



An Immersive Virtual Reality Experience for Learning Spanish

Sarah Garcia, Denis Laesker^(✉), Derek Caprio, Ronald Kauer, Jason Nguyen, and Marvin Andujar

University of South Florida, Tampa, FL 33620, USA
{sarahgarcia, dlaesker, derekcl, rkauer, jasonnl}@mail.usf.edu, andujarl@usf.edu

Abstract. While increases in the development of virtual reality (VR) have expanded the possibilities of learning foreign languages, research for computer assisted language learning in VR is lacking [1]. Virtual reality headsets continue to become more accessible to end users, and there is a current lack of educational language learning applications for such platforms. There are many tools and applications for language learning, however most lack the aspect of immersion – a proven method of learning a foreign language as seen in experiments involving overseas language experiences [2]. Therefore, this paper presents an interactive VR experience for learning Spanish through an immersive interaction and game-play, followed by the results from a focus group regarding user experience feedback. Additionally, this paper will detail the design and development of the current prototype. Results indicate that participants found this method of language learning to be more enjoyable than traditional methods because it did not feel like they were studying.

Keywords: Virtual reality · Language education · Gaming · Language learning · Game-based learning · Immersion

1 Introduction

Immersion has been shown to be a proven method for learning a foreign language, as seen in studies that involved overseas language experiences for college students, where students experienced statistically significant improvement in comprehension scores [2]. However, traveling and living overseas is costly, making it an unaffordable learning method to many. In spite of that, with the use of every day more affordable Virtual Reality technology, such as the Google Cardboard being used in education [3], one may take advantage of an immersive experience from the comfort of their own home. In this paper, we propose a VR experience that addresses the real-world cost limitation of traveling overseas to achieve immersion with another language, by providing a virtual environment where users can interact with objects and their surroundings in a different language for a lower price and at any desired time. The proposed VR language learning experience created in Unity and scripted for the Oculus Rift head-mounted display, takes advantage of the immersive capabilities offered by the Rift and extends them to the process of language learning. Also, using this application, learners can put

into practice in real-time what they have learned in the 3D environment. This 3D experience allows users to explore objects and their Spanish translations in a free-roaming practice round, then apply this knowledge during game-play. Game-play involves the use of a search-and-find format, where players are given items to search for within the scene. After finding the correct item, points are awarded and a new item to search for is given. Users can view their scores to measure their success as they correctly identify objects within their environment. Currently allowing for exploration of the inside of an average home and interaction with common household items, this application can be applied to most real-world locations.

This application can be used as a supplement to Spanish classes in schools. Participants from the focus group mentioned that they felt more engaged and believed they would be able to study for longer periods of time using our application than with traditional book-learning methods. This suggests that work assigned to be completed on our application has potential to be more effective than the former. Furthermore, teachers could assign game-play for the students to complete on their own in the classroom or at home.

An important consideration to take into account for use in schools is affordability, and this language learning experience can be adapted to different affordability levels. Although this application was built and tested with the Oculus Rift, it can easily be ported to other, more affordable, VR devices such as the Oculus GO or even devices such as the Samsung Gear or Google Cardboard, that utilize smartphones for their VR display. With the ubiquity of smartphones today, this would enable those who cannot afford more expensive devices such as the Oculus Rift to still be able to improve their Spanish by using the application. Users could download our application on their smartphone and pair it with an economically priced headset such as the Google Cardboard, thus allowing them to immerse themselves into a Spanish speaking environment from anywhere in the world. Additionally, this application can be used by individuals who are interested in learning Spanish on their own or outside of a classroom.

In this paper, we discuss the user interaction involved in the language learning experience, as well as detail methods used for its creation. We then discuss the methodologies used for a focus group conducted using the proposed language learning experience. This focus group was conducted in order to gain feedback on the current working prototype, so that data collected from users can be implemented into future iterations as well as to gain interest regarding the application. Feedback from the focus group is discussed, as well as potential use cases for the VR experience in the real world. The paper ends by discussing plans for future work to be completed.

2 Related Works

Employing the use of VR technology in educational applications has gained popularity in recent years [4]. However, although of its high potential, little research can be found regarding language learning in VR [5]. VR environments provide users with an interactive and immersive experience within an artificially generated virtual world [4]. As described in [6], interactivity is achieved by designing specialized interfaces and

providing users with real-time feedback. Immersion, on the other hand, is described as being divided into physical and mental immersion. The former can be accomplished by letting users navigate and control objects within the virtual environment. The latter is characterized by how engaged the user is.

