TEACHING with TECHNOLOGY 2020

Selected papers from the JALTCALL2020 Conference

JALTCALL 2020

Edo Forsythe, Editor-in-Chief

Associate Editors: Anthony Brian Gallagher
Douglas Jarrell
Branden Kirchmeyer
Louise Ohashi
James York

ISBN 978-4-901352-63-5

Teaching with Technology 2020
Selected papers from the JALTCALL2020 Conference

JALT CALL is a Special Interest Group (SIG) of the Japan Association for Language Teaching (JALT) that focuses on Computer-Assisted Language Learning (CALL) and technology in language learning.

http://jaltcall.org/

JALT is dedicated to the improvement of language teaching and learning.

JALT Central Office
Urban Edge Bldg. 5F, 1-37-9 Taito,
Taito-ku, Tokyo 110-0016, Japan

http://jalt.org/

First published in Japan in 2021 by the JALT CALL SIG.

Layout by Paul Mason

https://doi.org/10.37546/JALTSIG.CALL2020.8

8

A Comparison of the Affective Affordances of a Static and Interactive VR System on Learner FLA and Motivation

Koichi Shibata and James York, Tokyo Denki University

Abstract

This paper introduces a virtual reality (VR) system which was designed to promote English speaking proficiency as learners carry out collaborative information gap tasks. In a former study, a simpler system was developed to explore the effect of modality on learners' foreign language anxiety (FLA) where results suggested that anxiety was statistically significantly lower in the VR environment compared to a voice and video chat system. However, of three key affordances – presence, interactivity, and autonomy – the previous system only focused on presence. The current system features an interactive component also. In this paper, we present results of a study which compared the two systems (presence-only versus interactive system) with the aim of answering the question: Does more-fully utilizing the affordances of VR lower or increase learners' FLA?

In a counterbalanced design, 30 participants (15 pairs) completed a spot-the-difference task in two different VR environments: static-VR (former system) and interactive-VR (current system). Results of a post-experimental questionnaire suggested that there was no difference in participants' FLA for the two domains. However, a significant difference was found in terms of ease of communication and enjoyment which favored the interactive-VR mode. Additionally, compared to predictions that the interactive-VR task would be more cognitively demanding, it was considered simpler than the static-VR task. This suggests that using more of the affordances of VR by increasing interactivity further may make the embodied experience more life-like and therefore increase opportunities for learning. This paper introduces the system, implications for researchers and teachers, and future research directions.

本論文では、共同でinformation gap taskを実行する際に、英会話能力を促進するように設計された仮想現実（VR）システムを紹介します。以前の研究では、コミュニケーション方法が学習者の外国語不安感（FLA）に及ぼす影響を調査するために、シンプルなシステムが開発され、学習者の不安は、音声およびビデオチャットシステムと比較して、VR環境では統計的に有意に低いことが示唆されました。ただし、3つの主要な要素（没入感、対話性、および自律性）のうち、以前のシステムは没入感のみに焦点を合わせ

ていました。現在のシステムは、インタラクティブな機能を追加しました。この論文では、VRの要素をより十分に活用することで学生のFLAが低下または増加するのかという調査を行いました。

Keywords : **SCMC, VR, interactivity, motivation, anxiety**

Virtual reality (VR) exists along a continuum of augmented, mixed, and virtual realities, which refers to the level of immersion into a virtual environment each system provides (Hawkinson et al., 2017). Although originally conceptualized in the 1960s (see Sherman & Craig, 2018), VR did not become a popular consumer product until the last decade with the gradual and incremental improvements made in technology (both hardware and software). Stein (2019) provides a succinct history of developments over the last 10 years which include:

- Head-mounted display systems which require a desktop computer (HTC Vive, Oculus Rift)
- Augmented reality glasses (Google Glass)
- Augmented reality games which utilize smartphones GPS and gyroscope functions (Pokémon Go being one of the most popular games)
- Smartphone VR systems (Google Cardboard)
- Stand-alone VR systems (Oculus Quest)

In sum then, the development of VR technologies is growing, and becoming more affordable for consumers who are being provided with systems that are both cheaper and easier to use. Along with this trend of VR entering the consumer market then, research within CALL is also beginning to utilize the affordances of this new technology for language learning purposes.

The research outlined in this paper has the aim of developing a system which can be used to connect Japanese learners with native English speakers in a virtual environment to complete focused, closed-goal tasks (Ellis, 2003). The system has a specific focus on oral communication as this skill is most neglected or difficult to practice in classroom environments void of native speakers. The current paper introduces results of a study which sought to understand how the cognitive demands of completing a collaborative information gap task in a VR system could affect learners' foreign language anxiety (FLA) and motivation towards studying a foreign language. As this is a preliminary test of the system, native speakers were not a part of the study.

