Language Learning in Virtual Reality Environments: Past, Present, and Future

Tsun-Ju Lin[1] and Yu-Ju Lan[2]*

[1]Department of Applied Chinese, Kainan University, Taoyuan City, Taiwan // [2]Department of Applied Chinese Language and Culture, National Taiwan Normal University, Taipei, Taiwan // tsunjulin@gmail.com // yujulan@gmail.com

*Corresponding author

ABSTRACT

This study investigated the research trends in language learning in a virtual reality environment by conducting a content analysis of findings published in the literature from 2004 to 2013 in four top ranked computer-assisted language learning journals: *Language Learning & Technology*, *CALICO Journal*, *Computer Assisted Language Learning*, and *ReCALL*. Data from 29 articles were cross-analyzed in terms of research topics, technologies used, language learning settings, sample groups, and methodological approaches. It was found that the three most popular research topics for learners were interactive communication; behaviors, affections, and beliefs; and task-based instruction. However, the analysis results highlight the need for the inclusion of the impact of teacher. The data also revealed that more studies are utilizing triangulation of measurement processes to enable in-depth analysis. A trend of gathering data through informal learning procedures was also observed. This article concludes by highlighting particular fields related to VR in which further research is urgently needed.

Keywords

Virtual reality, MMOGs, Virtual worlds, Content analysis

Introduction

The recent rapid advances in and popularity of wireless communication and multimedia environments have resulted in language learning in a virtual reality (VR) environment [a so-called virtual learning environment (VLE)] receiving considerable attention in the past few years. In the early stage, VR generally involved only small representations of content area or domains designed to "bridge the gap between reality and abstract knowledge by the discovery method" (Lee, 1999, p. 72). In order to create such environments, the systems were commonly theme-based and comprised a full integration of artificial intelligence and a wide variety of social communication tools. Because of its features, VR was quickly introduced to promote authentic and immersive learning environments.

VR tools nowadays are a far cry from early two-dimensional (2D) text-based online VR environments, such as multiuser domains (MUDs) and MUD, object oriented (MOO), to which multiple users can be connected at the same time. Current VLEs have evolved to be more sophisticated and interactive with a high degree of visual appeal, allowing a wide range of interlocutors to communicate, cooperate, and compete through customized three-dimensional (3D) virtual spaces and avatars. Currently there are several VR tools in the market; Sykes et al. (2008) categorized VR based on its original design purposes into three types: open social virtualities, massively multiplayer online games (MMOGs), and synthetic immersive environments (SIEs). Open social virtualities, such as *Second Life* (*SL*), *OpenSimulator*, and *Active Worlds* (*AW*), allow the users to immerse themselves in a wide variety of social contexts, participating in individual and group activities, and letting them create and trade virtual properties and services with one another. MMOGs, such as *World of Warcraft* (*WoW*), emphasize more role-playing functions, and involve hundreds or even thousands of players cooperating and competing with each other simultaneously based on their selected race, class, or profession. However, Sykes et al. (2008) considered that the above two types of VR were originally developed for commercial and business use. In contrast, an SIE emphasizes its educational objectives by carefully incorporating pedagogy into the immersive spaces. Regardless of the original purposes for which VLEs were designed, researchers in the computer-assisted language learning (CALL) field have tried to employ pedagogical principles and practices that are innovative and theoretically grounded to understand the pedagogical values of VLEs in language learning.

Given the above background on how VR emerged and has been categorized, it is useful to define VR so as to reflect its development as a state-of-the-art technology. Based on a thorough review of the definitions proposed by other

researchers, Smart et al. (2007) offered the following core definition of VR: a system that aims to bring simulated real-life experiences, providing topography, movement, and physics that offer the illusion of being there.

Review of VR and simulation research

Several journal articles conducting meta-analysis have attempted to identify the characteristics and challenges of VR in technology-enhanced learning research. Lee (1999) analyzed 19 studies to determine the effectiveness of computer-simulated environments by investigating the relationship between 2 forms (i.e., pure and hybrid) and 2 modes of instructions (i.e., presentation and practice). That study revealed that within either the presentation or practice mode, hybrid simulation (which has expository instructional features) is much more effective than pure simulation (which does not integrate any specific directions and explanations). Schwienhorst (2002) examined the state of VR in second-language (L2) acquisition with the aim of identifying the common problems encountered and the role of VR in CALL research. Researchers in the 1990s seemed to agree that VR can promote constructivist activities and learner engagement (i.e., student involvement, self-monitoring, and self-evaluation). Additionally, VR can lead to "intrinsic motivation, more intercultural awareness, and a reduction of the affective filter" (Schwienhorst, 2002, p. 230). However, Schwienhorst (2002) claimed that VR research during that period was inadequate to allow any conclusive claims to be made.

Subsequent studies seemed to switch the focus to the effects of instruction involving computer simulations versus traditional instruction methods. Vogel et al. (2006) reviewed 32 studies and found that across all variables (e.g., age and gender) and situations (e.g., types of activities and the degree of image realism in the computer programs), the use of interactive simulations and games provided superior cognitive outcomes and more positive attitudes toward learning than did traditional teaching methods. Similar to previous studies, they also found that interactive simulations promoted self-directed learning and provided fail-safe learning environments. The authors also demonstrated that interactive simulations and games increase learners' knowledge more than traditional teaching methods across all ages. Liao and Chen (2007) examined 29 studies performed in Taiwan and also found that instruction involving computer simulations had more positive effects on learning by students than did traditional instruction methods. Their study further suggested that 3D virtual presentation may be more effective than 2D for students' learning due to the former providing more accurately simulated and authentic learning environments.

