Michaela Nuesser, Dongping Zheng* and Jin Dong

Optimizing affordances of high immersive virtual reality for language learning

https://doi.org/10.1515/jccall-2023-0006
Received March 11, 2023; accepted December 29, 2023; published online March 5, 2024

Abstract: Virtual reality (VR) technology has enjoyed a hype in the past years that made many realize the huge opportunity VR technology poses for language learning. Nevertheless, there has been a significant lack of research publications investigating the use of VR for language learning. In 2021 and 2022, however, we have seen a drastic change with a multitude of new research projects reported. In this article, we first present which common VR application features offer affordances for second language learning and then survey to what degree these features have been adopted in the reviewed studies. The main investigated affordances of VR that we identified to be beneficial to language learning are its presence, interactivity, embodiment, and emotional connection. Finally, we offer an outlook for researchers and developers to adopt and implement VR features for future language learning applications.

Keywords: affordances; virtual reality; language learning; emotional connection; interactivity; embodiment

1 Introduction

Virtual reality (VR) technology has created a variety of new ways for learners to learn foreign languages since recently finding its way into foreign language learning pedagogy. Currently, most proposed VR language learning applications have received little research attention from scholars not affiliated with the projects themselves. In this paper, we reviewed and analyzed VR language learning applications as described in design papers published between 2013 and 2022 in academic journals. Our aim was to identify to what degree the research projects afford

***Corresponding author: Dongping Zheng**, Department of Second Language Studies, University of Hawai'i at Mānoa, Honolulu, HI, USA, E-mail: zhengd@hawaii.edu
Michaela Nuesser, Department of Second Language Studies, University of Hawai'i at Mānoa, Honolulu, HI, USA, E-mail: mnuesser@hawaii.edu. https://orcid.org/0000-0003-1667-6921
Jin Dong, Department of East Asian Studies, Princeton University, Princeton, NJ, USA, E-mail: jd1398@princeton.edu

meaningful language learning and fully take advantage of the possibilities VR offers for language learning. We only included articles which describe three-dimensional (3D) worlds that can be accessed via a VR headset and that allow interaction between player and game objects. Thus, our selection criteria exclude 2D virtual worlds such as Second Life or Google Street View virtual applications. Contrary to our expectations, we saw most contributions to the field in journals not necessarily familiar to second language mainstream researchers, such as *Language Learning and Technology*, *Computer Assisted Language Learning*, *System*, or *CALICO*, but in journals of computer science like *IEEE* or *Interactive Learning Environments*.

In this paper, we define language learning following Zheng et al. (2017), who advocate for language learning as situated and interactive, emphasizing the importance of creating real or realistic learning environments. Real(istic) contexts provide learners with the opportunity to engage in interactions with objects and people in the environment as if they were doing things in which language plays a part in real world situations. Consequently, language learning applications should allow students to embrace linguistic and cultural semiotics in interaction with others and the environment. Zheng et al. (2018) provide this through opportunities for translanguaging (Wei, 2018) with co-players who speak each other's home and additional languages.

There are several Augmented Reality (AR) applications situating themselves within this line of language learning research that provides learning experiences with the intention to engage learners in exploiting affordances of designed elements for achieving social embodiment (e.g., Hellermann et al., 2019; Holden and Sykes, 2011; Zheng et al., 2018). There is only one VR application that does (Cui et al., 2022). However, this does not mean that other developed VR language learning projects do not offer situated and interactive language learning environments, which is why this review was conducted.

To introduce our review, we provide our working definition of virtual reality based on both computer science and language science standards. We then outline the selection criteria chosen for our review. Then, we present which common VR application features offer affordances for second language learning, before we survey to what degree these affordances have been used by VR researchers and developers of language learning games. This structure aligns with these two main review-oriented research questions: Which technological affordances does VR offer for language learning? And to what extent have the reviewed projects (2013–2022) employed the VR possibilities? In addition, we will elaborate the importance of the chosen definition of learning for virtual reality language learning applications and provide recommendations for future VR foreign language learning application developers.

2 Definitions of VR

Considering that computer sciences and language learning sciences have traditionally not shared much research, it is not surprising that their definitions of VR differ from one another. In the computer science field, according to Sherman and Craig (2018), VR is an artificial world that one experiences through sensory stimuli (such as sights and sounds) delivered by a computer and in which one's actions influence the environment's events to some extent. Technologically, it should include the following elements: three dimensions, a viewer-centered perspective, real-time interactivity, and spatialized audio.

In the language learning field, the concept of VR varies. While Garcia et al.'s (2019) definition aligns with those of the computer sciences, Lan (2020) and Godwin-Jones (2021) both include what Lan (2020) calls non-immersive VR – applications which place users in an environment viewed through a traditional desktop set-up that they can directly control with a mouse and keyboard. In these environments, users often explore worlds using avatars from a third-person perspective, while certain non-immersive VR platforms allow users to switch to a first-person view by using the mouse. Following the computer sciences' normative definition of VR and Garcia et al. (2019), we exclude desktop application versions of VR. Though this is uncommon in the field of language learning, we believe that the critical features of immersive VR – like head movement controlling a player's field of view and hand movement controlling rendered hands with the ability to grab or select virtual objects – allow for language learning opportunities distinct from those of non-immersive VR. Thus, in our working definition, we exclude applications like *Second Life*, which has received considerable language learning research attention in the past (e.g., Lan, 2014; Zheng, 2012).

