doi: 10.1111/jcal.13120
- Hwang, Y., and Lee, S. M. (2024). “Can we go to Tower Bridge?: teaching british culture from text to immersive realities. *Innovation Lang. Learn. Teach.* 1–12. doi: 10.1080/17501229.2024.2439406
- Jaensch, F., Csiszar, A., Scheifele, C., and Verl, A. (2018). “Digital twins of manufacturing systems as a base for machine learning,” in *25th International Conference on Mechatronics and Machine Vision in Practice* (New Jersey, NJ: IEEE), 1–6. doi: 10.1109/M2VIP.2018.8600844
- Jiang, H., Cheng, Y., Yang, J., et al. (2022). AI-powered chatbot communication with customers: dialogic interactions, satisfaction, engagement, and customer behavior. *Comp. Hum. Behav.* 134/107329, 1–14. doi: 10.1016/j.chb.2022.107329
- Kann, A. (2022). “Voice assistants have a plurilingualism problem,” in *4th Conference on Conversational User Interfaces (CUI 2022)*, (Glasgow, United Kingdom. ACM, New York, NY, USA), 5 pages. doi: 10.1145/3543829.3544526
- Khalid, S., Ullah, S., Ali, N., Alam, A., Rasheed, N., Fayaz, M., et al. (2021). The effect of combined aids on users’ performance in collaborative virtual environments. *Multi. Tools Appl.* 80, 9371–9391. doi: 10.1007/s11042-020-09953-9
- Kharkhurin, A. V., Koncha, V., and Charkhabi, M. (2023). Effects of plurilingualism and pluriculturalism on creativity: testing the mediating role of tolerance and intolerance of ambiguity. *Int. J. Multi.* 1–24. doi: 10.1080/14790718.2023.2242373
- Khojasteh, N., and Won, A. S. (2021). Working together on diverse tasks: A longitudinal study on individual workload, presence and emotional recognition in collaborative virtual environments. *Front. Virtual Reality* 2/643331, 1–24. doi: 10.3389/frvir.2021.643331
- Koenecke, A., Nam, A. J., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., et al. (2020). Racial disparities in automated speech recognition. *Proc. Natl. Acad. Sci. U. S. A.* 117, 7684–7689. doi: 10.1073/pnas.1915768117
- Krath, J., Schürmann, L., and von Korfes, H. F. O. (2021). Revealing the theoretical basis of gamification: a systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Comp. Hum. Behav.* 125:106963. doi: 10.1016/j.chb.2021.106963
- Kuhail, M. A., ElSary, A., Farooq, S., Alghamdi, A. (2022). Exploring immersive learning experiences: a survey. *Informatics* 9/4:75. doi: 10.3390/informatics9040075
- Kuteneva, I. E., Bystray, E. B., Molchanov, S. G., Seliverstova, I. A., and Semenova, M. L. (2021). Role of authentic materials in training future managers for intercultural communication through content and language integrated learning. *Nuances Est. Sobre Educ.* 32, 1–15. doi: 10.32930/nuances.v32i00.9121
- Lai, K. W. K., and Chen, H. J. H. (2022). An exploratory study on the accuracy of three speech recognition software programs for young Taiwanese EFL learners. *Int. Learn. Environ.* 32, 1582–1596. doi: 10.1080/10494820.2022.2122511
- Lamb, R. L., Annetta, L., Firestone, J., and Etopio, E. (2018). A meta-analysis with examination of moderators of student cognition, affect, and learning outcomes while using serious educational games, serious games, and simulations. *Comp. Hum. Behav.* 80, 158–167. doi: 10.1016/j.chb.2017.10.040
- Lang, M., and Müller, M. (2020). Von augmented reality bis KI - Die wichtigsten IT-Themen, die Sie für Ihr Unternehmen kennen müssen. *Hanser* 1–15. doi: 10.3139/9783446464353.fm
- Laranjo, L., Dunn, A. G., Tong, H. L., Kocballi, A. B., Chen, J., Bashir, R., et al. (2018). Conversational agents in healthcare: a systematic review. *J. Am. Med. Inf. Assoc.* 25, 1248–1258. doi: 10.1093/jamia/ocy072
- Lee, S., and Jeon, J. (2023). Addressing automatic speech recognition for ELT from the global Englishes perspective. *ELT J.* 77, 435–444. doi: 10.1093/elt/ccad038
- Li, K.-C., Chen, C.-T., Cheng, S.-Y., and Tsai, C.-W. (2016). The design of immersive English learning environment using augmented reality. *Univ. J. Educ. Res.* 4/9, 2076–2083. doi: 10.13189/ujer.2016.040919