It should be noted that among the four above-mentioned studies, only Schwienhorst (2002) emphasized the use of virtual worlds for language learning. Moreover, although Schwienhorst (2002) claimed to have conducted a meta-analysis, the article was actually more of a historical overview of relevant research and aimed to shed light on the use of VR in L2 acquisition. This is understandable given that in 2002 there were too few language studies involving VR to allow a meta-analysis to be conducted (aimed at establishing statistical significance with studies on a particular topic that have produced conflicting results); moreover, the number of examined studies was not mentioned in his article.

The subsequent decade saw a focus on the research trends in technology-based learning. Hsu et al. (2012) conducted a content analysis on five SSCI journals published from 2000 to 2009, with the results revealing that the proportion of articles on VLEs, digital games, and learning enhanced through the use of "intelligent" toys increased significantly between 2000–2004 and 2005–2009, from 0.81% to 3.82%. However, VR is still reported to be one of the least published research topics in the technology-based learning field. Similar to Hsu et al. (2012), Wang and Vásquez (2012) also reported VR as being less frequently explored technologies compared to other Web 2.0 tools in L2 acquisition.

The above-described results indicate that studies that analyze the appropriateness, application, and practices of VR, and its influence on language education are urgently needed. This situation prompted the present study to systematically review and synthesize the literature on language learning in VR from 2004 to 2013 in order to determine whether this research field has produced any conclusive data during the past 10 years. Reviewing the trends in VR may help to identify research interests and gaps, and further provide reference data for future research directions. To identify the overall research trends, this study categorized previous studies into five categories: research topics, technologies used, language learning settings, sample groups, and methodological approaches. The following four research questions (RQs) were specifically addressed by this study:

RQ1. What percentage of the articles published in the selected journals were related to VR?

RQ2. What topics related to language learning in VR were investigated in these journals from 2004 to 2013?
RQ3. What technologies have been used in VR studies?
RQ4. What methodologies were applied in VR studies from 2004 to 2013? How did the methodologies change from 2004 to 2013?

Method

The studies to be included in the content analysis were identified by performing computer searches of journal databases. Since it would have been difficult to include all relevant journals in the analysis, the following four top CALL-specific and education technology-related journals were selected: *Language Learning & Technology* (*LLT*), *CALICO Journal*, *CALL*, and *ReCALL*. Smith and Lafford (2009) evaluated these as the four highest quality English-language journals in the field of CALL. The examined journals were evaluated by 35 tenured CALL experts using a list of criteria, including the quality of the articles, significance of contribution to the field, review process, and originality. It was expected that a thorough analysis of high-quality articles in these 4 journals by the 35 experts would identify the most influential works in the CALL field.

In total, 811 articles appeared in the 4 journals from 2004 to 2013; this excluded columns, commentaries, book/media reviews, review studies, editorial materials, and letters. A content analysis was applied to these articles using the following keywords: simulation, VR, VLEs, social virtual worlds, MOOs, MUDs, and MMOGs. In a few cases where the keywords were insufficient, the researchers analyzed abstracts and full articles. This process yielded 29 empirical studies for inclusion in the analysis in this study. Based on the methodology of content analysis, descriptive statistics was utilized to classify the previously mentioned five categories, namely research topics, technologies used, language learning settings, research sample groups, and methodological approaches.

Research topics

Based on the methodology of inductive content analysis, the research topics related to L2 acquisition in CALL were first classified into four categories: learner differences, learning task, impact of the teacher, and environment. For each category, subcategories were classified. During the data analysis process, these categories and subcategories were refined continuously. The validity of these categories and subcategories was confirmed after all the articles were reviewed, as follows:

- *Learner differences*: Articles related to this research topic explore learner's linguistic and sociolinguistic competencies. Some articles also discussed the concept of metalinguistic awareness in terms of how a learner integrates form, meaning, and function to keep oneself using the normal target language. This category was further divided into four subcategories: interactive communication, motivation, cultural awareness and intelligence, and language proficiency.
- *Learning task:* This category includes articles that discuss or examine different types of instructional strategies and approaches used in VR. This category was further divided into three categories: task-based instruction (TBI), collaborative learning, and problem-solving.
- *Impact of the teacher:* This category includes studies that outline teachers' teaching experiences and perceptions in MUVEs. The two subcategories comprise task design, and perception and awareness.
- *Environment:* Articles that emphasize investigating the affordance of a VR environment were classified into this category.

Technologies used, language learning settings, research sample groups, and methodological approaches

First of all, the three types of VR proposed by Sykes et al. (2008) were extended to the following subcategories of research technologies used in this review: open social virtualities, MMOGs, SIEs, "others," and "not specified."
Second, after conducting several rounds of discussions, the following six subcategories were identified for the language learning settings: first language (L1) learning, L2 learning, foreign language (FL) learning, language exchange, "others," and "not specified." Third, discussions among researchers reached the following consensus for the eight categories for the research sample groups: elementary school, junior/senior high school, higher education, teachers, adults, special needs, "others," and "not specified." It should be noted that the "adults" sample group

included all adult participants without specifying whether they were teachers or participants in higher education. Finally, the same researchers discussed the usage of the subcategories for the methodological approaches, which produced the following finalized subcategories: quantitative, qualitative, and a combination of quantitative and qualitative approach (quan.+qual.). The data collection for each article is described later.