3 Paper selection criteria

In search of all recent articles that present 360-degree environment foreign language learning projects with VR typical interaction features, we conducted several keyword searches. Initially, we conducted a keyword search using the terms "virtual reality" and "language learning" with a filter for articles published between 2013 and 2020. We chose this time frame because Lin and Lan (2015) conducted a similar review of articles published between 2004 and 2013. Using Google Scholar, the search yielded 218,000 results. We found that many of them were not relevant because they included articles about Augmented Reality (AR) or programming - rather than natural languages.

In an attempt to narrow down the search results, we refined our search terms to "virtual reality, foreign language learning and HTC", "virtual reality, foreign language learning and Oculus Rift", and "virtual reality, foreign language learning and Meta Quest" since HTC, Oculus Rift (2013–2021) and Meta Quest (2021- present) are the main platforms for intractable VR design. These searches yielded 11,000, 11,800, and 362 results, respectively.

We further refined our selection by excluding review or planning papers to 28 papers of interest. After carefully separating experimental designs that focused on testing single features in VR environments from our intended learning environment experience studies, we selected eight articles that met all of our criteria, including having three dimensions and VR typical interaction features.

4 Establishing review criteria: affordances of VR for language learning

The great value for language learning that VR applications offer is linked to its clear potential to mimic the real world. In fact, it can even add language learning opportunities to the real world that are not usually possible like instant time and place travel, connecting geographically distant people, or trying out situations without real life impacts. This section introduces common VR application features like 360-degree view, interaction, simulation of tactile and haptic touch and grip, and narratives, and connects them to their possible affordances for foreign language learning.

One of the most popular features of VR is inarguably its potential to afford perceived feelings of presence and interactivity (Ryan, 2015; Walsh & Pawlowski, 2002). Presence is considered as "the subjective experience of being in one place or environment, even when one is physically situated in another" (Witmer & Singer, 1998). As Cowley and Fester (2017) write, presence contributes to being and telling. Presence makes one's individual contributions to the world about relationships and languaging with others. For language learners, this means that a perceived feeling of presence allows them to feel how they contribute to a new world in a new language they are learning, contribute to relationships, rather than choosing between different phrases in their mental capacity in tasks.

Interactivity can be described as the degree to which a user can modify the VR environment in real-time (Steuer, 1995). This is similar to real life affordances for interactivity with the immediate environments. VR applications are most immersive if they can afford experiences that are limited to the physical principles that govern physical reality (Bermejo & Hui, 2021). Cowley and Fester (2017) point out how this

interactivity with materials is part of the feeling of presence that allows individuals to contribute to the world.

Another key feature of virtual reality applications is its ability to simulate tactile and haptic touch and grip. Tactile and haptic simulation provides the feeling of experiencing kinesthetic sensations in a virtual environment (Bermejo & Hui, 2021). Bermejo and Hui (2021) provide the example that users can manipulate virtual items with their bare hands or with a joystick while feeling their surface and certain weight of the objects. With that, embodiment of the player character becomes more accessible. This is a crucial affordance for language learning: research focusing on language learning in young children, for example, found a high significance of human embodied experiences with the surrounding world in cognitive development (e.g., Mandler, 2007). Johnson (2017) supports this idea by stating that our language and mind cannot exist independently of our bodily experiences. Research has shown that embodying player avatars who have a certain role in the game environment leads to a multitude of language learning facilitators through prompting and inviting participation in the surrounding world (Zheng et al., 2009).

The similarities between virtual reality and reality also provide affordances for building emotional connection, which has frequently been used in psychotherapy (Herrera et al., 2018; Markowitz et al., 2018). We assume this emotional connection increases with complexity of the storyline within experiences that are constrained by time, location, and physical principles. The Quest Atlantis project (Barab et al., 2007) shows how a narrative or a back story is necessary in a game to scaffold participation. When researching the design of Quest Atlantis, Barab et al. (2007) found that narrative, inquiry, and inscriptions can support "consequential play". This is defined as play that supports the development of knowledge, skills, and attitudes that can be used in real-world contexts. While the article focuses primarily on the use of narratives in educational contexts, it does briefly touch on the relationship between narrative and emotions. The authors argue that narratives can be engaging and emotionally compelling for learners, and that emotions can play a critical role in motivating and supporting learning. Another article on the same project, Barab et al. (2010), suggests that narrative and emotional engagement are closely linked in educational contexts, particularly when using games as a curriculum tool. They argue that narratives help create a context for learning by providing a meaningful storyline, characters, and settings that learners can connect with on an emotional level. This emotional engagement helps to motivate learners to actively participate in the learning process. The authors further suggest that games are particularly effective at engaging learners emotionally because they often contain strong narratives, which can be experienced through immersive gameplay. As learners progress through the game and encounter challenges and obstacles, they

Table 1: VR application features and their affordances for language learning.

