Yang, J. C. et al. (Eds.) (2018). Proceedings of the 26th International Conference on Computers in Education. Philippines: Asia-Pacific Society for Computers in Education

Learning English Conversations in an Immersive Virtual Reality Environment

Ji-Yeon JEONG[a], Su-Jong SEO[a], Ye-Jin HAN[b], Sun-Jung KIM[b] & Heisawn JEONG[a, c*]
[a]*Department of Interaction Design, Hallym University, South Korea*
[b]*Department of Convergence Software, Hallym University, South Korea*
[c]*Department of Psychology, Hallym University, South Korea*
*heis@hallym.ac.kr

Abstract: Direct exposures to and contacts with the target language environments are useful in learning second languages. Unfortunately, it is not always possible for average learners to travel to another country or have a chance to converse with a native speakers of the language. Virtual Reality (VR) technology can provide a solution. We developed an immersive Virtual Reality (VR) English conversation training program in which learners can practice English conversations with a virtual character. We tested a pilot version of the program in this study. Six college students received the training either with a headset (HMD condition) or with a computer screen (non-HMD condition). We assessed their learning outcomes as well as their language anxiety and also asked them for suggestions about how to improve the environment. No significant differences in learning outcomes and language anxiety were found between the HMD and the Non-HMD conditions yet, but participants open-ended responses provided insights about how to further improve the VR training program.

Keywords: Technology Enhanced Language Learning(TELL), Immersive Virtual Reality(VR), Language Anxiety

1. Introduction

As the world becomes smaller, it is increasingly important to be able to communicate with people from other cultures. As English became a standard language for global communication, students around the world strive to learn and speak fluent English. Learning a foreign language such as English, however, is quite challenging for many students in Asia including Korea because of the differences in phonetics and grammatical structures as well as cultural backgrounds. Technology has played an important role in learning foreign languages since the invention of writing, audio-recordings, and other multimedia resources (Chun, Smith, & Kern, 2016). With the development of digital technology, an increasingly diverse range of technologies are used to help students to learn foreign languages. Personal computers and the Internet allow an instant access to electronic dictionaries so that learners can look up vocabularies as they read. Automatic speech recognition program provides instant feedback and helps learners to correct and adjust their pronunciations. Chat and social networking sites also provide a space for learners to interact with people speaking the target language (Golonka, Bowles, Frank, Richardson, & Freynik, 2012).

Virtual Reality (VR) technology is also making a rapid progress in this area. VR environments provide a realistic and yet safe environment for learners in which they explore and practice their skills. There exist a number of commercial VR training programs on the market. For example, a product called *Mondly* (https://www.mondly.com/), available in 28 languages, provides learners a VR environment in which they converse with virtual characters and receive instant feedback on their pronunciation and vocabulary. Another example is *Talkish* (https://talklish.com/main/), available only in Korea at the moment, in which users learn the English language by interacting with virtual characters. Learners play the role of a tourist, doctors, scientists in these environments and are usually given a specific task or missions to complete (e.g., ordering in a restaurant). Although still in its infancy, the use of VR technology for language learning is very promising. Practicing and learning such skills directly in the target culture is often too costly and/or

not feasible to most learners. In addition, the traditional method of teaching and learning second language emphasize the acquisition of grammatical knowledge and vocabularies. But VR provides an environment in which learners can practice and use these knowledges and skills. It is also socially safe because learners need not worry about how their failures and errors may be perceived by their interaction partners. According to Maclntyre and Gardner (1994), many second language learners experience language anxiety, that is, a kind of 'tension or fear' when they listen to or speak the second language. The language anxiety can be a significant obstacle in developing target language competencies. The virtual environment can provide them with a safe environment in which they can practice and learn their language skills without worrying too much about how their language abilities might be perceived. Given their potentials, it is useful to develop an effective VR training program for second language acquisition. The goal of the current study is to describe a prototype of VR training program and report on a pilot study to test its effectiveness.

2. Methodology

2.1 Participants

Six college students participated in the study. Participants were randomly assigned to HMD condition in which they used an immersive VR with a headset or the Non-HMD condition in which they used computer monitors (three participants in each condition).

2.2 VR conversation training program

The VR conversation training program consists of three scenarios (e.g., packing for a trip, potty-training puppies, working overtime) from a book about English conversations by Redstone, Cunningham, and Lee (2010). In each scenario, participants first listen/read a dialogue between two a man and a woman (see Figure 1). After participants read and listen to the dialogue, each turn of the dialogue is spoken one at a time, and participants are asked to repeat it. Participants can repeat this process until they feel ready to practice the conversation with a partner. In the conversation practice, participants practice the dialogue with a virtual character (see Figure 2). The virtual character initiates the conversation, to which participants generate their line of responses. Participants signal the beginning and end of their speech using a controller or keyboard. The system processes the recorded responses and presents the results of their speech recognition back on the screen in terms of accuracy. The accuracy represents the degree of match between users' response and the correct dialogue moves stored in the system. If the entire turn matches, 'Accuracy 100%' is displayed on the screen. If nothing matches, the accuracy is 0%. Participants can proceed to the next turn regardless of the accuracy. The conversation practice ends when participants speak their last turn. The same process was repeated for the three learning scenarios.