The reason that we focus on FLA is that it is known to negatively affect learner motivation to study a second or foreign language (Hewitt & Stephenson, 2012). It is considered one of the most hindering barriers to successful L2 acquisition, affecting a learner's motivation, and self-efficacy in the L2 (Dewaele & Thirtle 2009; Horwitz et al., 1986). The reason we focus on an oral collaborative task is because the context of this study is Japan, where the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) promote the development of practical English skills including being able to express oneself in English (MEXT, 2014, p.3). However, as Japan is a homogeneous nation (Ebuchi & Yokota, 2019) there are few opportunities for residents to interact with native speakers in face-to-face communication. Certainly, the Internet and other forms of computer-mediated communication (CMC) provide opportunities for learners to communicate with native speakers, but VR holds the most promise in terms of allowing for an embodied experience which has been called a "hybrid mode of communication which may offer the benefits of both face-to-face and computer-mediated modalities" (York et al., 2020, p.2).

Literature Review

Computer-mediated Communication in Language Learning Contexts

Before outlining the research on VR in language learning contexts, this section introduces the wider topic of computer-mediated communication (CMC) studies within CALL. There is a growing body of literature exploring the cognitive and affective affordances of CMC, which is further divided into asynchronous (ACMC) and synchronous (SCMC) modes. Here, the focus is on SCMC.

In relation to the present study, there are several meta-analyses on the effect of SCMC in language learning contexts (Lin, 2014; Aslan & Ciftci, 2019). Ziegler's (2016) meta-analysis examined 14 studies which utilized SCMC as a mode for interaction between learners. Findings of the meta-analysis suggested that both face-to-face (FTF) and SCMC interactions produced significant, positive effects on learners' language development. This confirms the notion that interaction – and in turn the Output Hypothesis – is beneficial to language development, but also that SCMC is not inferior to FTF communication. Furthermore, there was a small, positive advantage for SCMC over FTF for written production. However, this may be a moot point when considering the types of SCMC modes analyzed in this meta-analysis. SCMC does not imply oral interaction, and indeed, only three of the 14 included studies incorporated oral SCMC. Thus, if the SCMC mode is written, one would expect there to be an improvement in written output in comparison to FTF communication, which by necessity is oral.

Regarding the affective affordances of SCMC, text-based SCMC has been shown to reduce FLA in comparison to FTF communication (Abrams, 2003; Tudini, 2007). Furthermore, studies have shown that SCMC can promote more equal turn-taking opportunities (Warschauer, 1996) and improved willingness to communicate among peers (Reinders & Wattana, 2014; Yanguas & Flores, 2014) as well as with native speakers (Jauregi et al., 2012; Iino & Yabuta, 2015). Melchor-Couto (2017) conducted a study comparing the FLA of an experimental group who completed oral tasks in a virtual world to a control group that conducted the same tasks face-to-face. Results suggested that the FLA of participants in the experimental group went down over multiple meetings within the virtual world. This was attributed in part to the anonymity afforded by avatar-based communication (see also Dickey, 2005).

The Use of VR in Language Learning

As mentioned in the introduction to this paper, VR is currently a hot topic in CALL and broader educational fields (Liu et al., 2017). Several papers have explored hypothetical applications of VR and related technologies in language learning contexts (Bonner & Reinders, 2018; Hawkinson, et al., 2017; Alizadeh, 2019). However, there are few empirical studies on the cognitive or affective affordances of the technology. Of those that exist, there is an emphasis on exploring VR's effect on vocabulary acquisition (Cheng et al., 2017; Legault et al., 2019) where results claim successful vocabulary learning was due to the immersive nature of the environment, and provision of culturally relevant interactions.

York et al. (2020) compared the effect of three modes of SCMC on learners' FLA: oral SCMC, video SCMC and VR. All three modes were effective in reducing learners' FLA, but compared to oral and video SCMC modes, participants considered VR the most natural environment to communicate in, as well as being the most fun, and the most effective environment for language learning. However, limitations of the study were that the affordances of the VR domain were specifically *not* utilized to unify the complexity of the three chosen modes. The current study may be considered an extension, where an interactive element is added to the previous system. However, the efficacy of this new system in reducing learners FLA is unknown, hence the impetus for this study.

Task Complexity and Cognitive Demand

Robinson's (2011) Triadic Componential Framework for task conditions is a robust framework for assessing and making predictions regarding the *cognitive complexity* of language learning tasks, also known as the *Cognition Hypothesis*. It was developed over a number of iterations (see Robinson, 2000; 2005) and validated in several studies (Jackson & Suethanapornkul, 2013; Sasayama, 2016; York, 2019). The framework attempts to provide practitioners with the ability to predict the complexity of a language task and is separated into three major categories: task complexity, conditions and difficulty. For brevity, this paper focuses on the *task complexity* dimension. As can be seen in Table 1, task complexity is bifurcated into resource-directing and resource-dispersing dimensions. The addition of + and – signs also indicate that task complexity can be instrumentalized by a task designer by increasing or decreasing a certain condition. As a concrete example, increasing the number of elements that a learner has to manipulate in a task (*+few elements*) would increase its complexity along the resource-directing dimension. Thus, the complexity of a task is determined by the cognitive demands placed on learners.