Data coding and analysis

This study used the articles relevant to VR published in the four selected journals from 2004 to 2013 to investigate the research trends. Content analysis was utilized to classify the aforementioned categories.

A one-level coding process was used to categorize technologies used, language learning settings, sample groups, and research designs. Articles with only one major subcategory were coded to the designed subcategory, while those with more than one major subcategory or with subcategories other than the aforementioned ones were accounted as "others."

As mentioned above, many of the analyzed articles addressed more than one research topic. Thus, the research topic was coded into two levels, with the primary and matched subcategories coded as the first and second research topics, respectively. Any matching subcategories after the second research topic were excluded, and the field of the second research topic was coded "none" for articles with only one research topic.

The coding process was undertaken manually by the researchers. To obtain more reliable outcomes from coding, three researchers (two professors and one research assistant) in educational technology helped to code these studies based on the aforementioned categories. These researchers each had from 5 to 12 years of language teaching experience, knew more than one language, and had taught online courses in more than two different countries previously. All the articles followed the same coding process. The results were compared using Pearson correlation coefficient measures of reliability, which yielded $r = 0.84$. Any discrepancies were resolved by discussion. After the initial coding process, a descriptive analysis was conducted to report the data.

Results

Percentage of articles related to VR

The analysis by year and journal (see Table 1) revealed that only 3.6% of the articles published in the four journals from 2004 to 2013 were related to the field of VR: 0.8% in *LLT*, 1.9% in *CALICO Journal*, 4.7% in *CALL*, and 6.3% in *ReCALL*. The year with the highest percentage of related articles was 2012 (10.3%), while no related articles were published in 2004 and 2007. Figure 1 illustrates the increasing trend of articles related to VR in the four journals from 2004 to 2013.

Table 1. Percentage of VR studies in each journal during the analyzed period

|	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
LLT	0/13 (0)	0/14 (0)	0/12 (0)	0/13 (0)	0/12 (0)	0/12 (0)	0/10 (0)	1/12 (8.3)	0/11 (0)	0/19 (0)	1/128 (0.8)
CALICO	0/21 (0)	2/24 (0)	1/24 (4.1)	0/22 (0)	0/29 (0)	2/34 (0.6)	0/31 (0)	0/31 (0)	0/26 (0)	0/20 (0)	5/262 (1.9)
CALL	0/21 (0)	1/22 (4.5)	1/19 (5.2)	0/25 (0)	1/27 (3.7)	2/23 (8.7)	0/24 (0)	2/25 (8)	2/23 (8.7)	2/24 (8.3)	11/233 (4.7)
ReCALL	0/34 (0)	0/16 (0)	0/12 (0)	0/18 (0)	0/20 (0)	1/18 (5.6)	1/20 (5)	0/13 (0)	6/17 (35.2)	4/20 (20)	12/188 (6.3)
Total	0/89 (0)	3/76 (3.9)	2/67 (3)	0/78 (0)	1/88 (1.1)	5/87 (5.7)	1/85 (1.2)	3/81 (3.7)	8/77 (10.3)	6/83 (7.2)	29/811 (3.6)
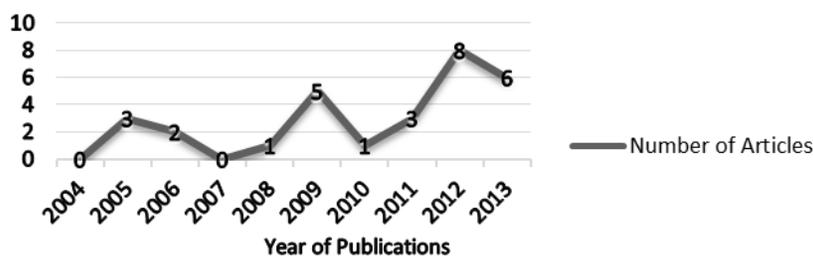
Figure 1. Publication trends from 2004 to 2013

Research topic analysis

As presented in Table 2, the most common research topic was learner differences ($n = 27$), followed by learning task ($n = 9$). On the other hand, the most common research topic subcategory was interactive communication ($n = 12$), followed by learner behaviors, affections, and beliefs ($n = 9$) and TBI ($n = 6$). The least common research topic subcategory was collaborative learning ($n = 1$).

Table 2. Thematic foci

Category	Subcategory	Frequency
A. Learner differences	A1. Interactive communication	12
	A2. Behaviors, affections, and beliefs	9
	A3. Cultural awareness and intelligence	3
	A4. Language proficiency	3
B. Learning task	B1. TBI	6
	B2. Collaborative learning	1
	B3. Problem-based learning	2
C. Impact of the teacher	C1. Task design	3
	C2. Perception and awareness	3
D. Environment	D1. Affordance	5

Note. In order to provide in-depth discussions, studies covered two or more categories were identified, and so the total frequency counts did not match the number of the reviewed articles.

