


## Listening tasks in virtual reality: A pilot study

Regina Kaplan-Rakowski<sup>a</sup> and Yongluan Ye<sup>b</sup>

<sup>a</sup>Department of Learning Technologies, University of North Texas, , [Regina.Kaplan-Rakowski@unt.edu](mailto:Regina.Kaplan-Rakowski@unt.edu) and <sup>b</sup>Department of Learning Technologies, University of North Texas, [YongluanYe@my.unt.edu](mailto:YongluanYe@my.unt.edu)

How to cite: Kaplan-Rakowski, R.; Ye, Y. (2023). Listening tasks in virtual reality: A pilot study. In *CALL for all Languages - EUROCALL 2023 Short Papers*. 15-18 August 2023, University of Iceland, Reykjavik. <https://doi.org/10.4995/EuroCALL2023.2023.16960>

---

### Abstract

*Listening comprehension is a crucial skill in Foreign Language (FL) learning, yet many FL learners encounter challenges in this area, leading to frustration and discouragement. To address the issue, this pilot study investigated whether high-immersion Virtual Reality (VR) could effectively facilitate attentive listening tasks for FL learners. This quasi-experimental, within-subject pilot study involved ten learners of Chinese from an underserved, urban high school in the United States. Participants experienced an animated story in VR with and without captions, then completed pre- and post-tests to assess language gains and listening comprehension. A questionnaire measured participants' perceptions of VR and the use of captions. Results indicated that VR positively impacted language gains and listening comprehension, showing improvement with both groups of participants. Although the group without captions slightly outperformed the group with captions, this difference was not statistically significant. Novice learners reported challenges in comprehending the VR story, highlighting their need for appropriate scaffolding. Overall, the findings suggest that VR can be an effective tool for practicing listening comprehension for FL learners, but careful consideration of content complexity and scaffolding is crucial for learners at different proficiency levels. This study lays the groundwork for further research on VR-assisted language learning.*

**Keywords:** *high-immersion Virtual Reality (VR), Virtual Reality-Assisted Language Learning (VRALL), listening comprehension, captions, language learning.*

---

## 1. Introduction

The development of listening skills is essential in Foreign Language (FL) learning, but many FL learners find listening comprehension challenging (Feyten, 1991; Graham, 2006). To succeed in listening comprehension tasks, FL learners need to identify individual words while forming mental representations. When failing to do so, learners often neglect the rest of the message, thus aggravating the accumulation of difficulties (Goh, 2000) and making the learners feel discouraged.

Given that learners find high-immersion Virtual Reality (VR) enjoyable and motivating (Kaplan-Rakowski & Gruber, 2022, 2023), the idea of providing listening comprehension tasks in VR to encourage learners to listen attentively is compelling although understudied (Dhimolea et al., 2022). Studies on listening skills attempted to use VR to prompt information retention (Lee, 2019; Pinto, 2019; Tai & Chen, 2021). Tai and Chen (2021) found that learners who experienced the interaction in VR retained significantly more knowledge than those who watched two-dimensional videos of the same interactions. Similar to Tai and Chen (2021), Pinto (2019) reported that learners had a higher sense of presence and engagement in VR, but VR and two-dimension (2D) video approach

made no significant difference in learners' knowledge retention. Meanwhile, the small sample size ( $N = 12$ ) and the short language input (~2 minutes) limited the generalizability of Pinto's (2019) findings.

Although previous studies on the effect of VR on listening comprehension yielded promising results (Lee, 2019; Pinto, 2019; Tai & Chen, 2021), they mainly focused on understanding dialogues instead of monologues. Monologues are essential for FL learners to function in the target language environment. However, monologues provide fewer cues that can aid comprehension of the materials (Fox Tree, 1999). Only Lee (2019) examined the effect of VR on monologues without incorporating the interactive functions of VR, limiting the impact of the technology.

The use of captions has shown significantly positive results in helping FL learners improve their listening comprehension (Ghasemholand & Nafissi, 2012; Hsu et al., 2013), but the impact of captions in VR is still uncertain. While VR provides more visual details to the learners, Mayer and Pilegard (2014) argued that too much information delivered through the visual channel may cause essential overload and impede learning outcomes. Adding captions to the rich visual information in VR may increase the cognitive load and hinder FL learners' listening comprehension.

### **Study Purpose and Research Questions**

To fill the gaps in the literature, this pilot study aimed to prepare for a more in-depth project exploring the effect of VR on listening comprehension and the effect of captions in VR (see Ye & Kaplan-Rakowski, 2023). The study seeks to answer the following questions: (1) Does VR promote language gains and listening comprehension? (2) How do captions affect listening comprehension in VR? and (3) Will captions cause essential overload or support listening in VR?

## **2. Method**

This pilot study was conducted in preparation for a larger project investigating the impact of VR on the development of listening comprehension skills (Ye & Kaplan-Rakowski, 2023). To answer the research questions, we designed a quasi-experimental, within-subject study. Parametric data were analyzed using independent  $t$ -tests, and nonparametric data were analyzed using descriptive statistics.