Table 1
Robinson's Triadic Componential Framework (2011)

Task complexity		Task conditions		Task difficulty	
Resource-directing	**Resource-dispersing**	**Participation**	**Participant**	**Affective variables**	**Ability variables**
+/– Here-and-Now	+/– Planning	+/– Open solution	+/– Same proficiency	H/L Openness	H/L Working memory
+/– Few elements	+/– Prior knowledge	+/– One-way flow	+/– Same gender	H/L Control of emotion	H/L Reasoning
+/– Spatial reasoning	+/– single task	+/– Convergent solution	+/– Familiarity	H/L Task motivation	H/L Task switching
+/– Causal reasoning	+/– Task structure	+/– Few participants	+/– Shared content knowledge	H/L Processing anxiety	H/L Aptitude
+/– Intentional reasoning	+/ Few steps	+/– Few contributions needed	+/– Equal status and role	H/L Willingness to communicate	H/L Field Independence
+/– Perspective-taking	+/– Independency of steps	+/– Negotiation not needed	+/– Shared cultural knowledge	H/L Self-efficacy	H/L Mind-reading

Robinson argues that cognitive complexity may have a positive effect on learners' output accuracy and complexity, but a negative effect on fluency as learners focus on producing complex language. Subsequently, due to the extra focus required, acquisition of the target language may also be improved. In relation to the current study, affectively, the increase in cognitive task complexity may have a negative effect on learners' motivation as they become overloaded by the cognitive demands of a task (Sweller, 2005; Paas et al., 2005; van Gog & Paas, 2008).

The topic of task complexity and cognitive demands has relevance to the present study in that we are testing two VR systems which, based on Robinson's Triadic Componential Framework, may be considered more or less complex environments. Thus, in this paper, we are addressing the question of how incorporating interactive elements into a VR language learning system may affect learner performance due to differing task complexities.

Methodology

The aim of this study is to uncover the benefit or hindrance of an interactive VR system on learners' FLA and motivation towards learning.

Participants

The study was carried out at a private university in Japan. 30 participants (15 pairs) volunteered to take part in the study. Their mean age was 20.9 (SD = 1.54). All participants were native speakers of Japanese. 27 of the participants were male and three were female which is not uncommon for this university (Table 2). In order to avoid a familiarity effect, participants were paired with partners that they meet for the first time when undertaking the VR tasks.

Table 2

Participant Details

Baseline Characteristic		n	%
Age			
	18	1	3.3
	19	4	13.3
	20	8	26.7
	21	7	23.3
	22	3	10.0
	23	6	20.0
	24	1	3.3
Gender			
	Male	27	90
	Female	3	10

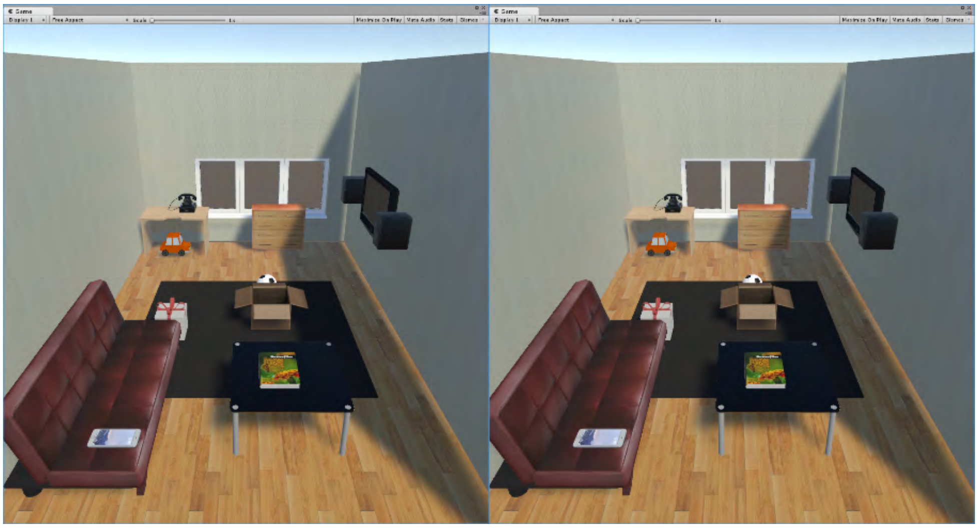
Instruments

VR system

Two VR systems were employed in this study. Both environments were designed as spot-the-difference information gap tasks to be completed by pairs of learners. Both systems present participants with their own doll house which features a number of items typically found in a living room. To create a spot-the-difference activity, three items appear in different places. Figure 1 shows this as the football, the car, and the magazine.

Figure 1

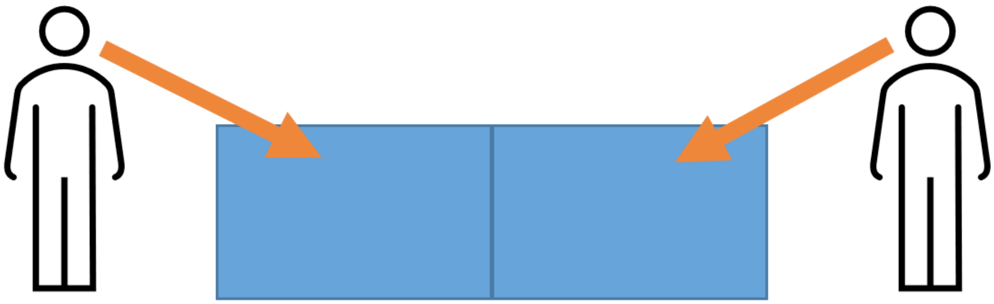
Example of the Doll Houses Presented to Each Participant



Participants are able to see each other standing behind their respective doll house, but not the content, as the doll houses were placed back to back (Figure 2).