VR application feature	Its affordance
360-degree view	Presence
Interaction possibilities with VR objects	Interactivity with the immediate learning environment
Tactile or haptic simulation	Embodiment
Coherent storyline	Emotional connection

become emotionally invested in the narrative and feel a sense of achievement when they overcome these challenges.

Bringing together the above concepts and our definition of language learning, we created a simplified framework (see Table 1), which we used to establish our review criteria, thus to operationalize our second research question, how much the reviewed projects employed the possibilities VR affords for language learning. We identified four most important VR features that we paired up with four affordances for language learning: VR's *360-degree view* mainly affords *presence*, interaction possibilities with VR game objects afford interactivity with the immediate learning environment, *tactile or haptic simulation* affords *embodiment of the player character*, and a *coherent storyline* affords an *emotional connection in context*. We conclude that a project that makes use of all four types of features is able to afford all four types of affordances for foreign language learning and thus would offer an optimal learning environment.

5 Findings: affordance possibilities implemented in the reviewed projects (2013–2022)

After reviewing the 28 VR language learning projects published between 2013 and 2022, we found 8 articles that met our criteria (an environment with three dimensions allowing for immersive-VR typical interaction features beyond its planning stage published in a research article between 2013 and 2022). In this findings section, we will report to what extent qualitatively these projects maximized the affordances of VR for language learning. The most important insight is that VR has great potential that is starting to be realized in varied forms even though much remains to be unexplored.

In this section, we focus on the four VR features that we identified as the most important affordances for creating meaningful language learning situations: (1)

360-degree view that encourages a feeling of presence; (2) voice, button, and motion command that allows for interactivity with the environment; (3) tactile and haptic simulation that aids embodiment of the player character; and (4) a storyline furthering a possible emotional connection between player and the game.

Looking at all reviewed papers combined, 360-degree view, as the hallmark of virtual reality, is the most exhausted feature in the reviewed projects that contributes to offering meaningful language learning situations. However, what sticks out the most is that quality of animation matters. While Garcia et al. (2019), Jia and Liu (2019), Ahlers et al. (2021), and Cai et al. (2021) show a 360-degree animation that is rather unrealistic – likely due to high production costs – others (Collins et al., 2019; Han, 2022; Urueta & Ogi, 2020) are closer to the animation quality of popular video games, while one of the reviewed projects, *Eduventure* by Lin et al. (2021), even makes use of (though rather low definition) 360-degree live action shots. Naturally, a more realistic environment contributes to greater feelings of presence, but well animated environments have proven themselves to foster immersion also. Eventually, if 360-degree camera technology improves and high-resolution capture becomes affordable, it might allow for the highest degree of virtually evoked presence.

Haptic simulation – for fostering embodiment – is one of the most exhausted VR features that help towards building meaningful language learning situations. This is especially the case when object handling is a necessity for progression through the game. The only projects that do not use haptic simulation are those that exclusively use voice commands (Han, 2022), which does not limit embodiment drastically as grabbing objects is not in focus in the story, or those that use buttons to be hovered over or clicked on with a cursor (Garcia et al., 2019; Jia & Liu, 2019; Lin et al., 2021). The latter does limit embodiment as, unlike real life, any haptic movement is substituted by cursor movement.

An aspect where most of the reviewed projects differ from one another is that of commands, i.e., voice, button, and motion commands that determine the degree of interaction with the environment. While the projects on the high end of the interactionality scale allow for interaction with AI characters or other player characters separated in space – Collins et al. (2019), Urueta & Ogi, (2020), Ahlers et al. (2021), and Han (2022), other projects only allow interaction through cursor movement to press buttons (Garcia et al., 2019; Jia & Liu, 2019; Lin et al., 2021).

The feature which can be regarded as one of the strongest to evoke meaningful language learning situations, a gripping storyline, is found the least in the reviewed projects. Indeed, projects range from mere exploration or character imitation games without a storyline to those which offer a theme and tasks but lack credibility through repetitiveness and absurdity. However, one game (Ahlers et al., 2021) exhibits a complex storyline in the form of an exit game, which – despite not

convincingly mirroring real life situations – can evoke an emotional connection that inspires players to play through the entire game.