An analysis of technologies used

The content analysis revealed that the technology used most often was open social virtualities (65.6%), while the least common were SIE (3.4%) and those in the "not specified" category (3.4%) (see Table 3). Research technologies used in open social virtualities from 2004 to 2008 employed various platforms, such as 2D text-based MOO ($n = 2$) and *AW* ($n = 2$). From 2009 to 2013, *SL* ($n = 12$) became the most utilized platform. *AW*, *Blue Mars Lite*, and *Wonderland* were each used once only. The results also showed that researchers demonstrated more interest in the MMOG type of VR platforms, such as *WoW* ($n = 3$) and *Civilization* ($n = 1$). While only one platform, *Quest Atlantis*, was classified as an SIE, four platforms were identified as "others." These platforms in the "others" category can be divided into 3D computer graphic/game development programs (e.g., *Unity* and *Autodesk 3Ds MAX*) and life simulation video games without a specified multiuser function (e.g., *The Sims* and *Sim Theme Park*).

Table 3. Percentage of articles related to specific MUVE technologies

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	*N* (%)
Open social virtualities	0	2	2	0	0	3	1	3	3	5	19 (65.6)
MMOGs	0	0	0	0	0	0	0	0	3	1	4 (13.8)
SIEs	0	0	0	0	0	1	0	0	0	0	1 (3.4)
Others	0	1	0	0	1	1	0	1	0	0	4 (13.8)
Not specified	0	0	0	0	0	0	0	0	1	0	1 (3.4)

The trends of VR in research methodology

RQ4 was answered by analyzing the language learning settings, sample groups, and methodological approaches (see Table 4). It is believed that dividing the analysis of applied methodologies into two five-year-periods (i.e., 2004–2008 and 2009–2013) will provide a better reference to help researchers make plans in the future. Therefore, the changes between 2004–2008 and 2009–2013 are particularly illustrated in this section. From 2004 to 2008, the language learning settings for language exchange ($n = 2$) and FL ($n = 2$) were utilized most, followed by L2 ($n = 1$), "others" ($n = 1$), and L1 ($n = 0$). From 2009 to 2013, the most common language learning setting was FL ($n = 16$), followed by "others" ($n = 6$) and L2 ($n = 1$). The number of articles on L1 learning remained 0 in second period, and on language exchange dropped to 0. Furthermore, the percentage of articles on FL increased from 33% to 70% between the two periods, and the language learning formats in the "others" category also increased from 17% to

26%. It is noted that the language learning formats in the "others" category were the target language for special purposes (e.g., business, architecture, and cultural awareness) or a combination of FL, L2, and L1.

Table 4. Analysis of the selected articles related to the study of VR

Year	Author(s)	Language learning settings/participants	Methodological approach	Data collection
2005	Hansson	FL. 10 7th-grade pupils	Qualitative	Text chats
	O'Rourke	Language exchange. 34 German and 24 Irish college students	Quan.+qual.	Text chats
	Rilling et al.	Others. 4 preservice language teachers	Qualitative	Learner feedback, teachers' self-reflection
2006	Peterson	FL. 24 college English learners in Japan	Qualitative	Text and audio transcripts, observation, questionnaires
	Schwienhorst & Borgia	Language exchange. 34 German and 26 Irish college students in 2000-2001. 18 German and 12 Irish college students in 2002-2003	Quantitative	Text chats, course statistical data
2008	Ranalli	L2. 9 intermediate-level English learners at a university	Quan.+qual.	Testing, questionnaire
2009	Deutschmann et al.	FL. A comparison of 2 oral proficiency courses aimed at doctoral students	Quan.+qual.	Audio recordings, questionnaires, group evaluations, personal interviews
	Ho et al.	L2. 45 12th-grade students	Qualitative	Screen recordings, questionnaires, interviews
	O'Brien et al.	FL. 1st semester: 42 German high-school students FL. 3rd semester: 33 German high-school students	Quantitative	Questionnaires
	Peterson	FL. 14 college English learners in Japan	Qualitative	Questionnaire, text chats, observations
	Zheng et al.	FL. 61 middle-school students in China	Quantitative	Testing, survey
2010	Peterson	FL. 7 college English learners in Japan	Qualitative	Questionnaire, text chats
2011	Collentine	FL. 58 college Spanish learners	Quan.+qual.	Text chats, user-tracking data (learner's movements, actions, and choices)
	Jauregi et al.	FL. 2 Spanish learners & 2 preservice teachers in universities	Qualitative	Verbal transcripts and interaction data, questionnaires, interviews
	Wehner et al.	FL. 40 Spanish learners at 1 university (*SL*, $n = 20$; non-*SL*, $n = 20$)	Quantitative	Survey
2012	Cornillie et al.	FL. 83 first-year university students and learners in high school	Quan.+qual.	Questionnaires, interviews, game logs
	Liang	FL. 11 college English learners	Quan.+qual.	Students' verbal and chat transcripts
	Liou	FL. 25 potential English teachers	Qualitative	Questionnaires, interviews
	Peterson(a)	FL. 8 college English learners in Japan	Qualitative	Text chats, observation, interviews, questionnaire
	Peterson(b)	FL. 4 college English learners in Japan	Qualitative	Text chats, observation, questionnaires, interviews
	Rama et al.	FL. 6 college English learners	Qualitative	Observation, interviews, chat

