Vocabulary Acquisition in New and Learned Contexts Using Immersive Virtual Reality

Thérèse Bergsma, Robby van Delden, Mariët Theune
Human Media Interaction - University of Twente, Enschede, the Netherlands
t.s.l.bergsma@gmail.com
r.w.vandelden@utwente.nl
m.theune@utwente.nl

Abstract: Vocabulary acquisition is important for learning a language. Language learners need multiple encounters with a word to add unknown word aspects to their word knowledge or to strengthen their knowledge of previously learned word aspects. Combining vocabulary acquisition with immersive virtual reality (IVR) makes it possible to have such multiple encounters. In IVR, objects representing target words can be shown in the same virtual environment as they were first encountered (*learned context*), but also in a new virtual environment (*new context*). To investigate learning words in either a new or learned context in IVR we built Wics (words in context system). We used Wics to carry out an experiment in which participants could learn 32 Japanese words, represented by 3D objects, in three sessions. In the second and third learning session all words were shown in either a new context, with a changed environment and a changed visual representation of the word objects, or a learned context, where only the placement of the objects was changed. Performance scores for an immediate and a delayed post-test (one week later) revealed no significant differences in performance between conditions. However, there was a tendency towards participants spending more time on learning the words in the learned context condition. We believe Wics to be the first IVR system to provide learners with multiple learning sessions to provide multiple encounters with the same words in different contexts, where learners have control over their own explorative learning, and where the post-tests are inside IVR.

Keywords: vocabulary acquisition, immersive VR, learned context, new context, dynamic representation, language learning

1. Introduction

Vocabulary acquisition research is complex and has a multi-dimensional nature. The research varies from theories on how we learn words to multimodal conversational analysis in the wild, where new technologies such as eye tracking facilitate other ways of analysis and gaming provides new forms of interaction (Loewen & Sato, 2017; Godwin-Jones, 2018). What can be learned about a word is more than just the meaning that is connected to it. A word has its own pronunciation and a specific text representation, but it also has other aspects such as in what context you can expect to encounter a word. If more word aspects are understood, then the depth of word knowledge increases (Nation, 2013). Multiple encounters with a word are necessary to add to a learner's word knowledge; this is called word recycling. Words can be encountered in a new context or a learned context (Jeong *et al.*, 2010). In the first case, a learner has not seen the word in the same context before and in the second case, the learner has previously encountered and processed the word in the same context. Here, a context is everything that surrounds a word when it is encountered. Jeong *et al.* (2010) compared videos showing words (text-based) in always the same context (learned context), with videos that visualized words in action (situation-based) in varying contexts (new context). The text-based participants performed exceedingly well when doing a text-based test, where they saw their learned context, but performed badly when they had to recall the word in a new context in a situation-based test. The situation-based learners did relatively well in both learned and new contexts.

An emerging focus within language learning is the combination of language learning with immersive virtual reality (IVR). IVR is a technology that allows many contexts to be easily generated. It provides a person with a full 360-degree experience by wearing either a head-mounted display or with surrounding projections. IVR has characteristics that create new opportunities for second language (L2) learners, including creating a feeling of interacting with an object inside the virtual environment (VE) or conversing with a non-player character (NPC) to counter learner anxiety (Collins, Vaughan and Cullen, 2020). IVR is also able to quickly bring variations into a VE or to change the whole VE. However, to our knowledge all previous research on vocabulary acquisition in IVR has worked with only one VE for each participant in each condition.

Building on the earlier mentioned necessity for multiple word encounters in different contexts for adding to the depth of word knowledge, our research question is: *What are the effects of recycling words in IVR in* *learned or new contexts on retrieving words when encountered in a new context?* We define *the context* as the virtual environment including specific visual representations of words as objects. Retrieving words is the ability to link these objects to the (phonetically) proper L2 word. To recycle words in IVR in new and learned contexts, and to evaluate this in a new context, we built a system called Wics, short for *words in contexts system*. We used it in an experiment to answer our research question. Below, we first describe related work to place our study and contributions in a wider context, then introduce our system and study, describe the results of this study and end with a discussion and conclusion.