Figure 2

Graphical Representation of the VR System Featuring Both Participants



Using the HTC Vive head-mounted display allowed for an immersive VR experience for both systems. Participants were able to look around the environment by moving their head and gesture by moving the controllers in their hands (Figure 3).

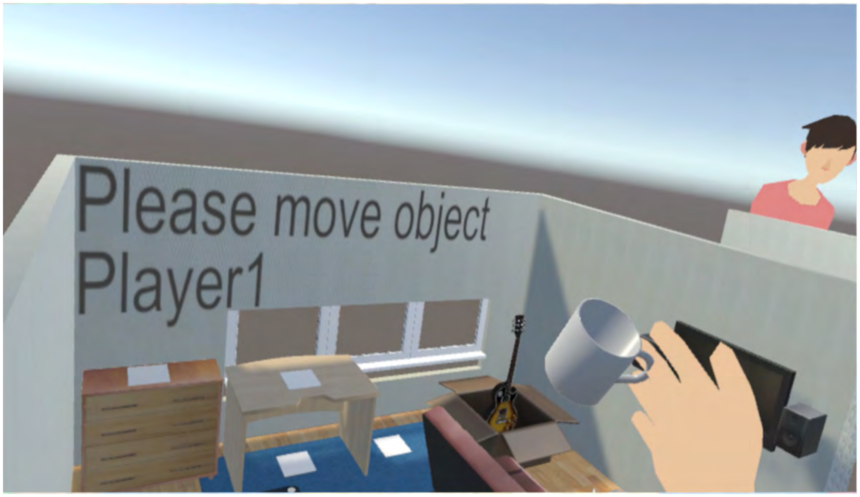
Figure 3

A Participant Gesturing with the Controller

The difference between the two systems is that in the previous *static-VR* system participants can only interact with the doll house by looking at it. They cannot move objects in the room. In this way, they must decide when they have found the three differences between their two rooms and end the task. With the current, *interactive-VR* system, participants can pick up and move objects in their room (Figure 4). Additionally, once participants have moved objects into the same position in both doll houses, a message is shown to tell them that they have completed the task.

Figure4

A Screenshot of a Participant Moving an Object in the Interactive-VR system



In summary then, the interactive-VR system was developed as an extension to the static-VR system. A comparison is shown in Table 3.

Table 3

A Comparison of the Static and Interactive VR Systems

Static-VR	Interactive-VR
Doll house	Doll house
Gestures via controllers	Gestures via controllers
	Moveable objects
	Notification when the task is complete

Pre-task

A pre-task was employed to prime learners for the main spot-the-difference task in the VR environments (Appendix A). It featured questions about living room vocabulary, prepositions, and a sample spot-the-difference task. The worksheet was created with images taken from the VR environment. Dyads were allocated 15 minutes to complete the pre-task.

Main Task

The main task was carried out using the VR systems described above. Participants were placed in separate rooms and given instructions on how to use the head-mounted display and controllers. Discord was utilized as the CMC tool, allowing participants to communicate orally as they carried out the task.

Although the two tasks are comparable, the interactive-VR system adds both interactivity (virtual tangibility) and feedback on correct task completion. In terms of Robinson's (2011) Triadic Componential Framework then, the interactive-VR system is considered to be of higher task complexity due to manipulations along the following dimensions:

- **Spatial reasoning:** participants must understand the virtual space to move the objects into the correct places.
- **Prior knowledge:** if participants are not familiar with VR, controllers, and headset, an additional cognitive demand is placed upon them by requiring them to pick up and move objects.
- **Few steps:** There are more steps to complete (moving items into place).

However, as the VR environment is fully immersive and an embodied experience, participants may find the interactive-VR activity intuitive and natural, as found in York, (2020) where a VR SCMC environment was considered an easier mode of communication than both oral and video SCMC.

Reduced FLCAS

A reduced version of Horwitz et al.'s (1986) foreign language classroom anxiety scale (FLCAS) was used in this study to measure participants' FLA at the pre- and post-experiment stages. The reduced version was based on Melchor-Couto (2017; 2018) which features only seven questions:

1. I never feel quite sure of myself when I am speaking in my foreign language class.
2. I do not worry about making mistakes in language class.
3. I start to panic when I must speak without preparation in language class.
4. In language class, I can get so nervous I forget things I know.
5. Even if I am well prepared for language class, I feel anxious about it
6. I feel confident when I speak in foreign language class.
7. I am afraid that the other students will laugh at me when I speak the foreign language.