- Lin, C.-J., Wang, W.-S., Lee, H.-Y., Huang, Y.-M., and Wu, T.-T. (2024). Interventions in STEM education through speech recognition-based learning analysis. *J. Educ. Comp. Res.* 63, 311–335. doi: 10.1177/07356331241307904
- Link, S., Mehrzad, M., and Rahimi, M. (2020). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Computer Assisted Language Learning*, 33, 1–30. doi: 10.1080/09588221.2020.1743323
- Liu, G., Fathi, J., and Rahimi, M. (2024). Using digital gamification to improve language achievement, foreign language enjoyment, and ideal L2 self: a case of English as a foreign language learner. *J. Comp. Assist. Learn.* 2024, 1–18. doi: 10.1111/jcal.12954
- Madni, A. M., Erwin, D., and Madni, A. (2019). “Exploiting digital twin technology to teach engineering fundamentals and afford real-world learning opportunities,” in *126th ASEE Annual Conference and Exposition, Tampa, Florida* (Washington, DC: ASEE). doi: 10.18260/1-2-32800
- Massey, A., Montoya, M., Samuel, B. M., and Windeler, J. (2024). Presence and team performance in synchronous collaborative virtual environments. *Small Group Res.* 55, 290–329. doi: 10.1177/10464964231185748
- Mee Mee, R. W., Shahdan, T. S. T., Ismail, M. R., Abd Ghani, K., Pek, L. S., Von, W. Y., et al. (2020). Role of gamification in classroom teaching: Pre-service teachers’ view. *Int. J. Eval. Res. Educ.* 9/3, 684–690. doi: 10.11591/ijere.v9i3.20622
- Miles, M., and Huberman, B. A. M. (1994). *Qualitative Data Analysis: An expanded Sourcebook, 2nd Edn.* London: Sage.

- Mills, N., Courtney, M., Dede, C., and Dressen, A. and Gant, R. (2020). Culture and vision in virtual reality narratives. *Foreign Lang. Ann.* 53, 733–760. doi: 10.1111/flan.12494
- Moustafa, F., and Steed, A. (2018). “A longitudinal study of small group interaction in social virtual reality,” in *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (New York, NY: Association for Computing Machinery), 1–10. doi: 10.1145/3281505.3281527
- Najafian, M., and Russel, M. (2020). Automatic accent identification as an analytical tool for accent robust automatic speech recognition. *Speech Commun.* 122, 44–55. doi: 10.1016/j.specom.2020.05.003
- Poppe, E., Brown, R., Recker, J., Johnson, D., and Vanderfeesten, I. (2017). Design and evaluation of virtual environments mechanisms to support remote collaboration on complex process diagrams. *Inf. Syst.* 66, 59–81. doi: 10.1016/j.is.2017.01.004
- Prasad, A., and Jyothi, P. (2020). “How accents confound: Probing for accent information in end-to-end speech recognition systems,” in *Annual Meeting of the Association for Computational Linguistics* (Stroudsburg, PA: Association for Computational Linguistics), 3739–3753. doi: 10.18653/v1/2020.acl-main.345
- Rauschnabel, P. A., Felix, R., Hinsch, C., Shahab, H., and Alt, F. (2022). What is XR? Towards augmented and virtual reality. *Comp. Hum. Behav.* 133, 1–18. doi: 10.1016/j.chb.2022.107289
- Rivas, S. F., Saiz, C., and Ossa, C. (2022). Metacognitive strategies and development of critical thinking in higher education. *Front. Psychol.* 13/913219, 1–13. doi: 10.3389/fpsyg.2022.913219
- Schoser, B. (2023). Framing artificial intelligence to neuromuscular disorders. *Curr. Opin. Neurol.* 36, 424–426. doi: 10.1097/WCO.0000000000001190
- Seidel, S., Yepes, G., Berente, N., and Nickerson, J. V. (2022). “Designing the metaverse,” in *Proceedings of the 55th Hawaii International Conference on System Sciences* (Honolulu: HICSS), 6699–6708. doi: 10.24251/HICSS.2022.811
- Selvakumar, S., and Sivakumar, P. (2023). Immerive learning: unlocking the future of education. *Thiagarajar Coll. Preceptors Edu Spect.* 5/1, 12–20. doi: 10.34293/eduspectra.v5i1-may23.003
- Senyigit, Y., and Okur, A. (2019). Speaking skill and pronunciation training in teaching Turkish for foreigners. *J. Mehmet Akif Ersoy Univ. Fac. Educ.* 52, 519–549.