2. Related work

For building Wics and setting up our study we looked at related work that, like ours, (i) presents participants with predetermined target words inside a VE in IVR to study and learn, and (ii) tests participants on words that they have encountered in the VE. This resulted in 16 related environments, not all discussed in this paper. Most closely related to our own work are projects that take an *exploration-based approach* to vocabulary acquisition. In these studies, participants can freely explore a VE in which target words are placed as visual objects. VEs typically correspond to familiar environments such as a living room (Ebert, Gupta and Makedon, 2016; Garcia *et al.*, 2019), kitchen (Legault *et al.*, 2019; Jia and Liu, 2019), zoo (Legault *et al.*, 2019), airport customs (Dobrova *et al.*, 2018), classroom (Cho, 2018) or supermarket (Collins, Vaughan and Cullen, 2020).

Various strategies are employed to make participants aware of target words within the environment. For example, exclamation marks hover over the target words (Ebert, Gupta and Makedon, 2016), objects are highlighted in blue (Garcia *et al.*, 2019), or gems float next to target words (Legault *et al.* 2019). In contrast, in another VE Legault *et al.* (2019) prevent drawing attention to the target words to encourage exploration, and only point out missed target words with arrows after some time has passed. Cho (2018) uses arrows on the floor to guide their participants past all target words, so participants are less free but can study the target words at their own pace. The commercial exploration-based game used for vocabulary acquisition by Alfadil (2020) places players in a VE (e.g. airport, cafe, cinema) where an NPC asks them to find specific objects.

Besides exploration, vocabulary acquisition can also be done through *conversation*. The IVR experience of Dobrova *et al.* (2018) consists of two VEs, one exploratory-based and one conversation-based. In the latter, participants interact with a virtual customs officer to learn new words and test their word knowledge. Other studies use a purely conversation-based approach. For example, the IVR environment of Cheng, Yang and Andersen (2017) is a Japanese teahouse with NPCs that provide different types of dialogue.

One line of research on vocabulary acquisition in IVR is based on *embodied cognition theory*, which revolves around the concept that cognitive processes (including language learning) are grounded in the body's interaction with the world (Vázquez *et al.*, 2018; Fuhrman *et al.*, 2021). For example, Vázquez *et al.* (2018) connect the movement of drawing specific symbols to target words.

Six of the studies compared word learning in IVR to learning in a non-IVR condition. The control condition in such studies varies between traditional methods (Legault *et al.*, 2019; Ebert, Gupta and Makedon, 2016; Vázquez *et al.* 2018; Alfadil 2020), a desktop equivalent (Cheng, Yang and Andersen, 2017), or gameplay videos (Tai, Chen and Todd, 2020). In some studies, participants in the IVR condition perform significantly better on the post-test (Legault *et al.* 2019; Alfadil, 2020; Tai, Chen and Todd, 2020). Sometimes they perform only slightly better (Cheng, Yang and Andersen, 2017), and in some studies they perform significantly worse than participants in the non-IVR condition (Ebert, Gupta and Makedon, 2016; Vázquez *et al.*, 2018). However, studies that looked at retention with a delayed post-test all found that IVR participants had a significantly higher retention rate than non-IVR participants (Ebert, Gupta and Makedon, 2016; Vázquez *et al.*, 2018; Tai, Chen and Todd, 2020). Thus, after some time had passed there was a larger word loss for non-IVR participants while IVR participants still remembered most of their learned words.

The discussed studies show promising results on retention of using IVR but did not recycle words or evaluate on recycling words. This raises the question if and how presenting learners with recycled words in either new or learned contexts might affect learning. This is what we focus on in our work.

3. Materials and method - experiment with Wics

Overall, the studies above indicate that IVR participants score slightly more often significantly better on post-tests than non-IVR participants, but the retention of the IVR group is always significantly higher. From this we conclude that a non-IVR condition can be excluded if the focus of a study lies on the long term effects of using IVR for vocabulary acquisition and participants are also tested on retention.