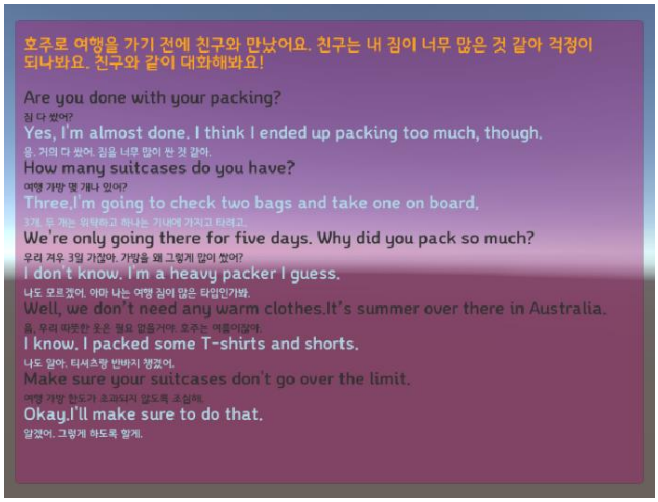

Figure 1. A dialogue sequence used in the VR Conversation Training Program

Figure 2. Conversation practice with a virtual character

The conversation training program was presented in immersive VR (HMD condition) or on a screen (non-HMD condition) in this study. In the HMD condition, participants viewed the whole learning materials with a HMD (head-mounted display) device (Oculus Rift CV1 model). Participants in the HMD condition controlled the sequence of the presentation with an Oculus controller. In Non-HMD condition, the program was presented on a 22 "monitor. Participants used a keyboard to control the program sequence. The system was empowered by Intel Core i7 CPU (Intel Core i7-2600 @ 3.40 Hz), 8GB memory, NVIDIA GeForce GTX 750M Ti graphics card, MS Windows 10. The computer was equipped with a 64-bit operating system. The training program was developed using Unity's game engine software (Unity 2018.2.1f1). For speech recognition, Microsoft Speech Platform from Microsoft was used. The n-gram algorithm is used to measure the similarity by comparing the correct answer with the user's answer. We used 4-gram as the phoneme.

2.3 Measurement Tools

2.3.1 Language Anxiety Questionnaire

FLCAS (Foreign Language Classroom Anxiety Scale) designed by Horwitz, Horwitz and Cope (1986) was used to assess participants level of language anxiety. The questionnaire consists of 16 statements (e.g., I would probably feel comfortable around native speakers of the foreign language) on a five-point Likert scale. This questionnaire was administered both before and after the training.

2.3.2 Presence Questionnaire

In order to measure the presence of virtual reality, we used the questionnaire designed by Witmer, Jerome, and Singer (2005), but excluded questions irrelevant to the current system (e.g., questions about haptic interface) were excluded. The questionnaire contained nine statements (e.g., I felt that the character in the virtual environment actually existed in front of me) on a 7-point Likert scale.

2.3.3 Pre-test

A pre-test was administered to assess participants' prior conversational proficiencies in English. The test questions were taken from several English conversation texts (Armstrong, & E2K, 2012; Hong, Nam, Nam, Lee, & Cha, 2015; Lee, 2007). Participants were asked to respond to questions in English (e.g., Have you ever thought about study in abroad?). They had to answer nine questions, with three questions in easy, mid-level, and difficult level each. As was in VR environment, they pressed a recording button when they begin speaking and pressed another button when they finished. Their recorded responses were later scored in terms of appropriateness and fluency on. Appropriateness of the response refers to how well the participant 's English responses are relevant to the situation, and fluency was assessed using indicators such as the relative speed of their responses and

amount/frequency of hesitations and pauses. Each item was scored using a 3 - point scale with the highest score for each question being 9. The scoring criteria were modified from Kim and Gong (2008).

2.3.4 Post-test

After all the participants finished the training, they took a post-test in which participants were asked to reproduce the conversation they learned during the training. One of the turns from the training dialogues (e.g., "You look tired. Did you have a long night?") was presented in Korean, to which participants responded with an appropriate turn in English. Participants' verbal responses were recorded and later scored in terms of accuracy (e.g., how well they reproduced the learned dialogues), appropriateness (e.g., how appropriate their responses were even though they may not be what they learned), and fluency (e.g., prompt responses with little hedging).

3. Procedure

The study was conducted in a lab. Participants, upon arrival at the lab, first answered a few questions about their demographic information (e.g., age, years of English education) and their prior experiences with English. For example, they answered how favorable their attitude toward English is and what is their preferred method of learning conversational English among (1) native speaker, (2) VR, (3) texts with audio, and (4) other methods. Participants then took the pre-test of English conversational proficiency test and completed the Language Anxiety questionnaire.