A breakdown of each of the eight applications and their use of the four possible affordances (presence, interactivity with the immediate learning environment, embodiment, and emotional connection) foregrounds the differences across the reviewed papers. We provide a thick description of each project characterizing the type of 3D technology being rendered, activities participants did deploy the technology features, actions participants engaged in, and aspects of language learning. The Spanish vocabulary learning game by Garcia et al. (2019) is a 360-degree animated environment which can evoke feelings of presence in the house or apartment scene constructed by the researchers. Though the animation quality is not state of the art equivalent, the intended three-dimensional room with furniture is visible from 360 degrees. As the purpose of the game is vocabulary learning, players are prompted to show where several furniture pieces or appliances (e.g., the fridge) are located. Interaction with the environment is executed through pointing, hovering, and clicking on objects, which then light up in different colors based on whether the correct object was chosen. Motion through the apartment is also done through controller command. After finding the correct item, the player hears the correct object name. With these commands, the project shows low degrees of interaction with the environment as its major feature, pointing, hovering, and clicking, are not interactions that remind of those in the real world. For that reason, we also cannot find tactile or haptic simulation, which is thought to help embodiment of the player character in possible meaningful language learning situations. In addition, the storyline of the game is also limited to a redundant scavenger hunt, which is likely not to foster a strong emotional connection between the player and the game. Garcia et al. (2019) thus convincingly proposed a presence-evoking vocabulary learning game, which could bring variety to often monotone practice. However, most VR features that we identified as optimal affordances for creating meaningful language learning situations are not exhausted.

Lin et al. (2021) is another project that fits our criteria. Their game *Eduventure* is an interactive, simulated tour to help students of English for Tourism Purposes. Interesting about this project is its 360-degree live action shot environment. Instead of an animated environment, all scenes in *Eduventure* were filmed with 360 cameras. Even though the resolution needs to improve to fully evoke the feeling of presence in the exact original place, the choice of a real environment seems like the best option to get close to a real environment. This can be especially helpful for place-based learning situations in which players should get accustomed to an environment which exists in real life (traveling to a target language country, for example). However, also because of the 360-degree film rather than an animated three-dimensional environment, *Eduventure* shows low levels of embodiment of a player character through

no tactile or haptic simulation. In addition, players interact with the environment merely by clicking buttons to replay audio narrations of the tour or pointing at spots in the environment to uncover hidden elements, not involving much realistic interaction with the environment. The storyline in Lin et al.'s (2021) VR game is similarly thin – the player simply follows a tour to different destinations, possibly fostering only weak emotional connection between player and game. It must be noted that the game is intended to be followed up by a worksheet task in which players practice giving destination information to a partner. Lin et al.'s (2021) VR language learning game is thus a great step towards technology that can bring learners to faraway places and should be built upon when creating new VR language learning opportunities in the future. Still, it does not make use of most VR features that we identified as optimal affordances for creating meaningful language learning situations.

GaeltechVR, a VR Irish language learning project by Collins et al. (2019), stands out by proposing a 360-degree animated environment specifically made "livable" by adding native-speaking AI characters to the environment and spatialized audio, i.e. audio you can only hear when close to the object or character they are coming from. With that, the feeling of presence can be comparably high. In addition, players can walk up to the AI characters (shop owners, for example) and grab objects as necessary for proceeding in the game. This rather high degree of possible interaction with the environment is further extended by tactile or haptic simulation as the player can grab animated objects. As Collins et al. (2019) also point out, a feeling of ownership or embodiment can thus be evoked through what Ehrsson et al. (2005) call the rubber hand illusion. With all these affordances for meaningful language learning situations, the storyline remains a simple scavenger hunt similar to that of Garcia et al. (2019). Even though GealtechVR prompts players to search for items through dialogue with shop owners that they walk up to – to start the scavenger hunt and to bring the objects they ask for, which can seem more realistic than the situations in most other reviewed games, the emotional connection between the player and the game may be still lack substance due to its repetitiveness as that can lead to a lack of engagement and interest. This repetitiveness of doing something unnecessary can lead to redundant gameplay, which can make the game feel monotonous and uninspired, which can reduce the player's emotional investment in the game. Similarly, a simple vocabulary recognition game may not have enough depth or complexity to create an emotional connection with the player, as it may not have meaningful or impactful choices and story elements that can foster an emotional response. Collins et al.'s (2019) GaeltechVR is thus a language learning game with great promise for success. Its complex environment that allows interaction is probably as meaningful and realistic as a mere vocabulary learning game can get. However, VR's potential for embodiment, presence, and interaction is not fully

exploited, therefore, the affordances of VR for creating a meaningful language learning environment, and emotional connections provided by an embedded gripping storyline are not fully utilized.

Ahlers et al. (2021) propose a tandem language learning game called Hololingo!. Their game is specifically interesting as it is a multiplayer game, meaning that several players, separated in space, can play in the same virtual environment. The 360-degree environment is also animated, though slightly less aesthetic than GaeltechVR. The feeling of presence given is added to by the second person joining the same space so that players can see and hear each other's player characters. Interaction opportunities with the environment are also exhausted quite well as not only coordination with the second player is made necessary for solving problems in the game but other interactions like flipping switches or grabbing and dropping objects are also required. With that, tactile and haptic simulation when grabbing objects largely contributes to feelings of embodiment of the player character. Hololingo! is also unique in that players proceed in the game by completing tasks which are part of an exit game based on an old German myth that originated in the same place the game is set in. Though it remains speculative how realistic an exit room is, a strong emotional, or at least motivational connection to the game, can be assumed. Ahlers et al. (2021) thus propose a game which stands out through prompting for coordination among multiple players (rather than focusing on input and/or output) and targeting more than vocabulary learning. Though not necessarily equipped with a realistic storyline, Hololingo! has the highest chance of evoking an emotional or motivational connection with the game. All its VR features combined afford a relatively great chance for creating a meaningful language learning situation for players.