Recycling words helps a learner to learn more word aspects and to better retrieve words from memory. During explicit learning words are often recycled in the same context, for example when using flash cards, but words in everyday use are applied in a variety of new contexts and not just one learned context. A first video-based comparison already showed the possible effectiveness of including such variation in learning (Jeong *et al.*, 2010). This opportunity of learning words in various contexts, combined with recycling words, fits well with the ease of IVR environments to represent objects in various ways and in an explorative approach. To test the effect of learning in a new context we thus set up a study with an IVR environment where we compared the use of a *new* to a *learned context*. Balancing the needed attention span, the time needed for implementation and testing, and the possible effect of the intervention, we chose two new environments for the new context intervention. In these new environments both the objects and the surroundings were changed.

Each word has word aspects, such as its pronunciation or other words that occur with that word, and knowing more word aspects increases a learner's word knowledge depth (Nation, 2013). There are too many word aspects to learn all at once, so we limited the number of word aspects to present to participants to five. Learners tend to first establish the form-meaning link of a word so they can read or hear the L2 word and understand the meaning (Schmitt, 2008). Therefore we provided participants with (i) the sound of a word, (ii) the written form of a word and (iii) the meaning of a word, so they would be able to establish a form-meaning link. As Wics shows different visual representations of words in different contexts, we also included (iv) the concept of a word. To also include (v) where and when the word is likely to be encountered, we chose environments, objects and their location to match with what can be expected in the real world.

The independent variable of the experiment is the context in which words are repeated for learning a second and third time (i.e. new or learned), see Table 1, and the dependent variables are the test scores for the post-test and delayed post-test, how often each target word is studied, and the time spent in the VEs. Between the environments, participants had to take a break to not get overwhelmed through extended IVR use and from learning. The delayed post-test took place one week later.

Table 1: Setup of the experiment for both contexts

Learned context	New context
Learning environment 1	
Environment A Object representations A Object locations A	
Learning environment 2	
Environment A Object representations A Object locations Y	Environment B Object representations B Object locations B
Learning environment 3	
Environment A Object representations A Object locations Z	Environment C Object representations C Object locations C
Posttest	
Environment D Object representations D Object locations D	
Delayed posttest	
Environment E Object representations E Object locations E	

3.1 Wics

To broaden our potential user pool, we used a social IVR environment for our system, limiting thresholds such as installing programs before participating. From the platforms that provided active world-building we chose *Neos VR* (https://neos.com/) for its all-in-one solution and ability to visualise the world-building directly within the environment in IVR. Alternatives such as *VRChat*, *Sansar*, *Rec Room*, and *AltspaceVR* have a larger user base but less interaction capabilities.

Based on the personal expertise of the first author we decided to investigate learning Japanese words, similar to seminal work from Rose and Billinghurst (1995) and later Cheng, Yang and Andersen (2017). Based on related work showing that some participants could learn as many as 30 words in 20 minutes (Legault *et al.*, 2019), but also to prevent clutter in a reasonably sized environment, we decided to have 32 target words. Similar to Ebert, Gupta and Makedon (2016) we linked words to virtual objects in the environment. Where they used exclamation marks to indicate target words we instead, to further reduce clutter, offered users the possibility to *point and click* the objects, triggering words both in spoken and written format together with a written English translation. To provide an indication of where any missed objects are in the world, we provided a toggle to highlight these. This way, users could go through the experience at their own pace in an exploration-based manner while minimising frustration.

3.2 Worlds, words, and items

We selected five publicly available worlds that could be edited in 3D, that were suitable to contain a collection of varied objects, and that looked different enough to provide varying contexts, see Figure 1. The order and role of each environment balanced low poly and more realistic visualizations, fitting variation in learning, revisiting, and testing.