### **2.1 Participants**

The study participants ( $N = 10$ ) came from low-income families residing in an urban area of the United States. Eleven participants consented to participate, but only ten (male = 6, female = 4) completed all parts of the study. All participants were 18 years old but varied greatly in their years of Chinese language education. The majority of participants had been learning Chinese for four to six years, with the least experienced participant having studied Chinese for one year and the most experienced having studied it for over ten years. This variety was to help gain data on the learning experience in VR from learners of different language proficiency.

### **2.2. Procedure**

The procedure used in the study had three stages. First, the participants received scaffolding (see Figure 1) on the story used in the experiment including explanations of the six high-frequency vocabulary items from the story and pictures showing the setting and the main characters. Afterward, participants took a pre-test to establish their language knowledge baseline.



Figure 1. Teacher providing scaffolding with screenshots from *The Line*. Photo by the authors.

Second, the participants experienced the target story in VR (see Figure 2) in two parts. This was an animated, romantic story which was embedded within the app called *The Line*. We chose this particular app because the story it offers has a repetitive structure that may help comprehension. Moreover, the interaction is simple and unlikely to cause motion sickness. The video version of this story can be viewed [here](#). The participants were randomly assigned to the caption or non-caption conditions to experience the first part of the story and then switch to the other condition to experience the second part of the story. See Figure 3 for a screenshot of the story.



Figure 2. Participant experiencing the virtual reality story *The Line*. Photo by the authors.



Figure 3. A screenshot of the interactive virtual reality story with captions

Third, after each intervention, the participants took a post-test to determine their language gains. Finally, the participants completed a questionnaire that collected demographic data and data that reflected the participants' experience in VR.

### 2.3. Instruments

The present study used four instruments: a pre-test, two post-tests, and a questionnaire. The pre-test and post-tests contained 14 multiple-choice and three translation questions, in which participants needed to listen to audio recordings to determine the meanings of words, phrases, and sentences. The post-tests also had three inferential questions regarding the story that could not be answered by merely watching the VR story. The participants' answers were scored and the accuracy rate was used in the analysis. The pre- and post-tests assessed only participants' listening comprehension, meaning that participants did not need to read Chinese in the tests.

The questionnaire collected participants' demographic data such as gender, age, and years of learning Chinese. To reflect the participants' experience using VR, the questionnaire used questions with a seven-point Likert scale. The questionnaire adapted four questions from the instrument developed by Leppink et al. (2014) to measure the cognitive load, three questions from the instrument by Ellis et al. (2021) to measure the positive effect of VR, and four questions from the instrument by Gandolfi et al. (2021) to measure presence.

### 3. Results and Discussion

Each test question underwent validation using Cronbach Alpha, which is a method to measure the reliability of questionnaires. It assesses whether scores from multiple items can reliably capture the target information. All pre-test questions were valid, but the post-test achieved only 78.6% validity because of three open-ended inferential questions. During the follow-up interview, several participants reported that those questions were too difficult because they asked for details that were mentioned in the narration only once. Therefore, before conducting the larger and main study (Ye & Kaplan-Rakowski, 2023), inferential questions were changed to multiple choice questions asking for opinions and feelings of the characters from the story.

The free-response portion of the post-tests was scored by two researchers following the same rubric: each correctly translated key item earned one point. Missed items or wrong translations earned zero points. When disputes occurred, the two researchers discussed the issue at hand until an agreement was achieved. The accuracy rate of each participant was analyzed using a *T*-test. The participants who experienced the story with captions on average

increased by 6.1 points on the listening comprehension test ( $SD = 9.96$ ), and those who experienced the VR story without captions increased by 8.1 points ( $SD = 12.16$ ). The non-caption group insignificantly outperformed the caption group [ $t(10) = 0.581, p = 0.078$ ]. In addition, two participants' scores were lower in one or both post-tests, which suggested that these learners guessed responses in the pre-test although the researchers urged them to choose the "I don't know" option when needed.

Nonparametric data were analyzed and the descriptive statistics (means, standard deviations, theoretical minimum, and maximum range of scores for each item) are presented in Table 1. Two participants at the novice level in Chinese reported greater difficulties comprehending the test questions and the VR story than the others. The 11<sup>th</sup> participant, who did not finish the study, decided to leave because the test was too intimidating. Therefore, the researchers decided to exclude novice Chinese learners from the main study.

**Table 1.** Descriptive statistics of the study constructs

N = 10	M	SD	Min	Max
Listening comprehension-VR with captions	6.10	9.96	-13.00	21.00
Listening comprehension-VR without captions	8.10	12.16	-16.00	26.00
Caption usefulness	5.28	1.21	4.00	6.25
VR positive effect	6.25	0.20	6.00	6.50
Cognitive load	4.33	0.57	3.33	4.67
Sense of presence	5.63	0.92	4.00	7.00

*Note.* Score ranges were 0 to 100 for listening comprehension tests, 1 to 7 for caption usefulness, 1 to 7 for positive affect, 1 to 7 for cognitive load, and 1 to 7 for sense of presence. Some theoretical minimal values are below zero. This is because some students did worse on the post-test than on the pre-test.