Jia and Liu (2019) present a VR language learning game in which kitchen words are to be learned. They state that they "aim to make full use of immersion and interaction" (p. 150). Even though the game has the ability to promote feelings of presence in the animated environment through its 360-degree three-dimensional scene, interaction with that environment is limited to looking around and hovering over and clicking on objects. With that, there is no tactile or haptic simulation. The intended storyline of "Words In Kitchen" is that of a little virtual character who asks the player to point at kitchen items. The only emotional connection that can be expected is to help the virtual character to find objects, which is thus expected to be comparably weak. Jia and Liu (2019, pp. 150–155) thus cannot stick to their promise to "make full use of immersion and interaction" (p. 150).

Han (2022) created another interesting game that allows players to interact with each other in one environment though spatially separated, making it a multiuser game. While there are no objects around that players interact with and thus no haptic simulations possible, interaction between player and environment comes from possible dialogue with AI and other player characters. With that, levels of possible

presence and interaction are high, and embodiment is not greatly compromised as object interaction is not necessary for proceeding in the game. The scenario itself, the learner going to an embassy to apply for a student visa and conducting an interview is not intrinsically emotionally evoking but mimics a real-life situation. In general, Han's (2022) game stands out as one that allows for interaction with real people and several AI characters. Players get to practice both speaking and coordinating. The game's potential for creating a meaningful language learning situation for players is thus mostly based on the complexity of coordination between players.

The game by Cai et al. (2021) is an animated 360-degree environment, which evokes feelings of presence in different social scenes. The aim of the game is to play the role of a character in a conversation by repeating memorized sentences. Players thus get to interact with the environment by talking to other characters, even though the interaction is limited to repetition of memorized phrases, meaning that it is less meaningful than if it were player driven. With that, there is no storyline that can be expected to further an emotional connection between player and game. Cai et al.'s (2021) VR language learning game thus makes use of VR technology to encourage mimicry of conversation, which is a novel way to promote familiarization with a foreign language. Still, most VR features that we identified as optimal affordances for creating meaningful language learning situations are not exploited.

Lastly, Urueta and Ogi's (2020) VR language learning game is unique in that it offers an animated environment to a teacher and student duo. The environment allows for creation of a sense of presence and possible object manipulation increases embodiment through tactile and haptic simulation. Players can interact not only with objects in the environment (those that are important to the storyline) but also with their teacher that can be spatially separated. The storyline can be influenced by the teacher playing the game and can thus range from evoking strong to evoking weak emotional connection. Generally, Urueta and Ogi's (2020) game stands out as one that allows for interaction with a real person, a teacher, and several AI characters. Players get to practice both speaking and coordinating (as long as the teacher fosters it). The game's potential for creating a meaningful language learning situation for players is thus mostly based on the complexity of coordination between teacher and student.

Our review shows that while there is a wide range of applications reported, most of them are centered on the lower end of the number of possible affordances for language learning. In general, the degree of VR features being used varies widely. In addition, we found few studies that fully exploited tactile or haptic simulation, 360 view, interaction with game objects, and a storyline setting to offer players presence, interactivity, embodiment, and an emotional connection to the game. Most research heavily relies on social presence as the main explanation for their reported effect of language learning. Social presence is generally defined as the degree to which users

perceive and experience the presence of others in that environment as real and socially related. Research found social presence has a positive effect on learners' motivation and engagement, which in turn is positively related to language learning (see Kang & Lee, 2018). However, how social presence provides affordances for learning and what specific features give rises to language activity are not fully understood.

We conjecture that there are mainly two reasons for the current state of VR research for language learning. One is that the studies mainly used perception-based motivation theories (Barrett et al., 2020) and survey methods (e.g., DeWitt et al., 2022). These are limited to participant self-report on their language learning experiences. This type of research addresses the learners' acceptances of a specific VR application but provides little insight on their affordances for learning or the design rationale, meaning their analysis does not address how design gives meaning to motivation and learning.

Secondly, the main theoretical framing of the design is Schmidt's (1990) Noticing Hypothesis (Lin et al., 2021; Sun et al., 2020). Therefore, most studies treated the designed VR environment merely as input, possibly overlooking many action possibilities afforded by VR as a learning environment. The differences between input and affordances were outlined by Van Lier (2000). Seeing a complex 3D space with designed artifacts as mere input may lead to convenient explanations such as cognitive overload when learners feel overwhelmed or cannot perceive the usefulness of a specific VR space. In our definition of learning, learners can participate socioculturally and ecologically (Zheng et al., 2017), become part of a space or a place, and do things with the world and/or others; learners' attention is paid to what they can act on, rather than doing a required task. The parts or elements that they could not pick up at the beginning, they may pick them up at the end or other team members can. Therefore, learning occurs when they are able to take an action or when they are able to collaborate with their team members who perceive and act on things from a different perspective.