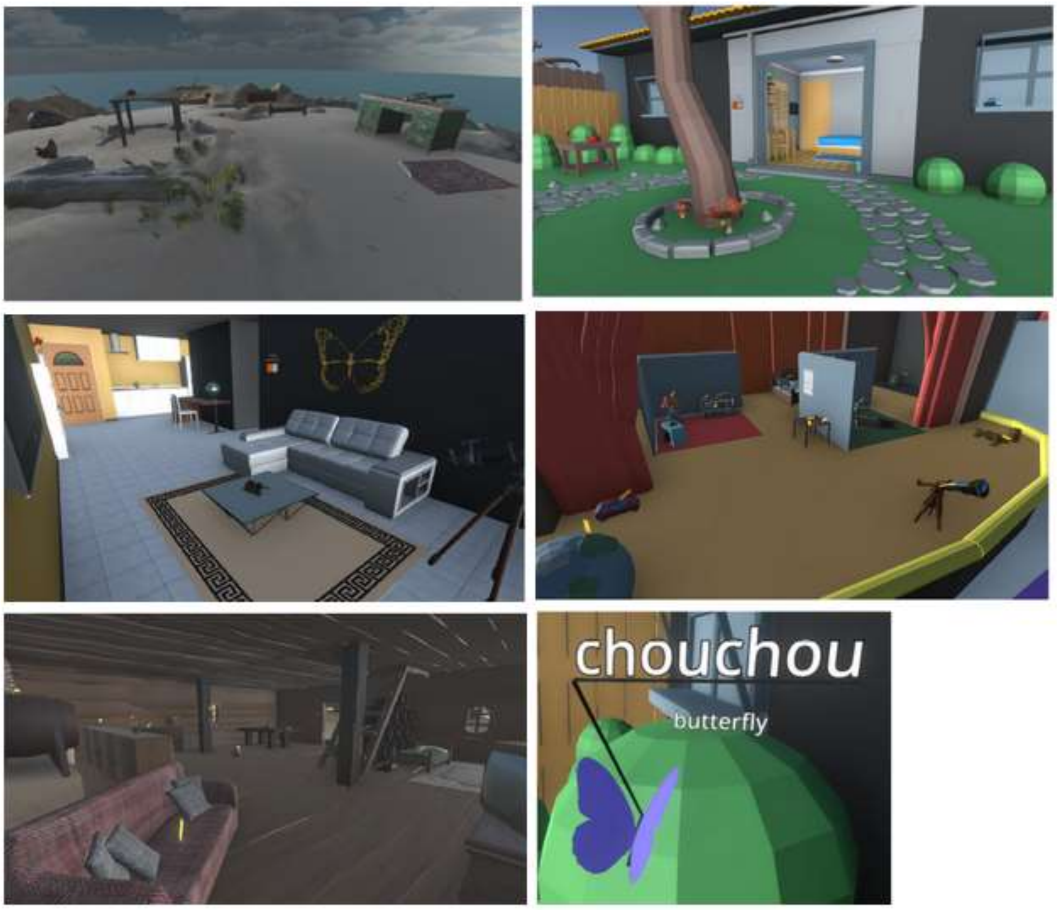

Figure 1: The selected worlds with 32 objects: A) an island, B) a bedroom with garden, C) an apartment, D) a theatre, and E) a barn, and an example of a triggered target word

To select the 32 target words, we excluded words of which multiple instances could be found in the environment (such as doors or stones, e.g. see B in Figure 1), and removed single objects that could be mistaken for a target word. We made sure to include short (one, two syllables), medium (three syllables), and long (four syllables) words and added a set of loan words from English (e.g. *naifu* for knife) to provide words that participants could quickly learn to create a feeling of success. Finally, pragmatically we needed at least five different visual representations of each object, which further reduced our word set. Our final word list was: avocado-*waninashi*, bag-*kaban*, ball-*tama*, bed-*beddo*, boat-*fune*, book-*hon*, bookcase-*hondana*, broom-*houki*, butterfly-*chouchou*, camera-*shashinki*, car-*kuruma*, chair/stool-*isu*, chest of drawers-*tansu*, coin-*kouka*, couch-*nagaisu*, desk-*tsukue*, earth-*chikyuu*, fish-*sakana*, flower-*hana*, garbage bin-*gomibako*, glasses-*megane*, key-*kagi*, knife-*naifu*, lamp-*ranpu*, mushroom-*kinoko*, pinecone-*matsukasa*, rug-*juutan*, shoe-*kutsu,* table-*tebburu*, teapot-*chabin*, telescope-*bouenkyou* and television-*terebi*.