				transcripts, journal entries
	Thorne et al.	Others. 32 Dutch & 32 American *WoW* gamers	Qualitative	In-/out-of-game texts, interviews, questionnaire
	Zheng et al.	Others. Three college-age English learners	Qualitative	Verbal transcripts and interaction data, observation
2013	Canto et al.	FL. 36 Spanish learners at the University of Utrecht and 14 preservice teachers at the University of Valencia	Quantitative	Observation, verbal testing, questionnaires
	Dooly & Sadler	Others. Student-teachers studying to become language teachers in Spain and the USA	Qualitative	Classroom and screen recordings, verbal and text transcripts, forum posts, email exchanges, self-evaluation, field notes, etc.
	Ryu	Others. 6 adults whose L1 is not English	Qualitative	Interaction data, observation, interviews
	Shih	FL. 4 college English learners in Taiwan	Quan.+qual.	Cultural knowledge tests, interviews, observation, blog entries
	Wigham & Chanier(a)	Others. 17 architecture college students	Qualitative	Screen recordings, observation, verbal chats, questionnaires
	Wigham & Chanier(b)	Others. 17 architecture college students	Quantitative	Screen recordings, text and audio transcripts

Table 4 indicates that from 2004 to 2008 the most common research sample groups were in higher education ($n = 4$), followed by teachers ($n = 1$) and elementary school ($n = 1$). There were no research groups in the junior/senior high school, adults, special needs, and "others" categories during the initial period. From 2009 to 2013, research samples in higher education were still used for most of the VR research papers ($n = 13$), followed by "others" ($n = 4$), junior/senior high school ($n = 3$), teachers ($n = 2$), and adults ($n = 1$), with none for elementary school ($n = 0$) and special needs ($n = 0$). Furthermore, declining trends were evident between the two periods in the elementary school (from 17% to 0%), higher education (from 66% to 57%), and teachers (from 17% to 9%) research subcategories, while researchers demonstrated increased interest in "others" (from 0% to 17%), junior/senior high school (from 0% to 13%), and adults (from 0% to 4%).

The results indicated that from 2004 to 2008, the quantitative methodological approach was found in most of the VR publications ($n = 3$), followed by the quan.+qual. ($n = 2$) and quantitative ($n = 1$) approaches. From 2009 to 2013, the top ranked research design was still qualitative ($n = 13$), followed by quan.+qual. ($n = 6$), and quantitative ($n = 4$). The research interest in quan.+qual. decreased from 33% to 26% between the two periods, while that in the qualitative methodological approach increased from 50% to 57%, while the quantitative research approach remained at 17%.

Discussion

This study aimed to identify the trends in VR research articles in terms of their research topics, technologies used, language learning settings, sample groups, and methodological approaches. The results are discussed below.

VR trends in selected journals and research topics

The findings show that the number of VR studies of language learning gradually increased from 2004 to 2013. This is consistent with Hsu et al. (2012) reporting that the number of articles on technology-based learning concerning "digital games and intelligent toy enhanced learning" increased significantly between 2000–2004 and 2005–2009 (from 0.81% to 3.82%).

The content analysis revealed that the most common research topic from 2004 to 2013 was learner differences—interactive communication. Previous VR research has investigated the role of metalinguistic awareness (O'Rourke, 2005), the kind of interaction management strategies employed by learners (Peterson, 2006), and the factors influencing the use of these strategies in VLEs (Peterson, 2009). Based on this valuable groundwork, some follow-up studies have suggested that 3D VLEs affect learner behaviors in terms of linguistic complexity, accuracy, and correct feedback (e.g., Thorne et al., 2012; Wigham & Chanier, 2013b). As more 3D technologies were developed, the use of nonverbal communication (e.g., the modalities of proxemics, kinesics, and avatar appearance) has been highlighted to overcome verbal miscommunication in virtual worlds (Wigham & Chanier, 2013a). It is expected that future research will focus on examining the effectiveness of learner interactive communication in different VR contexts.

Additionally, learner behaviors, affections, and beliefs have also received considerable attention between the two 5-year periods. The related studies have consistently found that VLEs increase learner autonomy and self-efficacy, reduce learning anxiety, and foster creativity. VR presents a realistic virtual space and visible "classmates" who assist students in gaining a sense of participation and building emotional bonds (positive or negative) with their collaborative partners (e.g., Collentine, 2011; Deutschmann et al., 2009; Peterson, 2012b). In contrast, negative perspectives have also been reported, mostly associated with the use of specific VR tools that had user-side and server-side issues. Overall, a growing trend of individual differences and preferences regarding social identities and metacognition in virtual worlds should be expected as more learner control functions become available with the use of these latest advanced technologies (Hsu et al., 2012).

The topics in the learning task—TBI category received particular attention from VR researchers between the two periods. TBI is characterized by activities that are generally theme venues and emphasize engaging learners in meaningful, goal-directed communication to collaboratively solve problems and complete assigned tasks. Several studies found that conducting TBI in VLEs, which involves authenticity and collaborative elements, has a direct impact on learner participation, engagement, and the amount of negotiation (e.g., Peterson, 2012a). Since TBI has a wide range of theoretical groundings (e.g., interactionist theory, sociocultural theories, and experiential learning theory), more research into integrating TBI in VR is expected.