Post-experiment Questionnaire

Upon task completion, the participants completed the FLCAS and were also asked to give their perceptions of learning within each environment (Table 34). Five measures were utilized based on Satar and Ozdener (2008) and were weighted from 1 to 10, 10 being a strong indication of agreement and 1 disagreement with each statement. In addition, five measures were also included to understand how each environment affected participants' cognitive load, such as in deHaan et al. (2010). Responses to these statements were then used to ascertain the accuracy of task complexity predictions. For example, Statement 1 is an indicator of the cognitive demands of a system where participants' cognitive capacity may be overloaded by the system meaning that they have no capacity left to attend to learning goals (Lim et al., 2006).

Table 4
Measures Used in the Post-Experiment Questionnaire

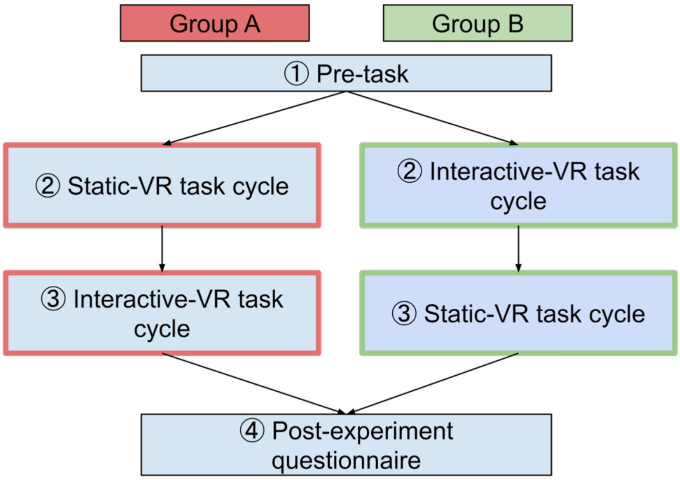
Number	Statement
1	It was easy to speak English when using this system.
2	It was fun to learn with this system.
3	It was easy to complete the task with this system.
4	I think this system is an effective way of learning English.
5	I did not feel anxious learning English with this system.

Procedure

The experiment was a counterbalanced, repeated-measures design where participants completed a spot-the-difference task using both systems. The procedure can be seen graphically in Figure 5. The 15 pairs were assigned to either Group A or Group B. First they both completed the pre-task (1). As part of the pre-task, participants also answered the FLCAS questions. In order to avoid a familiarity effect the two groups completed the tasks in a different order ((2) and (3)). The FLCAS was administered post-task. Finally, both groups completed the post-experiment questionnaire (4).

Figure 5

The Experiment Procedure Flowchart



Data Analysis

A one-way repeated measures ANOVA was used to compare mean scores for the pre-experiment, and post-task FLA scores. An alpha level of p = .05 was set for all statistical tests. The statistical analysis software used was IBM's SPSS 24. A paired samples *t*-test was used to determine if participants' perceptions of the two systems were affected by interactivity.

Research Questions

Based on the study design above, the following questions are explored in this paper:

1. Which system is more effective at reducing learners' FLA?
2. How does interactivity affect learners' perception of task difficulty and enjoyment?

Results

RQ1: FLA Scores

FLA scores measured at the pre-experiment and post-task levels producing three sets of data. Descriptive statistics are available in Table 5. Both the static-VR and interactive-VR systems reduced participants FLA in comparison to their pre-experiment FLA scores.

Table 5

Descriptive Statistics for FLA Scores

FLA score	Mean	Std. Deviation
Pre-experiment	20.04	3.75
Post-task Static-VR	17.13	4.36
Post-task Interactive-VR	16.96	3.90

Following the descriptive statistics, a repeated measures ANOVA was run on the data. Mauchly's test of sphericity indicated that the assumption of sphericity was met and that there was an interaction between FLA scores and when the questionnaire was completed $\chi^2(2) = 4.14$, $p = .11$. Pairwise comparisons revealed that there was a significant difference in FLA mean scores between the pre-experiment and both post-task questionnaires. However, there was no statistically significant difference between mean scores for the two VR modes (see Table 6). This indicates that both systems were effective in lowering participants FLA, but that neither system was more effective than the other.

Table 6

Pairwise Comparisons of Mean FLA Score for All Three Stages of the Experiment

(I) Questionnaire	(J) Questionnaire	Mean Difference (I-J)	Std. Error	Sig.
Pre-experiment	Static-VR	2.917*	0.645	.00
	Interactive-VR	3.083*	0.507	.00
Static-VR	Pre-experiment	–2.917*	0.645	.00
	Interactive-VR	0.167	0.457	1
Interactive-VR	Pre-experiment	–3.083*	0.507	.00
	Static-VR	-0.167	0.457	1

RQ2: Learner Perceptions of the Two Systems

A post-experiment questionnaire explored participant perceptions of learning within each system, with a focus on cognitive demands. Descriptive statistics are provided in Table 7. Of note are two things: 1) all mean scores are higher for the interactive-VR system and 2) all standard deviations are smaller for the interactive-VR. This indicates that converse to predictions, the additional cognitive demands posed by the interactive-VR system were not perceived as such by the participants. A detailed explanation follows.