In the two test worlds, all objects have a label assigned with a text field, see Figures 1D and 1E. After pointing and clicking on it, a text can be typed in the field using the platform's virtual keyboard. Participants are transported between the different worlds through portals. Inside the worlds, participants can choose their preferred manner of locomotion (e.g. teleport or walk with a controller) and whether to perform the test sitting or standing. Furthermore, there is a small introduction room, where the participants can learn how to activate objects and find not yet found items. The Wics environments are publicly accessible in Neos VR, which is a free platform available through Steam, using the following URL: https://tinyurl.com/4mzth2bh.

3.3 Participants and procedure

Experienced IVR participants were sought within the Neos VR community. A small description of the experiment was provided, including specifics regarding the expected duration (an hour) and how to sign up. A promotion room in Neos VR contained a textual description of the experiment, a link to the consent form, three L2 word objects to learn, a portal visualisation and a promotional poster. Interested participants could open a consent form and an information document with detailed specifics about the experiment. Participants were also sought through word-of-mouth. Our study was approved by our faculty's ethics committee (RP 2021-223).

Of in total 26 participants who started the experiment, four did not complete the delayed post-test and were excluded from the results. Six participants in the learned context condition and nine participants in the new context condition were experienced IVR users. They participated through Neos VR without further assistance from a researcher. Five participants in the learned context condition and two participants in the new context condition were non-experienced IVR users. They participated in a physical room with a researcher present. Age ranges were 21-25 years (eight participants), 26-30 years (eleven participants), 56-60 years (two participants) and 61-65 years (one participant)

3.4 Equipment

Neos VR user participants used their own hardware for Neos VR. The other participants travelled to a location with a room with an HTC VIVE (2016). The researcher kept an eye on what happened in IVR from a desktop in one corner. The space was cleared so the participant could move their arms around freely. A chair stood in the middle of the IVR play area for the participant to sit on if they preferred.

3.5 Stepwise procedure for the participants

Non-experienced IVR participants received a short explanation regarding the IVR introduction room they would start in. While the participants then went through the introduction room, the researcher stayed present in the physical room, monitoring them and offering guidance when necessary. Once the participant felt like they understood the interactions, the researcher left the physical room. Experienced participants, who did the experiment in Neos VR, received the experiment world through the Neos VR message system. They also started in the middle of the introduction room. When the participants had finished the introduction, the portal to the first world (learning environment) was opened.

The participants then were free to explore the environment and activate target objects to hear and read their L2 word plus an English translation. After the participants had found all the target objects in the learning environment, they were forced to take a short break of minimally three and maximally 15 minutes in a virtual break room, where they could watch a relaxing video. Participants could also take off their headset.

Figure 2: The toggle button to indicate missed items and open a portal to the next world

The experiment ended in a saving room filled with information in two parts: first, asking the participant if they wanted to get reminders about the second part of the experiment that must be done a week later; second, explaining to the participant how to send their data to the researcher before closing the experiment world. The data the participants needed to send were their filled-in answers from the test environment and logs of use of the learning worlds. After doing so, they were reminded that they should no longer actively try to learn or repeat the words in their head.

3.6 Analysis

Points were assigned by the first author: 1 point if the answer was phonetically correct, 0.5 points if it was almost correct (i.e. one syllable wrong or missing), similar to Ebert, Gupta and Makedon (2016) and Fuhrman *et al.* (2021). If participants already knew words prior to participating, we corrected the score to not include those words, dividing the initial score by 32 minus the N words that were known prior to participating. For example, if a participant knew six words beforehand, and scored 18 points on the words they did not know beforehand, then their performance score percentage was 18/(32−6) 100% = 69.23%, whereas their score would have been 24/32 100% = 75% without the correction.

The performance and retention scores seem to be close to normally distributed continuous data but are expected to break parametric assumptions due to a skewed data set. Therefore, we treated the performance scores as ordinal data and used a non-parametric test to compare the independent new and learned context conditions: the Mann-Whitney U test (Morgan, 2017). We also used the Mann-Whitney U test to compare the total time spent in the learning environments, the total time spent in the break rooms, the performance scores for the post-test and word retention.

4. Results

Wics worked mostly as intended and was received well. Only two of the 22 participants experienced a world crash in Neos VR. In both cases, the preceding VEs were quickly traversed a second time until the point where the crash happened. The times inside the learning VEs and break rooms of these two participants were left out of the results, while their test scores were included.