- Serrano-Ausejo, E., and Marell-Olsson, E. (2024). Opportunities and challenges of using immersive technologies to support students’ spatial ability and 21st-century skills in K-12 education. *Educ. Inf. Technol.* 29, 5571–5597. doi: 10.1007/s10639-023-11981-5
- Shadiev, R., Xueying, W., and Huang, Y. M. (2020). Promoting intercultural competence in a learning activity supported by virtual reality technology. *Int. Rev. Res. Open Distance Learn.* 21, 157–174. doi: 10.19173/irrodl.v21i3.4752
- She, M., Xiao, M., and Zhao, Y. (2023). Technological implication of the digital twin approach on the intelligent education system. *Int. J. Hum. Robot.* 20. doi: 10.1142/S0219843622500050
- Shruthi, H. L., Radhakrishnan, A., Veigas, A. D., Railis, D. J., and Dines, R. S. (2025). Analyzing pedagogy and education in English language teaching using information and communication technology. *Educ. Inf. Technol.* 2025, 1–23. doi: 10.1007/s10639-025-13439-2
- Sierra Rativa, A., Postma, M., and Van Zaenen, M. (2020). The influence of game character appearance on empathy and immersion: virtual non-robotic versus robotic animals. *Simul. Gaming* 51, 685–711. doi: 10.1177/1046878120926694
- Singh, K. N., Chanu, Y. J., and Pangsatbam, H. (2024). MECOS: a bilingual manipuri-english spontaneous code-switching speech corpus for automatic speech recognition. *Comp. Speech Lang.* 87/101627, 1–14. doi: 10.1016/j.csl.2024.101627
- Skulmowski, A., and Rey, G. D. (2018). Embodied learning: introducing a taxonomy based on bodily engagement and task integration. *Cognit. Res. Princ. Implic.* 3/6, 1–10. doi: 10.1186/s41235-018-0092-9
- Soruc, A. (2016). “Developing phonological awareness and improving orthographic knowledge for teaching pronunciation,” in *The Theory and Practice of English Language Teaching*, eds. B. Inan Karagul and D. Yuksel (Kocaeli: KUV Publications), 8–20.