Table 7

Descriptive Statistics and *t*-test Results for the Post-Experiment Questionnaire Measures

Measure	Mean Scores (*SD*)			
	Static-VR	Interactive-VR	Mean diff.	Sig.
1. It was easy to speak English when using this system.	7.63 (1.59)	8.57 (1.40)	0.93	.002
2. It was fun to learn with this system.	8.77 (1.68)	9.53 (0.9)	0.77	.005
3. It was easy to complete the task with this system.	7.30 (2.22)	8.07 (1.86)	0.77	.032
4. I think this system is an effective way of learning English.	8.40 (1.42)	8.80 (1.19)	0.40	.07
5. I did not feel anxious learning English with this system.	7.57 (2.42)	7.80 (2.21)	0.23	.36

Paired-samples *t*-tests were conducted to compare participants' perceptions of the VR systems for each of the measures on the post-experiment questionnaire. There was a significant difference in the mean scores for three of the five measures. Focusing on measures 1 and 3, results revealed that participants perceived it significantly easier to speak English in the interactive-VR mode ($p < .001$). The interactive-VR task was also considered easier to complete than the static-VR task ($p < .05$). This suggests that the increase in *spatial reasoning* and *steps required* for task completion did not translate into an increase in cognitive task complexity as perceived by participants.

There was also a statistically significant difference in mean scores for measure 2 ($p < .05$) which suggests, again, that increasing the interactive affordances of VR was not perceived as a hindrance but a positive, enjoyable addition. Finally, the differences in mean scores between the systems for measure 4 approached significance ($p = .07$), which suggests that the interactive-VR system was considered a more effective learning environment than the static-VR system. Adding weight to the results of the FLCAS mean scores analysis, there is no significant difference in mean scores for measure 5 here either.

Conclusion

The impetus for the current study was a previous study of ours in which we compared three modes of SCMC on learners' FLA (York et al., 2020). In that study, differences in the three SCMC environments were kept as minimal as possible. This meant that the VR system in the previous study did not utilize the interactive affordances of VR. However, results suggested that compared to oral and video SCMC, the VR environment was the most effective in reducing learners' FLA (York et al., 2020). In the current study, interactive elements were added which were a cause for concern in terms of task complexity. Adding interactive elements to the system was predicted to increase task complexity based on Robinson's Triadic Componential Framework (2011). However, conversely, the results suggest that additional interactivity created a favorable context for embodied learning which was both enjoyable and easier to communicate than the previous system.

Implications for Designers

For designers, the results of this study suggest that the incorporation of VR's interactive affordances may be beneficial rather than a hindrance for language learning. The provision of feedback (both in terms of interaction with the system, and as indicators of progress) may reduce cognitive demand and increase motivation to learn.

In FTF communication, recasts are considered one of the most frequent types of feedback given by teachers and effective in promoting language development (Long, 2014, p. 27). Within a VR environment, feedback can be given from interlocutors as in a classroom, but also from the environment itself, and based on results of this study, should be included where possible to aid in comprehension.

Implications for Teachers

Assuming the trend in affordable, portable, and stand-alone VR systems continues, it will not be long before devices like the HTC Vive will be implementable in classroom teaching. With such, teachers will be able to offer their students access to immersive environments that can be utilized for language learning purposes: access to native speakers, interactive tasks that take advantage of kinesthetic learning, and embodied co-constructed interaction between peers (see Steffensen, 2015). Results of the current study suggest that not only will learners enjoy using such technology, but that if tasks utilize the affordances of the environment, will be considered less cognitively demanding than tasks done in non-interactive systems.

Future Research

The positive results of this study act as a catalyst for investigating the affective and cognitive benefits of additional VR affordances. From an instructional design perspective, the addition of a time limit and multiple levels of tasks to clear – making the system more like a typical video game – would allow for an individualized experience which matches a learners' developmental stage. This would also keep learners in a flow state (Csikszentmihalyi, 1990). As a concrete example, if a pair of learners could not complete a level within a specific time limit, an easier level would be presented instead. Alternatively, if a pair completed a level in a very short period of time, they could be presented with a much more complex task upon completion (i.e., increasing the number of elements in the spot-the-difference task used here). Finally, as the system is already network-enabled, it is possible to enter the environment from anywhere in the world, and thus, one step that we are keen to explore in future research is how FLA is affected when participants' interlocutors are 1) from another country, or 2) native speakers.

References

Abrams, Z.I. (2003). The effect of synchronous and asynchronous CMC on oral performance in German. *The Modern Language Journal*, *87*(2), 157–167.

Alizadeh, M. (2019). Augmented/virtual reality promises for ELT practitioners. *JALT2018 – Diversity and Inclusion*, *2018*(1), 360. **https://doi.org/10.37546/jaltpcp2018-48**

Aslan, E., & Ciftci, H. (2019). Synthesizing research on learner perceptions of CMC Use in EFL/ESL writing. *CALICO Journal*, *36*(2), 100–118. **https://doi.org/10.1558/cj.34818**

Bonner, E., & Reinders, H. (2018). Augmented and virtual reality in the language classroom: Practical ideas. *Teaching English with Technology, 18*(3), 33–53.