Table 2 shows the max score, score, and performance percentage per participant per condition, as well as the median and standard deviation between participants for both conditions. The max score refers to the maximum score a participant could get given the number of words they knew beforehand. The participants' performance on the post-test did not differ significantly between the learned context condition (Mdn = 89.06) and the new context condition (Mdn = 79.69), U = 58.5, z = -0.1286, p = 0.8977. Table 3 shows the max score, score, performance percentage, and retention percentage per participant per condition, as well as the median and standard deviation between participants for both conditions. Figure 3 shows the box plots of both the performance and the retention percentages of both conditions. The participants' performance on the delayed post-test (retention) did not differ significantly between the learned context condition (Mdn = 76.92) and the new context condition (Mdn = 72.73), U = 75, z = 0.9053, p = 0.3653. Table 4 shows the total time spent in rounded-down minutes in the learn and break environments per participant per condition, as well as the median and standard deviation between participants for both conditions. Time spent by participants in the learning environments did not differ significantly between the learned context condition (Mdn = 47.5) and the new context condition (Mdn = 26) during the experiment, U = 73, z = 1.6996, p = 0.08921. Time spent by participants in the break environments also did not differ significantly between the learned context condition (Mdn = 23.5) and the new context condition (Mdn = 16) during the experiment, U = 74, z = 1.7789, p = 0.0752.

Table 2: Post-test scores and performance of participants per condition

Learned				New			
Participant	Max score	Score	Perf. (%)	Participant	Max score	Score	Perf. (%)
L1	32	14	43.75	N1	25	15.5	62.00
L2	25	24	96.00	N2	32	25	78.13
L3	32	28.5	89.06	N3	32	25.5	79.69
L4	25	19.5	78.00	N4	31	30	96.77
L5	32	16.5	51.56	N5	32	30.5	95.31
L6	32	28.5	89.06	N6	20	19	95.00
L7	32	29.5	92.19	N7	32	31.5	98.44
L8	32	30	93.75	N8	32	7	21.88
L9	32	30,5	95.31	N9	28	24	85.71
L10	32	6.5	20.31	N10	32	17.5	54.69
L11	32	30	93.75	N11	28	16.5	58.93
Median	32	28.5	89.06	Median	32	24	79.69
Mean	30.73	23.41	76.61	Mean	29.45	22.00	75.14

Table 3: Delayed post-test scores and performance of participants per condition

Learned					New				
	Max score	Score	Performance (%)	Retention (%)		Max score	Score	Performance (%)	Retention (%)
L1	32	14	43.75	100.00	N1	25	13	52.00	83.87
L2	25	19	76.00	79.17	N2	32	12	37.50	48.00
L3	32	22	68.75	77.19	N3	32	14.5	45.31	56.86
L4	25	18.5	74.00	94.87	N4	31	22.5	72.58	75.00
L5	32	9	28.13	54.55	N5	32	23.5	73.44	77.05
L6	32	18	56.25	63.16	N6	20	16	80.00	84.21
L7	32	22.5	70.31	76.27	N7	32	30	93.75	95.24
L8	32	22	68.75	73.33	N8	32	3.5	10.94	50.00
L9	32	30	93.75	98.36	N9	28	16	57.14	66.67
L10	32	5	15.63	76.92	N10	32	9.5	29.69	54.29
L11	32	10.5	32.81	35.00	N11	28	12	42.86	72.73
Median	32	18.5	68.75	76.92	Median	32	14.5	52.00	72.73
Mean	30.73	17.32	57.10	75.35	Mean	29.45	15.68	54.11	69.45

Table 4: Time spent in the learn and break environments per participant and condition

Learned			New		
	Learn	Break		Learn	Break
L1	-	-	N1	45	35
L2	59	27	N2	31	13
L3	32	19	N3	20	9
L4	19	19	N4	50	30
L5	71	27	N5	38	23
L6	52	29	N6	25	17
L7	65	20	N7	24	15
L8	54	40	N8	-	-
L9	43	18	N9	24	12
L10	9	20	N10	18	11
L11	36	37	N11	27	22
Total	440	256	Total	302	187
Median	47.5	23.5	Median	26	16