The use of VR tools in CALL

Open social virtualities was found to be the most popular research VR tool, with a consistently high level of interest across the selected studies. Earlier VR research focused more on MOO and *AW*, whereas studies performed between 2009 and 2013 switched to the use of *SL*. This finding suggests that text-based 2D VLEs are no longer able to satisfy the needs of practitioners. Rather, 3D VR tools with the features of multiple communication channels and high visual appeal are better suited to creating tasks with greater degrees of interaction and collaboration, and thus can provide multiple modalities for input, output, and feedback. As a consequence of being widely investigated in the CALL field, we suggest that *SL* provides more useful features for language learning than do other open social virtualities (Warburton, 2009). Additionally, the *ReCALL* special issue in 2012 prompted a sudden increase in the use of MMOG platforms, although this was followed by only one study in 2013; this implies that we are still in the exploration stage of MMOG technology in the CALL field. Future trends should be tracked to see how MMOGs can be used for language education in the future.

Notwithstanding the positive research results mentioned above, several negative attitudes and both user-side and server-side problems were mentioned throughout most of the studies. In terms of user-related issues, novice *SL* users seem to consistently report negative experiences associated with significant investments of time and energy, complicated in-world interfaces, and difficulty in writing scripts and modeling behaviors. Server-side issues included downtime, frequent updates, lag, and large monetary investments (if wishing to be more than a mere spectator) and hardware requirements (e.g., good graphic cards and a high-speed Internet connection) (e.g., Ho et al., 2009; Peterson, 2011). Fortunately, some VR software companies have been collaborating with other companies that operate in different fields of arts and sciences (e.g., topography and physics) in the development of 3D VLEs in order to make them more user-friendly. Also, the suppliers of VR platforms such as *SL* have offered cost discounts for educational and nonprofit institutions. User-related issues could also be solved through long-term training by joining education-focused online communities, such as the *EUROCALL/CALICO Joint Virtual Worlds Special Interest Group* (http://www.eurocall-languages.org/sigs/vw) and the *Virtual Worlds Education Roundtable*

(http://www.vwer.org/). However, these associations are more focused on teachers than on learners. As Hauck and Youngs (2008) claimed, there is a high degree of risk in assuming that students will automatically be familiar with the affordances provided by an online environment. Several studies have claimed that it is essential to provide students with well-organized prestudy technical training to ensure that in-world learning will be effective.

The trends of VR in language learning settings, sample groups, and research designs

The FL setting was found to be the most common in the VR publications analyzed. This is not surprising, given that FL learners often do not have ready access to a suitable environment in which to practice and use the target language. VLEs can overcome this difficulty by providing an immersive and authentic environment to socially interact with native speakers. The data also showed that the "others" category was the second most common research language learning setting, appearing in four out of six articles published in 2013. This implies that language learning environments are becoming more diverse. The learning settings available in cyberspace cannot be divided into traditional ones, such as FL, L2, or L1; rather, a virtual environment can be a combination of all other environments, in which users are not constrained to their physical locations—they can mingle together without physical frontiers.

It is noteworthy that two of the earlier studies investigated whether the use of the MOO medium and tandem exchange—based on mutual language exchange between partners who are native speakers of the target languages—affect on the negotiation of meaning (O'Rourke, 2005), and whether the implementation of MOO improves the balance in bilingualism (Schwienhorst & Borgia, 2006). Surprisingly, this language learning format has not received much attention in recent years, even though the combination of VLEs with tandem courses would make it easy to create language partnerships across countries.

The current study also found that higher education was the sample group utilized most in VR publications, with sample groups in elementary and junior/senior high schools being explored less frequently over the last 10 years. Several factors could explain this research trend. First, this result reflects Rankin, Gold, and Gooch (2006) finding that lower-level language learners experience difficulties dealing with multiple competencies required by the 3D VLEs and can experience cognitive overload. Second, some VLEs have age restrictions, such as 16+ years for *SL* and 13+ years for *WoW*. Most VR platforms are commercial products, and the instructors have little control over who interacts with students outside class times and what locations they visit in a virtual world. Safety therefore remains a concern, especially in K-12 education settings. To avoid these problems, *SL* has divided its regions into general, moderate, and adult regions to avoid adult content being available to users younger than 18 years. Additionally, some software programs have been developed for use by people of all ages by carefully incorporating pedagogy into the immersive spaces (e.g., *AW Educational Universe* and *Croquelandia*).

Finally, the findings reveal that the qualitative methodological approach was dominant in the VR studies, and that nearly one-third of the articles used quan.+qual. techniques. This is a reflection of the complexities involved and the multiple sources of the problems addressed in these studies. Regarding the instrumentation used for data collection, the relevant studies performed over the past 5 years have overcome earlier research deficiencies, such as solely utilizing text chats or transcribed audio tapes, which had resulted in a lack of nonlinguistic and paralinguistic information (O'Rourke, 2005). As presented in Table 4, more recent studies have employed screen recordings or embedded user-tracking systems with triangulated data collection to strengthen their credibility (e.g., Collentine, 2011; Wigham & Chanier, 2013a). Spector et al. (2014) reported that it has become common to employ qualitative or quan.+qual. methods with triangulation techniques to investigate and explain complicated issues encountered in the field of education technology research. Regarding the procedure of data collection, recent attention has emphasized at-home tasks (Liou, 2012) and the beyond-game culture (Ryu, 2013) for language learning in virtual worlds, rather than multichannel in-class data collection. This is consistent with Peterson (2012b) predicting that there will be an increase in the number of education studies focusing on informal outside-class language learning rather than on formal in-class language learning. This situation has resulted in the recruitment of volunteers as research participants directly in VLEs becoming a new trend of data collection (see Ryu, 2013).