6 Conclusions

From our above analyses, we conclude three main points:

Firstly, if a VR application is not embedded in a storyline, it can lack the emotional engagement that comes from feeling invested in the characters and their journeys. A well-written and compelling storyline can create a sense of immersion and connection to the environment and its agents, making the experience more engaging and memorable, and – importantly – making language

learning effortless as both negotiation for meaning and action (Zheng et al., 2009) is required to proceed in the storyline. Without this emotional connection, meaning remains to be information that stays at the cognitive domain. Good experiences involve mind, body, and heart (Ames, 2011). When learners are holistically immersed in the learning environment for which VR plays a part due to its technological affordances as we identified in Table 1, we can say language learning is sustainable.

Secondly, multiuser VR applications, as opposed to single user applications, can create a more immersive and realistic language-learning experience by allowing players to interact with other human beings in addition to interacting with the virtual environment. These interactions can mimic real-life interactivity and provide opportunities for language activity among real people. Additionally, these interactions can foster a sense of social presence, which can further enhance the emotional connection to the task and the language being learned.

Thirdly, most VR games do not fully exploit the affordances of VR technology for language learning, often focusing on single aspects of language learning such as vocabulary or grammar. This approach can fall short in creating meaningful and realistic situations that can foster a deeper understanding of the language. To truly benefit from the immersive nature of VR, games should aim to create rich, authentic narratives that can provide opportunities for languaging and negotiation of meaning, rather than just isolated linguistic exercises.

Finding this large number of recent research projects published surrounding the topic of VR on language learning indicates that VR technology has become more and more accessible to education. However, while acknowledging the effort of carrying out a research project that requires large scale planning in terms of budgets, selecting appropriate VR technology, and designing tasks and activities, we were disappointed to find that, to date, VR technology is not revolutionizing language learning the way it was promised. Throughout our analysis, we were able to identify problems in terms of the gap between possible VR affordances for language learning and current research limitations.

For language learning researchers, we suggest that the first and most important step in developing a VR language learning project might be to experience a VR application themselves. Then, they should formulate a research agenda, integrating first person experience with a theoretical framing and methodology.

If first-person experiencing is not possible, we suggest to carefully study VR application features first, and then design activities, tasks, and objects to optimize the use of these features. Based on our review and analysis, most studies involved a design of certain tasks to some extent, however, the designs were not saliently described, nor were they accounted for as the outcome variables. As result, studies in

forms of survey research mostly asked questions such as the perceived usefulness of VR, which does not allow us to reach a reliable understanding of *how* VR can exchange language learning.

Virtual environment technology itself without designed elements either in forms of activities and tasks or physical layout cannot provide affordances for learning, interaction, or coordination (Zheng, 2012; Zheng & Newgarden, 2012). Without these elements, we are not able to make a bridge between research findings on second language learning and pedagogical intervention, which leads to little contribution to theory building. Furthermore, without describing the design and rationalizing the relationship between design and outcome variables, researchers will not be able to test how to exploit VR technology for learning, in other words, optimizing the features of VR to create new learning experiences as well as new theories for understanding such new learning.

More specifically, responding to the paucity of theoretical building in VR affordances for language learning, we will provide solutions to aforementioned problems of overuse of VR for minimal effects. For example, to fully actualize an information gap task (York et al., 2021) in a VR environment, York and their co-authors' task can be modified to manipulate objects rather than just telling the differences between two observed rooms. The purpose of manipulation can be briefly addressed in the task design, which leaves room for learners to explore how the ability of being able to see different perspectives of the same room in HTC Vive affords language use. Consequently, the manipulation can be studied using design-based research (Cui et al., 2022; Reeves & McKenney, 2013) to tease out exactly what kinds of manipulation lead to the desired learning outcome. By the same token, manipulatary action and elaboration of the design by Urueta and Ogi (2020) can be enhanced with narratives that engage learners emotionally, and researchers can test which specific features of VR technology can contribute to the storylines to increase motivation and engagement in the learning process. Eventually, the goal is to create new understanding of how VR enhanced storylines afford language learning and lead to new theories of language learning in light of VR technologies.

References

Ahlers, T., Bumann, C., Kölle, R., & Lazović, M. (2021). Hololingo! – a game-based social virtual reality application for foreign language tandem learning. In A. Kienle, A. Harrer, J. M. Haake & A. Lingnau (Eds.), *DELFI 2021* (pp. 37–48). Gesellschaft für Informatik e.V.

Ames, R. T. (2011). *Confucian role ethics: A vocabulary*. The Chinese University Press.

Barab, S. A., Gresalfi, M., Dodge, T., & Ingram-Goble, A. (2010). Narratizing disciplines and disciplinizing narratives: Games as 21st century curriculum. *International Journal of Gaming and Computer-Mediated Simulations*, *2*(1), 17–30.