2.1 Active Learning and the Search-and-Find Format

The use of the search-and-find format, where users are given prompts to search an environment for specific items, as an active learning technique has shown various benefits in the learning process. For instance, a study using this format, where undergraduates learned how to conduct research through exploring databases and reputable websites has proven to be both engaging and cognitively challenging [7]. *VEC3D* [7] is an immersive and interactive web-based online virtual environment (*VEC3D* website) of an English classroom that employs the use of a “Virtual Scavenger Hunt” as one of its goal-based scenarios. In this scenario, participants act as scavengers competing against a countdown timer to locate and name virtual objects scattered throughout the virtual environment. Moreover, a one-year ethnographic study with Taiwanese undergraduate students majoring in English language teaching with various levels of proficiency was conducted. Results showed that students felt motivated to use communication strategies, felt comfortable in the environment, and demonstrated positive attitudes towards the application. Differing from *VEC3D*, our created game employs the use of VR and the Oculus Rift head mounted display, increasing the notion of immersion.

2.2 Virtual Reality for Language Learning

Immersion plays an important role in learning a new language [4]. This immersion can be achieved by travelling to a place where the target language is predominately spoken, however, such relocation can be very costly. Therefore, the benefits of using VR to provide learners with immersive environments for a small fraction of the cost would only be beneficial. For instance, by adapting a 3D video game called *Crystallize* to work in VR, users were able to learn when to bow in Japanese greetings [9]. This was done through detecting a change in angle using an Oculus VR Headset (signifying the bowing one would do in the real world) when the user is presented with an in-game prompt from a non-playable character (NPC). Moreover, a formative study with 68 participants using both VR and non-VR versions of the game was conducted to determine whether participants would learn when and how to bow. Users were first taken through a tutorial that gradually increased the time in which the system would alert them when to bow. Learning was measured by the ability of a user to perform an unprompted bow (bowing without explicitly being told to do so). Results suggest that users were able to learn how to bow and felt more involved in Japanese culture when interacting in the VR version.

Among other language learning systems created, *SeLL* [10] is an English learning system aimed at improving oral skills that combines VR, speech recognition and pronunciation assessment through the use of artificial intelligence (AI) in order to provide an acoustic and visual immersive experience to its users. The system allows its

users to interact with the virtual environment, which includes various activities designed for oral communication with either other learners or virtual characters equipped with intelligent dialogues. Moreover, the system provides users with feedback on their pronunciation, fluency, and expression at the end of learning.

3 User Interaction

After setting up the Oculus Rift hardware necessary, users launch the game application. Users begin their experience in a free-roaming environment, in which they interact with the environment with no defined gameplay goals to familiarize themselves with their surroundings and touch controller navigation. The left-hand controller joystick is used to walk around in the environment and the right-hand controller joystick is used to look around, moving the screen in the direction of the joystick 30°. Participants can also look around by physically turning their head. After users have familiarized themselves with the virtual environment, the gameplay section can be started. Here, users are prompted with both audible (implemented from Google Translate) and readable cues such as “¿dónde está la cama?” (“where is the bed?”) or “¿dónde está la televisión?” (“where is the television?”), indicating which object they must search for next. The readable cue can be found near the bottom of the user’s screen, and their current score is found towards the top, as seen in Fig. 1.



Fig. 1. Game-play depicting readable cues given to the user (bottom of image) and their current score (top of image).

After reading the cue, users would then search for the prompted item and select it by hovering over it with the gaze pointer and clicking the ‘A’ Button. Time taken to find all search-and find items lasts as long as needed to find all items in the scavenger

hunt, usually about five to ten minutes. When searching for the needed item, users are able to see which items can be selected by hovering over them with the gaze pointer. Items that can be selected become highlighted blue, as seen in Fig. 2 on the left. If an item selected is incorrect, the item will instead highlight red, as seen in Fig. 2 on the right. After finding the correct item, its name would be audibly played back once again to the user and their score would increase by 10 points before being prompted to find the next item. This continues until all objects were found. Figure 3 shows the state diagram of the game, depicting the continuous flow of tasks given to the user until all objects in the sequence are found.