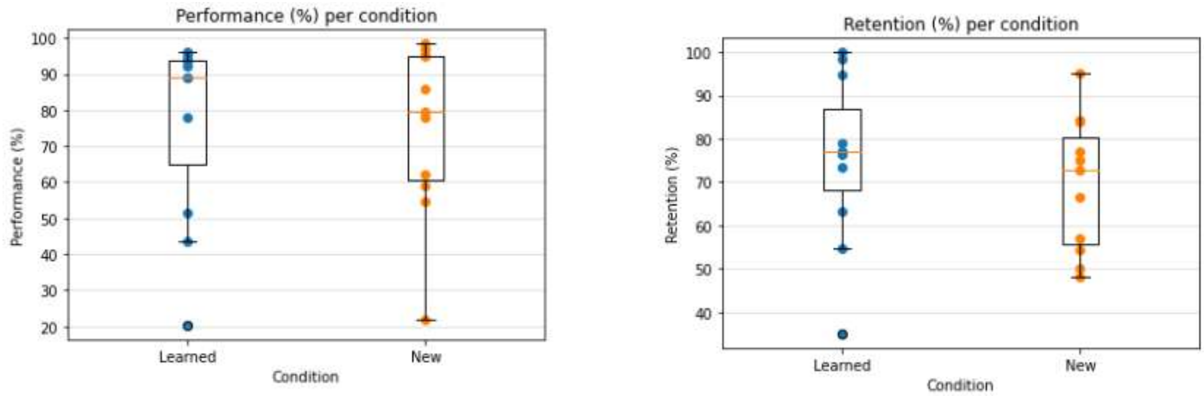

Figure 3: Box plot of both the performance (left) and retention (right) percentages of both conditions

5. Discussion and conclusion

In this research, we compared recycling words in IVR in a new context, where each of the three learning sessions took place in a different VE with changed visual representations of the words, and in a learned context, which was kept the same for each learning session. The two post-tests happened in a new context, also within IVR: the first post-test immediately after learning, and the delayed post-test a week after the first. We built a system called Wics to do the experiment. Each learning environment in Wics contained the same 32 target words, represented as visual objects. The participants could explore the environment and activate target objects to hear and read their L2 word (in Japanese) plus an English translation. Comments by the participants indicated the system was overall well received. For instance, participant N6 mentioned: *"It was a very interesting experience - certainly richer than simply trying to learn a list of words!"* With mean scores of 77% (learned context) and 75% (new context), results regarding the new words learned were similar and even slightly better than those by Cheng, Yang and Andersen (2017), the other IVR for learning Japanese, where participants had a mean score of 4.77 out of 8 possible words = 60%.

We found no significant difference between conditions for performance on the post-test, nor on retention, measured via the delayed post-test. This differs from the expected generalization of the results from Jeong *et al.* (2010), which showed participants who watched text-based videos within the same context (learned context) performed badly when they had to recall the word in a new context in a situation-based test. However, we see that our learned context group spent substantially more time in the learning environment. This difference was not significant at 95% confidence but would be at 90% confidence. We used Spearman's Rho to informally investigate a possible correlation between total learning time and performance on the post-test ($r_s = 0.33434$, $p = 0.14965$) and the total learning time and retention ($r_s = -0.03009$, $p = 0.8998$). The results suggest that a longer learning time might have benefited performance on the first post-test, but not retention. This also fits with the response of one participant *"I spent a lot more time in the first one, so it's not too surprising to me that my memory of that one is stronger than the subsequent two."* (N6). This leaves open the possibility that learning in a new context prepared participants better for retrieving words in the post-test (in a new context), but that this benefit was offset by participants in the learned context condition learning for a longer period. Furthermore, we used the same type of audio for all contexts, whereas some participants indicated an unanticipated auditory memory (e.g. N5: *"I'm generally not a very visual thinker, so for most of them I was just trying to recall the sounds of the words."*).

Perhaps our IVR provides too many new word aspects at once, where the visual representation word aspect is not one of the primary word aspects stored in memory for learners (Nation, 2013), and is only stored minimally and mostly lost over time. Perhaps offering learners ample opportunity to process words between learning sessions leads to multiple word aspects being stored more equally in memory.