Cheng, A., Yang, L., & Andersen, E. (2017). Teaching language and culture with a virtual reality game. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. CHI'17: CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery. **https://doi.org/10.1145/3025453.3025857**

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row.

deHaan, J., Reed, W.M., & Kuwada, K. (2010). The effect of interactivity with a music video game on second language vocabulary recall. *Language Learning & Technology*, *14*(2), 74–94.

Dewaele, J.M., & Thirtle, H. (2009). Why do some young learners drop foreign languages? A focus on learner-internal variables. *International Journal of Bilingual Education and Bilingualism, 12*(6), 635–649. **https://doi.org/10.1080/13670050802549656**

Dickey, M.D. (2005). Three-dimensional virtual worlds and distance learning: Two case studies of active worlds as a medium for distance education. *British Journal of Educational Technology*, *36*(3): 439–451.

Ebuchi, T., & Yokota, Y. (2019). Japan immigration hits record high as foreign talent fills gaps. *Nikkei Asian Review*. **https://asia.nikkei.com/Spotlight/Japan-immigration/Japan-immigration-hits-record-high-as-foreign-talent-fills-gaps**

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

Hawkinson, E., Mehran, P., & Alizadeh, M. (2017). Using MAVR to bring new dimensions to the classroom. *The Language Teacher*, *41*(3), 30–32.

Hewitt, E., & Stephenson, J. (2012). Foreign language anxiety and oral exam performance: A replication of Phillip's MLJ Study. *The Modern Language Journal*, *96*(2): 170–189. **https://doi.org/10.1111/j.1540-4781.2011.01174.x**

Horwitz, E.K., Horwitz, M.B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125–132.

Iino, A., & Yabuta, Y. (2015). The effects of video SCMC on English proficiency, speaking performance and willingness to communicate. In F. Helm, L. Bradley, M. Guarda, & S. Thouësny (Eds.), *Critical CALL – Proceedings of the 2015 EUROCALL conference, Padova, Italy* (pp. 254–260). Research-publishing.net. **http://dx.doi.org/10.14705/rpnet.2015.000342**

Jackson, D.O., & Suethanapornkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning*, *63*(2), 330–367. **https://doi.org/10.1111/lang.12008**

Jauregi, K., de Graaff, R., van den Bergh, H., & Kriz, M. (2012). Native/non-native speaker interactions through video-web communication: A clue for enhancing motivation? *Computer Assisted Language Learning*, *25*(1), 1–19. **https://doi.org/10.1080/09588221.2011.582587**

Legault, J., Zhao, J., Chi, Y.-A., Chen, W., Klippel, A., & Li, P. (2019). Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages, 4(1)*, 13. **http://dx.doi.org/10.3390/languages4010013**

Lim, C.P., Nonis, D., & Hedberg, J. (2006). Gaming in a 3D multiuser virtual environment: Engaging students in science lessons. *British Journal of Educational Technology*, *37*(2), 211–231.

Lin, H. (2014). Establishing an empirical link between computer-mediated communication (CMC) and SLA: A meta-analysis of the research. *Language Learning and Technology*, *18*(3), 120–147.

Liu, D., Dede, C., Huang, R., & Richards, J. (2017). *Virtual, augmented, and mixed realities in education*. Springer.

Long, M. (2014). Second language acquisition and task-based language teaching. John Wiley & Sons.

Melchor-Couto, S. (2017). Foreign language anxiety levels in Second Life oral interaction. *ReCALL*, *29*(1), 99–119.

Melchor-Couto, S. (2018). Virtual world anonymity and foreign language oral interaction. *ReCALL*, *30*(2), 232.

Ministry of Education, Culture, Sports, Science, and Technology. (2014). *On Integrated Reforms in High School and University Education and University Entrance Examination Aimed at Realizing a High School and University Articulation System Appropriate for a New Era*. **https://www.mext.go.jp/component/english/__icsFiles/afieldfile/2015/03/31/1353908_1.pdf**

Paas, F., Tuovinen, J.E., Van Merriënboer, J.J.G., & Darabi, A.A. (2005). A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction. *Educational Technology Research and Development*, *53*(3), 25–34. **https://doi.org/10.1007/BF02504795**

Reinders, H., & Wattana, S. (2014). Can I Say Something? The effects of digital game play on willingness to communicate. *Language Learning & Technology*, *18*(2), 101–123. **https://doi.org/10125/44372**