Barab, S. A., Sadler, T. D., Heiselt, C., Hickey, D., & Zuiker, S. (2007). Relating narrative, inquiry, and inscriptions: Supporting consequential play. *Journal of Science Education and Technology*, *16*(1), 59–82.

Barrett, A., Pack, A., Guo, Y., & Wang, N. (2020). Technology acceptance model and multi-user virtual reality learning environments for Chinese language education. *Interactive Learning Environments*, *31*(3), 1–18.

Bermejo, C., & Hui, P. (2021). A survey on haptic technologies for mobile augmented reality. *ACM Computing Surveys*, *54*(9), 1–35.

Cai, J. Y., Wang, R. F., Wang, C. Y., Ye, X. D., & Li, X. Z. (2021). The influence of learners' cognitive style and testing environment supported by virtual reality on English-speaking learning achievement. *Sustainability*, *13*(21), 1–14.

Collins, N., Vaughan, B., Cullen, C., & Gardner, K. (2019). Gaeltechvr: Measuring the impact of an immersive virtual environment to promote situated identity in Irish language learning. *Journal for Virtual Worlds Research*, *12*(3) 1–16

Cowley, S. J., & Fester, M. T. (2017). Voices in human agency. The power of presence. *Reti, saperi, linguaggi*, *4*(2), 303–320.

Cui, Z., Dong, J., Liu, Y., Tangiyev, D. M., Nokes, J., Nuesser, M., Tang, S., Zhang, H. & Zheng, D. (2022). Design of a distributed language learning environment: Contextualizing Chinese language learning in a Panda reserve virtual reality. In Y.-J. Lan & S. Grant (Eds.), *Contextual Language Learning: Real Language Learning on the Continuum from Virtuality to Reality* (pp. 39–70). Springer.

DeWitt, D., Chan, S. F., & Loban, R. (2022). Virtual reality for developing intercultural communication competence in Mandarin as a Foreign language. *Educational Technology Research & Development*, *70*(2), 615–638.

Ehrsson, H. H., Holmes, N. P., & Passingham, R. E. (2005). Touching a rubber hand: Feeling of body ownership is associated with activity in multisensory brain areas. *Journal of Neuroscience*, *25*(45), 10564–10573.

Garcia, S., Laesker, D., Caprio, D., Kauer, R., Nguyen, J., & Andujar, M. (2019). An immersive virtual reality experience for learning Spanish. In *Learning and Collaboration Technologies. Ubiquitous and Virtual Environments for Learning and Collaboration: 6th International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, proceedings, part II 21* (pp. 151–161). Springer.

Godwin-Jones, R. (2021). Evolving technologies for language learning. *Language Learning & Technology*, *25*(3), 6–26.

Han, L. (2022). Students' daily English situational teaching based on virtual reality technology. *Mobile Information Systems*, *2022*, 1–13.

Hellermann, J., Thorne, S. L., & Haley, J. (2019). Building socio-environmental infrastructures for learning. In J., Hellermann (Ed.), et al., *Analytic Research on Learning in Action: The Complex Ecology of Second Language Interaction "In the Wild"* (pp. 193–218). Springer.

Herrera, F., Bailenson, J., Weisz, E., Ogle, E., & Zaki, J. (2018). Building long-term empathy: a large-scale comparison of traditional and virtual reality perspective-taking. *PLoS One*, *13*(10), 1–37.

Holden, C. L., & Sykes, J. M. (2011). Leveraging mobile games for place-based language learning. *International Journal of Game-Based Learning (IJGBL)*, *1*(2), 1–18.

Jia, T., & Liu, Y. (2019). Words in Kitchen: An instance of leveraging virtual reality technology to learn vocabulary. In *2019 iEEE international symposium on mixed and augmented reality adjunct (ISMAR-Adjunct)* (pp. 150–155). IEEE.

Johnson, M. (2017). *Embodied mind, meaning, and reason: How our bodies give rise to understanding*. University of Chicago Press.

Kang, Y. J., & Lee, W. J. (2018). Effects of sense of control and social presence on customer experience and e-service quality. *Information Development*, *34*(3), 242–260.

Lan, Y. J. (2014). Does second life improve Mandarin learning by overseas Chinese students? *Language Learning & Technology*, *18*(2), 36–56.

Lan, Y. J. (2020). Immersion, interaction and experience-oriented learning: Bringing virtual reality into FL learning. *Language Learning & Technology*, *24*(1), 1–15.

Lin, T. J., & Lan, Y. J. (2015). Language learning in virtual reality environments: Past, present, and future. *Journal of Educational Technology & Society*, *18*(4), 486–497.

Lin, V., Barrett, N. E., Liu, G. Z., Chen, N. S., & Jong, M. S. Y. (2021). Supporting dyadic learning of English for tourism purposes with scenery-based virtual reality. *Computer Assisted Language Learning*, *36*(5–6), 1–37.

Mandler, J. M. (2007). *Foundations of mind: Origins of conceptual thought*. Oxford University Press.

Markowitz, D. M., Laha, R., Perone, B. P., Pea, R. D., & Bailenson, J. N. (2018). Immersive virtual reality field trips facilitate learning about climate change. *Frontiers in Psychology*, *9*, 1–20.