To our knowledge, Wics is the first system to provide learners with multiple learning sessions, provide control over learning by activating an indicator for missed words, and offer post-tests inside IVR. For our study, we built on a comprehensive overview of IVR-based language acquisition with 16 papers (cf. the literature review by Palmeira *et al.* (2020), encountered later in our research, with nine papers). A full description of our literature review is outside the scope of this conference paper, so we refer the reader to an underlying master thesis (Bergsma, 2022). Similar to related studies, our participants showed enthusiasm for learning in Wics, discovering which virtual environment would await them next. The possibility of IVR to change VEs with ease should not be overlooked by language learning researchers as it can keep motivation up while providing learners with many contexts to learn, whether they switch contexts consecutively or repeat one context a few times. Furthermore, as participants showed a clear enthusiasm for going back to IVR to do the delayed post-test, other studies might consider doing their post-test(s) also in IVR, to base the results on a test method that complements the learning method that is studied.

References

Alfadil, M. (2020) "Effectiveness of virtual reality game in foreign language vocabulary acquisition", *Computers & Education*, Vol 153, pp. 103-893.

Bergsma, T.S.L. (2022) *Vocabulary acquisition in new and learned contexts using immersive virtual reality.* MSc thesis. University of Twente.

Cheng, A., Yang, L. and Andersen, E. (2017) "Teaching language and culture with a virtual reality game", *CHI 2017*, pp 541–549.

Cho, Y. (2018) *How spatial presence in VR affects memory retention and motivation on second language learning: A comparison of desktop and immersive VR-based learning.* MA thesis. Syracuse University.

Collins, N., Vaughan, B. and Cullen, C. (2020) "Designing contextually: An investigation of design-based research to promote situated Irish language identity through virtual reality", *6th International Conference of the Immersive Learning Research Network*, pp 147–154.

Dobrova, V. *et al.* (2018) "Virtual and augmented reality in language acquisition", *International Conference on the Theory and Practice of Personality Formation in Modern Society*, pp 218–223.

Ebert, D., Gupta, S. and Makedon, F. (2016) "Ogma: A virtual reality language acquisition system", *9th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, pp 1–5.

Fuhrman, O. *et al.* (2021) "The moving learner: Object manipulation in virtual reality improves vocabulary learning", *Journal of Computer Assisted Learning*, Vol 37, pp 672–683.

Garcia, S. *et al.* (2019) "An immersive virtual reality experience for learning Spanish", *CHI 2019*, pp 151–161.

Godwin-Jones, R. (2018) "Contextualized vocabulary learning", *Language Learning & Technology*, 22(3), pp 1–19.

Jeong, H. *et al.* (2010) "Learning second language vocabulary: neural dissociation of situation-based learning and text-based learning", *Neuroimage*, Vol 50, No. 2, pp 802–809.

Jia, T. and Liu, Y. (2019) "Words in kitchen: An instance of leveraging virtual reality technology to learn vocabulary", *IEEE International Symposium on Mixed and Augmented Reality Adjunct*, pp 150–155.

Legault, J. *et al.* (2019) "Structural brain changes as a function of second language vocabulary training: Effects of learning context", *Brain and Cognition*, Vol 134, pp 90–102.

Loewen, S. and Sato, M. (2017) *The Routledge handbook of instructed second language acquisition*, Routledge, New York.

Nation, I.S.P. (2013) *Learning vocabulary in another language*, Cambridge University Press, Cambridge.

Palmeira, E.G.Q. *et al.* (2020) "The use of immersive virtual reality for vocabulary acquisition: A systematic literature review", *Anais do XXXI Simpósio Brasileiro de Informática na Educação*, pp 532–541.

Schmitt, N. (2008), “Instructed second language vocabulary learning”, *Language Teaching Research*, Vol 12, No. 3, pp. 329–363.

Tai, T., Chen, H.H. and Todd, G. (2020) “The impact of a virtual reality app on adolescent EFL learners’ vocabulary learning”, *Computer Assisted Language Learning*, pp 1–26.