Robinson, P. (2000). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, *22*(1), 27–55.
Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching*, *43*(1), 45–107.
Robinson, P. (2011). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance,* (pp. 3–38). John Benjamins Publishing.
Sasayama, S. (2016). Is a 'complex' task really complex? Validating the assumption of cognitive task complexity. *Modern Language Journal*, *100*, 231–254.
Satar, H.M. & Ozdener, N. (2008). The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat. *The Modern Language Journal*, *92*(4), 595–613.
Sherman, W.R., & Craig, A.B. (2018). *Understanding virtual reality: Interface, application, and design*. Morgan Kaufmann.
Stein, S. (2019, December 17). *VR in the 2010s: My decade with things on my face*. Cnet.com. **https://www.cnet.com/news/arvr-2009-2019-my-decade-with-things-on-my-face/**
Steffensen, S.V. (2015). Distributed language and dialogism: Notes on non-locality, sense-making and interactivity, *Language Sciences*, *50*, 105–119.
Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R.E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). Cambridge University Press.
Tudini, V. (2007), Negotiation and intercultural learning in Italian native speaker chat rooms. *The Modern Language Journal*, *91*, 577–601. **https://doi.org/10.1111/j.1540-4781.2007.00624.x**
van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist*, *43*, 16–26.
Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal, 13*(2), 7–26.
Yanguas, Í., & Flores, A. (2014). Learners' willingness to communicate in face-to-face versus oral computer-mediated communication. *The JALT CALL Journal*, *10*(2), 83–103.
York, J., Shibata, K., Tokutake, H., & Nakayama, H. (2020). Effect of SCMC on foreign language anxiety and learning experience: A comparison of voice, video, and VR-based oral interaction. *ReCALL*, *33*(1), 1–22. **https://doi.org/10.1017/S0958344020000154**
Ziegler, N. (2016). Synchronous computer-mediated communication and interaction. *Studies in Second Language Acquisition*, *38*(3), 553–586. **https://doi.org/10.1017/S027226311500025X**

Appendix A

Spot the difference task: Warm up

In this lesson, you will work with a partner. You will compare two pictures of the same room. There are six items placed in the room, but some of them are in a different place to your partner. This is a spot-the-difference activity. You will have to decide which of the items are the same and which are different.

Activity 1: What is in your living room?

Think of the objects that you find in a living room and make a list of 10 things here:

____________________ ____________________

____________________ ____________________

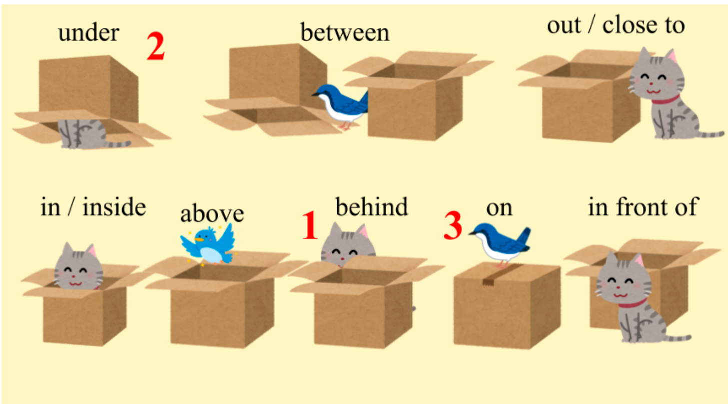
____________________ ____________________

____________________ ____________________

____________________ ____________________

Activity 2: Positions of place

Look at the following picture. The cat 🐱 and bird 🐦 are in different positions in relation to the box.

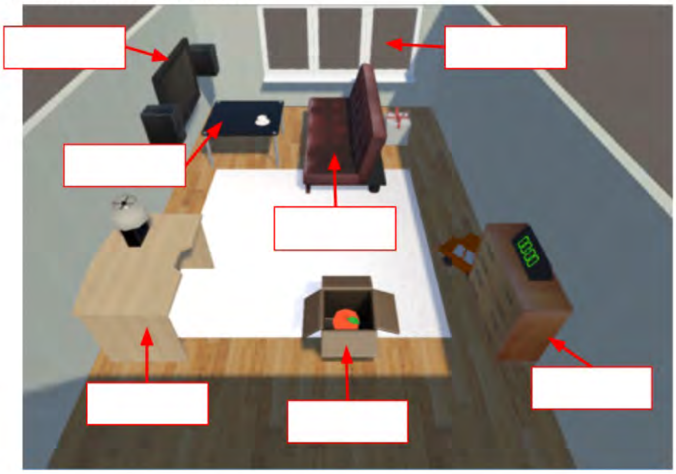


Please write sentences for the three pictures labelled 1, 2 and 3. For example: *The cat is in the box.*

1. ____
2. ____
3. ____

Activity 3: Label the items.

Please label the items marked with a textbox:



Activity 4: Where are the objects?

In this room, there are the following objects:

 Apple	 Lamp	 Clock
 Car	 Present	 Cup

Please mark the following sentences correct or incorrect.
Please use a ✓ for correct sentences.
Please use a ✗ for incorrect sentences.
For example:

The apple is on the box. ____X_____

1. The clock is on the chest of drawers. __________
2. The car is in front of the chest of drawers. __________
3. The present is on the sofa. __________
4. The lamp is on the desk. __________
5. The cup is in front of the tv. __________

Authors' Bios

Koichi Shibata is a master's student studying information sciences at Tokyo Denki University. He conducts research on the application of VR in language education.

James York is a lecturer at Tokyo Denki University where he researches the application of games and play in language teaching contexts. He also co-edits *Ludic Language Pedagogy*, an open-access journal.