Conclusion and future prospects

This study investigated trends in VR research from 2004 to 2013 in the four top CALL journals. The research topics that have been widely investigated during this period (interactive communication; behaviors, affections, and beliefs; and TBI) are consistent with VR being advocated as a promising arena for language learners. The findings reveal that FL educators in higher education are now widely expected to deliver their language instruction with the aid of VR technologies. Regarding the methodologies, qualitative and quan.+qual. methods with triangulated data collection will continue to provide rigorous and in-depth analysis of language learning in VLEs. We believe more advanced and new methods of data collection, such as brain activity and eye tracker data, as well as informal language learning will evolve to more accurately capture in-world language activities.

However, the content analysis performed in this study revealed several areas in which considerable efforts are urgently needed. First, only a small proportion of the included articles examined teachers' perspectives and the awareness of task design. As Peterson (2011) pointed out, a most effective role of the teacher in VLEs has yet to be clarified. As a result, little is known from empirical research about how instructors' roles change in a VR classroom, teachers' decision-making on how to integrate pedagogical activities into VLEs by utilizing the strengths of VR, and how to motivate teachers to adopt and continue using VLEs when teaching (e.g., the method of teacher preparation). Second, regarding the less developed language learning settings and sample groups, studies of language for specific purposes and the tandem principle of bilingualism are worth exploring further. As mentioned above, the TBI approach in theme-based venues can facilitate bilingual language exchanges and language learning for specific purposes (e.g., tourism and medical applications). The positive results obtained when using VR are consistent with 3D VLEs having the advantage of multichannel communication, which can effectively reduce learning barriers (e.g., anxiety and inhibition) (Wehner et al., 2011). Thus, it was felt that virtual worlds have particular potential in assisting language learners with special needs. For example, students with autism may benefit from learning languages in virtual worlds since these students might respond better in a low-stress environment where they have more time to master tasks. We believe that the continuing explorations in the VR field will lead to the development of valuable new perspectives about and methods of language development.

Finally, this study did not consider relevant conference proceedings and other language and technology-related publications. Future research should expand the data sources, as well as consider other journals and other analysis techniques (e.g., citation analysis) so as to include more of the available literature.

Acknowledgements

References

Canto, S., Jauregi, K., & Bergh, H. (2013). Integrating cross-cultural interaction through video-communication and virtual worlds in foreign language teaching programs: Is there an added value?. *ReCALL, 25*(1), 105-121.

Collentine, K. (2011). Learner autonomy in a task-based 3D world and production. *Language Learning & Technology, 15*(3), 50-67.

Cornillie, F., Clarebout, G., & Desmet, P. (2012). Between learning and playing? Exploring learners' perceptions of corrective feedback in an immersive game for English pragmatics. *ReCALL, 24*(3), 257-278.

Deutschmann, M., Panichi, L., & Molka-Danielsen, J. (2009). Designing oral participation in Second Life: A comparative study of two language proficiency courses. *ReCALL, 21*(2), 206-226.

Dooly, M., & Sadler, R. (2013). Filling in the gaps: Linking theory and practice through telecollaboration in teacher education. *ReCALL, 25*(1), 4-29.

Hansson, T. (2005). English as a second language on a virtual platform: Tradition and innovation in a new medium. *Computer Assisted Language Learning, 18(1),* 63-79.

Hauck, M., & Youngs, B. L. (2008). Telecollaboration in multimodal environments: The impact on task design and learner interaction. *Computer Assisted Language Learning, 21*(2), 87-124.

Ho, C. M. L., Rappa, N. A., & Chee, Y. S. (2009). Designing and implementing virtual enactive role-play and structured argumentation: Promises and pitfalls. *Computer Assisted Language Learning, 22*(5), 381-408.

Hsu, Y. C., Ho, H. N. J., Tsai, C. C., Hwang, G. J., Chu, H. C., Wang, C. Y., & Chen, N. S. (2012). Research trends in technology-based learning from 2000 to 2009: A Content analysis of publications in selected journals. *Educational Technology & Society, 15*(2), 354-370.

Jauregi, K., Canto, S., de Graaff, R., Koenraad, T., & Moonen, M. (2011). Verbal interaction in Second Life: Towards a pedagogic framework for task design. *Computer Assisted Language Learning, 24*(1), 77-101.

Lee, J. (1999). Effectiveness of computer-based instructional simulation: A Meta-analysis. *International Journal of Instructional Media, 26*(1), 71-85.

Liang, M. Y. (2012). Foreign ludicity in online role-playing games. *Computer Assisted Language Learning, 25*(5), 455-473.

Liao, Y. K., & Chen, Y. W. (2007). The Effect of computer simulation instruction on student learning: A Meta-analysis of studies in Taiwan. *Journal of Information Technology and Applications, 2*(2), 69-79.