Reeves, T. C., & McKenney, S. (2013). Computer assisted language learning and design-based research: Increased complexity for sure, enhanced impact perhaps. In *Design-based research in CALL* (pp. 9–21). CALICO.

Ryan, M. L. (2015). *Narrative as virtual reality 2: Revisiting immersion and interactivity in literature and electronic media*. Johns Hopkins University Press.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied linguistics*, *11*(2), 129–158.

Sherman, W. R., & Craig, A. B. (2018). *Understanding virtual reality: Interface, application, and design*. Morgan Kaufmann.

Steuer, J. (1995). Defining virtual reality: Dimensions determining telepresence. In F. Biocca & M. R. Levy (Eds.), *Communication in the Age of Virtual Reality* (pp. 33–56). Lawrence Erlbaum.

Sun, C., Yao, Y., Wang, R., & Ye, X. (2020). A study on the influence of scene reality of VR environment on English learners' learning engagement and learning effectiveness. In *2020 IEEE 2nd International Conference on Computer Science and Educational Informatization (CSEI)* (pp. 181–185). IEEE.

Urueta, S. H., & Ogi, T. (2020). A TEFL virtual reality system for high-presence distance learning. In L. Barolli, H. Nishino, T. Enokido & K. Takizawa (Eds.), *Advances in Networked-based Information Systems: The 22nd International Conference on Network-Based Information Systems (NBiS-2019)* (pp. 359–368). Springer.

Van Lier, L. (2000). From input to affordance: Social-interactive learning from an ecological perspective. In L. P. J. M. De Bot, R. B. Ginsberg & C. Kramsch (Eds.), *Sociocultural Theory and Second Language Learning* (pp. 245–259). Oxford University Press.

Walsh, K. R., & Pawlowski, S. D. (2002). Virtual reality: A technology in need of IS research. *Communications of the Association for Information Systems*, *8*(1), 297–313.

Wei, L. (2018). Translanguaging as a practical theory of language. *Applied linguistics*, *39*(1), 9–30.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3), 225–240.

York, J., Shibata, K., Tokutake, H., & Nakayama, H. (2021). Effect of SCMC on foreign language anxiety and learning experience: A comparison of voice, video, and VR-based oral interaction. *ReCALL*, *33*(1), 49–70.

Zheng, D. (2012). Caring in the dynamics of design and languaging: Exploring second language learning in 3D virtual spaces. *Language Sciences*, *34*(5), 543–558.

Zheng, D., Liu, Y., Lambert, A., Lu, A., Tomei, J., & Holden, D. (2018). An ecological community becoming: Language learning as first-order experiencing with place and mobile technologies. *Linguistics and Education*, *44*, 45–57.

Zheng, D., & Newgarden, K. (2012). Rethinking language learning: Virtual worlds as a catalyst for change. *International Journal of Learning and Media*, *3*(2), 13–16.
Zheng, D., Schmidt, M. M., Hu, Y., Liu, M., & Hsu, J. (2017). Eco-dialogical learning and translanguaging in open-ended 3D virtual learning environments: Where place, time, and objects matter. *Australasian Journal of Educational Technology*, *33*(5), 107–122.
Zheng, D., Young, M. F., Wagner, M. M., & Brewer, R. A. (2009). Negotiation for action: English language learning in game-based virtual worlds. *The Modern Language Journal*, *93*(4), 489–511.

[Post online publication correction added on 7 May 2025: journal title "*Language, Learning and Technology*" was updated to "*Language Learning & Technology*" throughout the reference list.]

Bionotes

Michaela Nuesser
Department of Second Language Studies, University of Hawai'i at Mānoa, Honolulu, HI, USA
mnuesser@hawaii.edu
https://orcid.org/0000-0003-1667-6921

Michaela Nuesser is a PhD candidate in Second Language Studies from University of Hawai'i at Mānoa. Her research interests include language teaching and technology, distributed language, and quantitative and qualitative approaches to linguistic relativity.

Dongping Zheng
Department of Second Language Studies, University of Hawai'i at Mānoa, Honolulu, HI, USA
zhengd@hawaii.edu

Dongping Zheng received her PhD in Educational Psychology from the University of Connecticut and is currently an associate professor at the University of Hawai'i at Mānoa. Her research interests include distributed language, ecolinguistics, design-based research, virtual reality, humane AI, and mobile-enabled place-based learning. Her forthcoming co-authored book is 分布认知视阈下大学英语教学生态环境的研究与构建 (*Ecological niche construction for college English education research: A distributed cognition perspective*), Northeast Normal University Publishing House.

Jin Dong
Department of East Asian Studies, Princeton University, Princeton, NJ, USA
jd1398@princeton.edu

Jin Dong is currently a PhD student in the Department of East Asian Languages and Literatures at the University of Hawai'i at Mānoa, and a lecturer at Princeton University. Her research interests include (trans)languaging, distributed language, virtual world/reality, and game-assisted language learning.