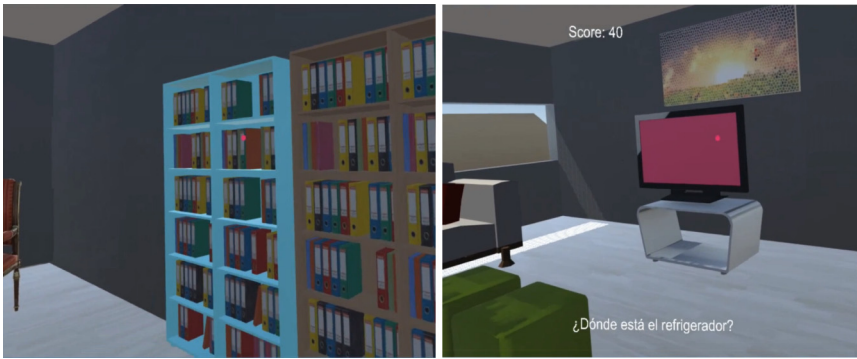


Fig. 2. Game-play depicting the changing of items when hovered over as correct (left) and incorrect (right).

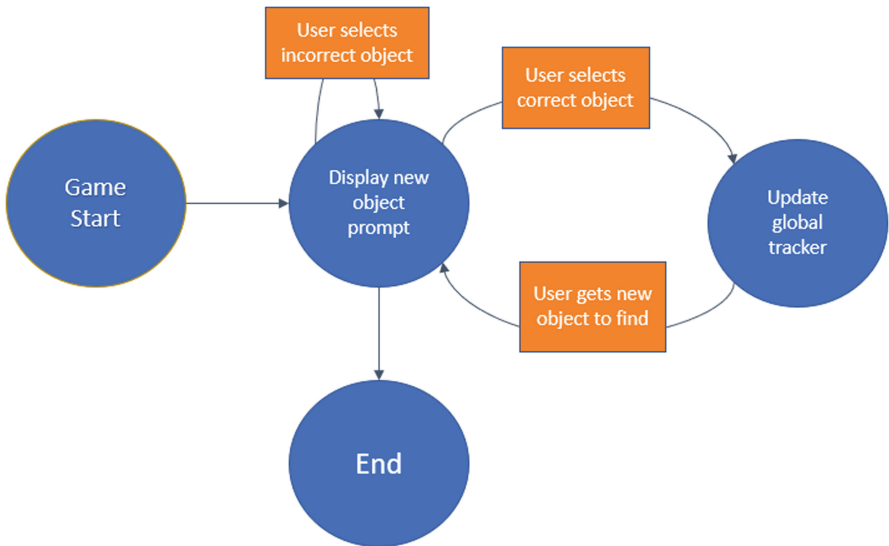


Fig. 3. Diagram of user game-play states.

4 Experience Creation

4.1 Hardware

Development of the system was performed using the Unity Video Game Engine due to its popularity and gradual learning curve. Although the created system was initially built to be used with the Oculus Rift VR headset and its complementary Oculus Touch Controllers (Fig. 4), the application is also portable to other VR headsets. The headset used is equipped with a 3.5in, OLED display, 2160×1200 resolution, and 90 Hz refresh rate.



Fig. 4. User with Oculus Rift and touch controllers.

4.2 Software Architecture

A modular and scalable software architecture was developed for this project. This allows for future developers to quickly understand the architecture and add, or remove, to and from the project as needed. All aspects of the environment and game-flow are controlled by the GameManager component, as seen in Fig. 5. This component is responsible for keeping track of the current target item, which prompt is to be displayed, and which audio files are to be played for each prompt and item. The AudioManager component allows audio files to be played and heard by the player.

Items, prompts, and audio are all objects in the game. These are stored in three global dynamic arrays that can be seen and modified through the Unity Inspector

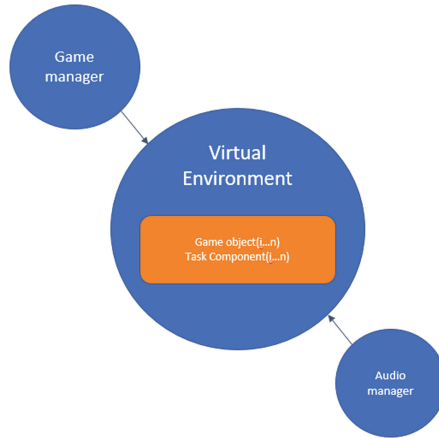


Fig. 5. Depiction of the software architecture used in creation of the game-play

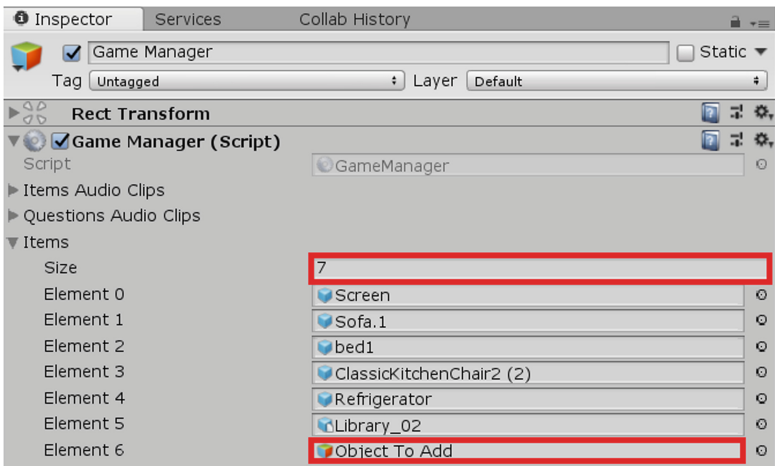


Fig. 6. Elements within the Game Manager seen in the Unity Inspector.