Liou, H. C. (2012). The Roles of Second Life in a college computer-assisted language learning (CALL) course in Taiwan, ROC. *Computer Assisted Language Learning, 25*(4), 365-382.

O'Brien, M. G., Levy, R., & Orich, A. (2009). Virtual immersion: The Role of CAVE and PC technology. *CALICO Journal, 26*(2), 337-362.

O'Rourke, B. (2005). Form-focused interaction in online tandem learning. *CALICO Journal, 22*(3), 433-466.

Peterson, M. (2006). Learner interaction management in an avatar and chat-based virtual world. *Computer Assisted Language Learning, 19*(1), 79-103.

Peterson, M. (2009). Learner interaction in synchronous CMC: A sociocultural perspective. *Computer Assisted Language Learning, 22*(4), 303-321.

Peterson, M. (2010). Learner participation patterns and strategy use in Second Life: An Exploratory case study. *ReCALL, 22*(3), 273-292.

Peterson, M. (2011). Towards a research agenda for the use of three-dimensional virtual worlds in language learning. *CALICO Journal, 29*(1), 67-80.

Peterson, M. (2012a). EFL learner collaborative interaction in Second Life. *ReCALL, 24*(1), 20-39.

Peterson, M. (2012b). Learner interaction in a massively multiplayer online role playing game (MMORPG): A Sociocultural discourse analysis. *ReCALL, 24*(3), 361-380.

Rama, P. S., Black, R. W., van Es, E. & Warshauer, M. (2012). Affordance for second language learning in World of Warcraft. *ReCALL, 24*(3), 322-338.

Ranalli, J. (2008). Learning English with the Sims: Exploiting authentic computer simulation games for L2 learning. *Computer Assisted Language Learning, 21*(5), 441-455.

Rankin, Y., Gold, R., & Gooch, B. (2006). 3D role-playing games as language learning tools. In E. Gröller & L. Szirmay-Kalos (Eds.), *Proceedings of Eurographics, 25*(3), 211-225.

Rilling, S., Dahlman, A., Dodson, S., Boyles, C., & Pazvant, O. (2005). Connecting CALL theory and practice in preservice teacher education and beyond: Process and products. *CALICO Journal, 22*(2), 213-235.

Ryu, D. (2013). Play to learn, learn to play: Language learning through gaming culture. *ReCALL, 25*(2), 286-301.

Schwienhorst, K. (2002). The State of VR: A Meta-analysis of virtual reality tools in second language acquisition. *Computer Assisted Language Learning, 15*(3), 221-239.

Schwienhorst, K. & Borgia, A. (2006). Monitoring bilingualism: Pedagogical implications of the bilingual tandem analyser. *CALICO Journal, 23*(2), 349-362.

Shih, Y. C. (2013). A Virtual walk through London: Culture learning through a cultural immersion experience. *Computer Assisted Language Learning*, 1-22. doi:10.1080/09588221.2013.851703

Smart, J., Cascio, J., & Paffendof, J. (2007). Metaverse roadmap: Pathway to the 3D web. *Metaverse: A Cross-industry public foresight project.* Retrieved from http://www.metaverseroadmap.org/MetaverseRoadmapOverview.pdf

Smith, B., & Lafford, B. A. (2009). The Evaluation of scholarly activity in computer-assisted language learning. *Modern Language Journal, 93*(1), 868-883.

Spector, J. M., Merrill, M. D., Elen, J., & Bishop, M. J. (Eds.) (2014). *Handbook of research for educational communication and technology.* New York, NY: Springer. doi:10.1007/978-1-4614-3185-5

Sykes, J. M., Oskoz, A., & Thorne, S. L. (2008). Web 2.0, synthetic immersive environments and mobile resources for language education. *CALICO Journal, 25*(3), 528-546.

Thorne, S. L., Fischer, I., & Lu, X. (2012). The Semiotic ecology and linguistic complexity of an online game world. *ReCALL, 24*(3), 279-301.

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulation for learning: A Meta-analysis. *Journal of Educational Computing Research, 34*(3), 229-243.

Wang, S., & Vásquez, C. (2012). Web 2.0 and second language learning: What does the research tell us?. *CALICO Journal, 29*(3), 412-430.

Warburton, S. (2009). Second Life in higher education: Assessing the potential for and the barriers to deploying virtual worlds in learning and teaching. *British Journal of Educational Technology, 40*(3), 414-426.

Wigham, C. R., & Chanier, T. (2013a). A Study of verbal and nonverbal communication in Second Life: The ARCHI21 experience. *ReCALL, 25*(1), 63-84.

Wigham, C. R., & Chanier, T. (2013b). Interactions between text chat and audio modalities for L2 communication and feedback in the synthetic world Second Life. *Computer Assisted Language Learning*, 28(3), 260-283.

Wehner, A. K., Gump, A. W., & Downey, S. (2011). The Effects of Second Life on the motivation of undergraduate students learning a foreign language. *Computer Assisted Language Learning, 24*(3), 277-289.

Zheng, D., Newgarden, K., Young, M. F. (2012). Multimodal analysis of language learning in World of Warcraft play: Languaging as values-realizing. *ReCALL, 24*(3), 339-360.

Zheng, D., Young, M. F., Brewer, R. A., & Wagner, M. (2009). Attitude and self-efficacy change: English language learning in virtual worlds. *CALICO Journal, 27*(1), 205-231.