During the training, participants learned and practiced three scenarios of English conversations. Participants in the HMD condition interacted with the system while wearing a headgear, whereas those in the non-HMD condition interacted with the virtual character on the screen. Only one of the three participants in the HMD condition have used HMD before the experiment. All participants were asked to practice the manipulation of the controller and get comfortable with the devices. Training began when participants indicated that they are ready.

On the completion of the training, participants filled out the presence questionnaire and took the post-test, which was followed by the Language Anxiety Questionnaire and questions about their attitudes and preferred methods. Lastly, we asked whether participants have any feedback and suggestions about the VR system.

The study took from 40 minutes (non-HMD condition) to an hour (HMD condition) on average. Participants in the HMD condition needed extra time to get used to the VR environment.

4. Results

Given the small sample size, we describe the main descriptive differences between the two conditions. Participants in the HMD condition had higher level of pre-test conversational proficiency ($M = 7.37$, $SD = 1.39$) than those in the non-HMD condition ($M = 5.56$, $SD = 2.08$). Their post-test performances were about the same ($M = 9.78$, $SD = 1.94$ in the HMD condition versus $M = 9.70$, $SD = 1.58$ in the non-HMD condition).

The two conditions did not differ much in their language anxiety before the training: Language anxiety score was 3.38 ($SD = 1.43$) in the HMD condition and 3.38 ($SD = 1.12$) in the non-HMD condition before training. Anxiety level did not change much after the training, although the level of language anxiety was somewhat reduced in the non-HMD condition ($M = 3.25$, $SD = 1.43$) as compared to HMD condition ($M = 3.37$; $SD = 1.19$). It seems that participants who practiced the conversation with the computer screen felt less anxious after training, whereas participants who used headgear did not. It is likely that the unfamiliarity of the devices might have prevented them from lessening their anxiety.

As for the presence of the VR questionnaire, the score was slightly higher in the HMD condition ($M = 2.40$, $SD = 2.15$) than the Non-HMD condition ($M = 1.88$, $SD = 1.69$), suggesting that the immersive VR environment was perceived to be more realistic than the screen environment.

Participants' most preferred method of learning English conversation was practicing with native speaker partners followed by VR before the study. This preference did not change after the study, but one of the participants changed her mind after the VR experience. Learning English conversation in VR was not an option for her before the study, but she was willing to learn with VR afterward.

In the open-ended response about the system, participants indicated that the VR character was unrealistic and/or distracting at times. They also indicated that the training contents need improvement as well as the technical aspects of the program. One of the participants commented that "The program seemed to focus on memorizing the conversation rather than practicing spontaneous English conversations".

5. Discussion

The goal of our research is to develop a VR conversation training program in English. We tested the prototype in a pilot study. The results are inconclusive because of the small sample size, but it appears that learning English conversation in the HMD condition and the non-HMD condition did not produce any differences in all the measures we examined. Participants are used to learning from computer screens. One might expect that participants' performance would suffer because of the unfamiliarity of the immersive VR devices. The fact that participants in the HMD condition learned about the same in spite of the unfamiliarity suggests that the immersive VR may outperform the non-HMD condition with further design improvements.

The results of the study suggest a few areas for further improvement in the design and use of the environment. First, immersive VR requires headset, which means that participants with glasses would suffer from low vision. Upgrading the graphics card is an option, but we may also make the display bigger or ask participants to wear lenses during the training. Second, the contents of the training programs need to be expanded with more diverse scenarios and instructional activities. At the moment the training relies on simple practice model, but more diverse instructional activities and activities sequences are needed to make the conversational training effective. With these and other improvements, we hope and expect that immersive VR can play a bigger role in second language learning in the future.

Acknowledgements

This research was supported by a Basic Science Research Program through the National Research Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF 2016R1D1A1B03935697 and NRF 2017R1D1A1B03035576).

References

Armstrong, N. & E2K (2012). Shiny stones. Darakwon: Seoul.

Chun, D., Kern, R., & Smith, B. (2016). Technology in language use, language teaching, and language learning. *The Modern Language Journal, 100*(S1), 64-80.

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: a review of technology types and their effectiveness. *Computer Assisted Language Learning, 27*(1), 70-105.

Hong, M. P., Nam, H. M. V. V., Nam, T. H., Lee, C. W., & Cha, M. K. (2015). High school English conversation. Visang: Seoul.

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

Kim, Y. C. & Gong, J. S. (2008). A study on the effect of using rubric through students` participation on speaking ability. *Studies in English Education, 13*(1), 66-98.

Lee, J. K. (2007). Try again! English conversation that starts again with middle school textbooks. Gilbut: Seoul.

Maclntyre. P. & Gardner. R. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44*(2), 283-305.
Redstone, C., Cunningham, G. & Lee, H. S. (2010). Everyday English 365. Logos books: Seoul.
Witmer. B. G., Jerome. C. J, & Singer. M. J. (2005, June). The factor structure of the presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 14(3), 298-312.