component. This facilitates adding more objects to be managed by the GameManager, therefore, increasing the number of items a user would need to find during gameplay. Figure 6 shows the simplicity of this process. The size of the array, or the number of objects needed, can be easily modified by setting the “Size” to the desired number. In addition, objects can be added by being dragged into an element slot. Moreover, this also makes it easier to expand the 3D environment as seen fit. Player movement and interaction with the environment was achieved by using the pre-built OVRPlayerController provided by the Oculus Software Developer Kit. The 3D environment was created using Unity primitives as well as Asset Packages obtained from the Unity Asset Store. The GameManager uses a global tracker to oversee a player’s

progress within the scavenger hunt. All objects that are part of the scavenger hunt contain a Task component, which determines whether a clicked object is the correct one, or not. In case the former happens, the global tracker is updated and the GameManager handles which prompts are to be displayed next. If the latter happens, the item clicked is highlighted in red denoting “wrong item”.

5 Methodology

5.1 Study Design

We conducted a focus group to collect qualitative data regarding the current prototype of the immersive experience to gain insight to improve the capability to teach a foreign language. This focus group consisted of participants completing a pre-questionnaire, trying the current prototype, and finally a focus group session of a group of four. The pre-questionnaire was used to gain information regarding general demographic information, experience with learning a new language and prior usage of any virtual reality system. Participants then tested the virtual environment one at a time. A screen recording of each participant’s time in the virtual environment was saved for later analysis of how participants interacted with the prototype. Afterwards, participants took part in a focus group guided by an interviewer that lasted approximately one hour, to discuss their experience and provide feedback regarding the prototype they tested.

5.2 Participants

Data was collected with a total of four participants, one female and three males, all between the ages of 18 to 24. Three of the participants were Asian and one was African-American. All were undergraduates at the University of South Florida and had some experience in learning a foreign language, though none were fluent in Spanish. Three of the four participants had experience with video games, which may show their ability and speed to adapt to the controllers. Additionally, three of the four participants had experience using a virtual reality headset.

5.3 Procedure

After being put into a group of four, participants were asked to:

1. Complete a pre-experiment survey before starting to interact with the virtual environment.
2. Participants entered the virtual environment, one at a time, for ten minutes each. Each participant was first directed to roam freely around the environment for five minutes. This was done so that participants who did not have experience with virtual reality would have the level of familiarity needed to properly interact with the application and provide valuable feedback.
3. After this, the participant was prompted in Spanish to search for specific objects in the virtual environment. This search-and-find phase lasted an additional five minutes for each participant.

After all the four participants had their chance to interact with the virtual environment and find all necessary objects, the focus group was conducted. During the focus group participants were asked questions regarding their experience with the virtual environment. These included how natural they found the interaction, what they thought of this immersive method of language learning, their likes, dislikes and suggestions to improving the prototype.

6 Results

6.1 Environment Interaction

During analysis of screen recordings taken of each participant while using the prototype, it was noticed that participants experienced confusion when trying to select certain items. This confusion occurred when there were multiple instances of that object in the room, however only one of them could be selected to complete the search-and-find task. This was seen when two participants had trouble when trying to select the couch item in the home, as two couches were present but only one could be selected (i.e. only one would highlight blue). Additionally, participants chose different methods of exploring their surroundings. Two participants primarily used the right-hand touch controller joystick to change their view of the room, while the other two instead chose to physically turn their head in order to see other areas.