- Stankova, E., Chlumska, R., and Zerzanova, D. (2022). The relationship between native and foreign language speaking proficiency in university students. *J. Lang. Educ.* 8/2, 122–139. doi: 10.17323/jle.2022.11501
- Suh, A., Cheung, C. M., and Ahuja, M. and Wagner, C. (2017). Gamification in the workplace: the central role of the aesthetic experience. *J. Manage. Inf. Syst.* 34, 268–305. doi: 10.1080/07421222.2017.1297642
- Sun, Y., Zhang, Q., Bao, J., Lu, Y., and Liu, S. (2024). Empowering digital twins with large language models for global temporal feature learning. *J. Manuf. Syst.* 74, 83–99. doi: 10.1016/j.jmsy.2024.02.015
- Takshara, K. S., Bhuvanewari, G., Kumar, T. S. P. (2025). The game of language learning and rewiring biocognitive receptors. *MethodsX* 14:103143. doi: 10.1016/j.mex.2024.103143
- Takuchi, N., and Hanks, E. (2024). Social virtual reality for L2 Spanish development: learning how to interact with others in a high-immersion virtual space. *Modern Lang.* 108/4, 954–975. doi: 10.1111/modl.12968
- Tanenbaum, T. J., Hartoonian, N., and Bryan, J. (2020). “How do I make this thing smile?”: an inventory of expressive nonverbal communication in commercial social virtual reality platforms,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY: ACM), 1–13. doi: 10.1145/3313831.3376606
- Tang, F., Chen, X., Zhao, M., and Kato, N. (2023). The roadmap of communication and networking in 6G for the metaverse. *IEEE Wireless Commun.* 30, 72–81. doi: 10.1109/MWC.019.2100721
- Tao, F., Sui, F., Liu, A., Qi, Q., Zhang, M., Song, B., et al. (2019). Digital twin-driven product design framework. *Int. J. Prod. Res.* 57, 3935–3953. doi: 10.1080/00207543.2018.1443229
- Thanyawatpokin, B., and Vollmer, C. (2022). “Language learner Identity and games and gamification in the language learning classroom: observations from the Japanese context,” in *Individual and Contextual Factors in the English Language Classroom*, eds. R. Al-Mahrooqi and C. J. Denman (Cham: Springer), 323–344. doi: 10.1007/978-3-030-91881-1\_16
- Thompson, G., and McKinley, J. (2018). “Integration of content and language learning,” in *The TESOL Encyclopedia of English Language Teaching, 1st Edn*, eds. J. I. Lontos, M. Delli Carpini, and S. Abrar-ul-Hassan (Hoboken, NJ: Wiley). doi: 10.1002/9781118784235.eelt0634
- Tillman, R., and Louwerse, M. (2018). Estimating emotions through language statistics and embodied cognition. *J. Psycholing. Res.* 47, 159–167. doi: 10.1007/s10936-017-9522-y
- Titova, S. V., and Temuryan, K. T. (2024). Intelligent agents in teaching foreign languages: typology, opportunities, challenges. *Lang. Cult.* 2024/65, 262–287. doi: 10.17223/19996195/65/12
- Wu, Y. J. A., Lan, Y. J., Huang, S. B. P., and Lin, Y. T. R. (2019). Enhancing medical students’ communicative skills in a 3D virtual world. *J. Educ. Technol. Soc.* 22, 18–32.
- Yu, H., and Nazir, S. (2021). Role of 5G and artificial intelligence for research and transformation of English situational teaching in higher studies. *Mobile Inf. Syst.* 2021, 1–16. doi: 10.1155/2021/3773414
- Zhang, H., and Lin, Y. (2022). “Improve few-shot voice cloning using multi-modal learning,” *ICASSP 2022, Singapore* (New Jersey, NJ: IEEE), 8317–8321. doi: 10.1109/ICASSP43922.2022.9746233
- Zhang, Z., and Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon* 10/3, 1–13. doi: 10.1016/j.heliyon.2024.e25370
- Zhao, J.-H., Chen, Z.-W., and Yang, Q.-F. (2024). I do and I understand: a virtual reality-supported collaborative design-assessing activity for EFL students. *System* 121:103213. doi: 10.1016/j.system.2023.103213
- Zhao, Y., Jiang, J., Chen, Y., Liu, R., Yang, Y., Xue, X., et al. (2022). Metaverse: perspectives from graphics, interactions and visualization. *Visual Inf.* 6, 56–67. doi: 10.1016/j.visinf.2022.03.002
- Zhuang, C. B., Liu, J. H., Xiong, H., Ding, X., Liu, S., Weng, G., et al. (2017). Connotation, architecture and trends of product digital twin. *Comp. Integ. Manuf. Syst.* 23, 753–768. doi: 10.13196/j.cims.2017.04.010