6.2 Game-Play Feedback

All participants reported in the survey that they would be interested in using a virtual environment to learn a foreign language. During the focus group, participants often mentioned that they found language learning in virtual-reality to be more fun than traditional learning methods. Remarks regarding the environment included that this method of learning was “*almost like playing a game*” and that they could “*study longer without taking a break*”. Participants also expressed that they believed learning while immersed in a virtual environment helped them to better remember the names of the objects since they were able to interact with them. Feedback from the focus group also suggests that the User Interface (UI) design can be improved, as participants had mixed reactions regarding the placement of the score and search-and-find task prompts displayed on the screen. Some participants viewed it as intrusive, describing them as “*on their face*”, others did not notice the score display at all during gameplay, implying it was not very visible. Participants also mentioned a desire for a more complex environment, in order to learn harder words and make the environment feel more realistic by adding a backdrop around the virtual environment. In addition to more complex environments, higher level vocabulary and more objects were also commonly suggested by the focus group participants. While all participants found their movement in the virtual environment to be realistic and intuitive, three of the four were critical of the panning feature used to look around using the Oculus handheld remote, stating that it was not smooth. The panning of the character and camera 30° done by use of the joystick located on the right-hand controller was described as too sharp of a turn, causing them nausea every time the function was used. Focus group’s participants also

discussed that the directions for search-and-find tasks seemed to have “come out of nowhere” and did not like that the reading of the prompt came from an unseen narrator. This feedback can be remedied with the implementation of transition scenes between tasks and the potential implementation of a character in view that speaks the tasks to the user so that prompts given feel less sudden. All participants expressed enthusiasm about this method of language learning, finding it more entertaining and effective than traditional study methods.

7 Conclusion

Feedback from the focus group indicates that participants find language learning in a virtual environment using our prototype to be more enjoyable than traditional language learning methods. Positive responses imply that the use of a language learning system such as this is wanted, although studies regarding language retention and vocabulary gain are needed to investigate its success in teaching information to users. The feedback suggests improvements such as re-design of the UI, a more spacious virtual environment with a more diverse list of objects. The results from the focus group will be incorporated in future iterations of the prototype to create a more immersive environment for language learning.

8 Future Work

As this is the first iteration of the application, improvements will be made using feedback gathered from the focus group. These improvements include using changes in design choices, such as switching over to the use of touch controllers for item selection instead of using the user’s line of vision, and improvements regarding character view-panning to reduce player nausea. Furthermore, participants indicated to increase the amount of scavenger hunt tasks and the environment map size. Increase in map size includes the addition of other types of locations one might encounter often or find useful vocabulary in, such as a grocery store or restaurant. Participants also expressed that the addition of differing difficulty levels would be beneficial. Implementation of elements to make the reading of new search-and-find prompts more user-friendly can be added, such as addition of transition scenes and a friendly avatar speaking so that new task announcements do not feel as sudden. Finally, the addition of a voice recording feature, where the user would be able to record and hear playback of themselves saying item names would be useful to assist users in their pronunciation skills while allowing them to compare their tries to correct pronunciations.

References

1. Lin, T.J., Lan, Y.J.: Language learning in virtual reality environments: past, present, and future. *J. Educ. Technol. Soc.* **18**(4), 486–497 (2015)
2. Savage, B.L., Hughes, H.Z.: How does short-term foreign language immersion stimulate language learning? *Front.: Interdisc. J. Study Abroad* **24**, 103–120 (2014)

3. Parmaxi, A., Stylianou, K., Zaphiris, P.: Leveraging virtual trips in Google expeditions to elevate students' social exploration. In: Bernhaupt, R., Dalvi, G., Joshi, A., Balkrishan, D.K., O'Neill, J., Winckler, M. (eds.) INTERACT 2017. LNCS, vol. 10516, pp. 368–371. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68059-0_32
4. Huang, H.M., Rauch, U., Liaw, S.S.: Investigating learners' attitudes toward virtual reality learning environments: based on a constructivist approach. *Comput. Educ.* **55**(3), 1171–1182 (2010)
5. Ebert, D., Gupta, S., Makedon, F.: Ogma: a virtual reality language acquisition system. In *Proceedings of the 9th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, p. 66. ACM, June 2016
6. Huang, H.M., Liaw, S.S., Chen, W.T., Teng, Y.C.: Developing a virtual reality learning environment for medical education. In: *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, pp. 1320–1329. Association for the Advancement of Computing in Education (AACE), October 2009
7. Jones, J.A., Smith, S., Royster, M.: The scavenger hunt as an active learning technique. *NACTA J.* **61**(1), 94 (2017)
8. Shih, Y.C., Yang, M.T.: A collaborative virtual environment for situated language learning using VEC3D. *J. Educ. Technol. Soc.* **11**(1), 56–68 (2008)
9. Cheng, A., Yang, L., Andersen, E.: Teaching language and culture with a virtual reality game. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 541–549. ACM (2017)
10. Guo, J., et al.: SeLL: second language learning paired with VR and AI. In: *SIGGRAPH Asia 2017 Symposium on Education*, p. 7. ACM, November 2017