The results indicate that VR can support listening comprehension amongst Chinese FL learners, which answers our first research question. This learning outcome may be explained by the participants' enjoyment in using VR as a learning media and that VR created a reality in which participants felt present. To answer the second question, although a majority of the participants considered captions helpful in aiding comprehension, the learning outcomes of the caption group were insignificantly lower than of the non-caption group. This result may have been caused by some participants' low Chinese proficiency level and their limited ability to process the captions at the pace of the VR story. This study did not measure reading comprehension that may be enhanced with captions.

This pilot study was based on a small sample size. Future studies should reexamine the claims with more participants. Moreover, because of the cost of VR devices, VR may need to outperform other lower-cost media (e.g. computers, tables) in supporting listening comprehension for the technology to be deemed desirable to learners and educators. Therefore, future research should examine the impact of alternative media on listening comprehension.

#### 4. Conclusions

This study explored the effect of VR on listening comprehension and validated the methodology used in the main study (Ye & Kaplan-Rakowski, 2023). Overall, participants reacted positively toward learning language with VR, and the learning outcome suggested that the VR experience supported their learning. Learning outcomes with different media will be examined in the main study.

While all participants reported a positive effect toward VR, those at the novice level in Chinese found the tests too difficult. One participant complained about the large quantity of unfamiliar vocabulary and the scaffolding, which, although helpful, was insufficient. Given previous research on learning vocabulary in VR (Papin & Kaplan-



Rakowski, 2022), vocabulary and grammar may need to be limited in input materials and assessments for novice language learners.

## Acknowledgements

We extend our thanks to all the study participants for their willingness to participate, share insights, and dedicate their time to advance our research.

## References

- Christmann, A., & Van Aelst, S. (2006). Robust estimation of Cronbach's alpha. *Journal of Multivariate Analysis*, 97(7), 1660–1674.
- Dhimolea, T. K., Kaplan-Rakowski, R., & Lin, L. (2022). A systematic review of virtual reality language learning. *TechTrends*, 66, 810–824. <https://doi.org/10.1007/s11528-022-00717-w>
- Ellis, M. E., Downey, J. P., Chen, A. N., & Lu, H. (2021). Why Taiwanese seniors use technology. *Asia Pacific Management Review*, 26(3), 149–159. <https://doi.org/10.1016/j.apmr.2021.01.001>
- Feyten, C. M. (1991). The power of listening ability: An overlooked dimension in language acquisition. *The Modern Language Journal*, 75(2), 173–180. <https://doi.org/10.2307/328825>
- Fox Tree, J. E. (1999). Listening in on monologues and dialogues. *Discourse Processes*, 27(1), 35–53.
- Gandolfi, E., Kosko, K. W., & Ferdig, R. E. (2021). Situating presence within extended reality for teacher training: Validation of the extended Reality Presence Scale (XRPS) in preservice teacher use of immersive 360 video. *British Journal of Educational Technology*, 52(2), 824–841.
- Ghasemboland, F., & Nafissi, Z. (2012). The effects of using English captions on Iranian EFL students' listening comprehension. *Procedia, Social, and Behavioral Sciences*, 64, 105–112. <https://doi.org/10.1016/j.sbspro.2012.11.013>
- Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, 28(1), 55–75. [https://doi.org/10.1016/S0346-251X\(99\)00060-3](https://doi.org/10.1016/S0346-251X(99)00060-3)
- Graham, S. (2006). Listening comprehension: The learners' perspective. *System*, 34(2), 165–182.
- Hsu, C., Hwang, G., Chang, Y., & Chang, C. (2013). Effects of video caption modes on English listening comprehension and vocabulary acquisition using handheld devices. *Educational Technology & Society*, 16(1), 403–414.
- Kaplan-Rakowski, R., & Gruber, A. (2023). An experimental study on reading in high-immersion virtual reality. *British Journal of Educational Technology*, 00, 1–19. <https://doi.org/10.1111/bjet.13392>
- Kaplan-Rakowski, R., & Gruber, A. (2022). Motivation and reading in high-immersion virtual reality. In B. Arnbjörnsdóttir, B. Bédi, L. Bradley, K. Friðriksdóttir, H. Garðarsdóttir, S. Thouëсны, & M. J. Whelpton (Eds), *Intelligent CALL, granular systems, and learner data: Short papers from EUROCALL 2022* (pp. 208–213). Research-publishing.net. <https://doi.org/10.14705/rpnet.2022.61.1460>
- Lee, A. (2019). Using virtual reality to test academic listening proficiency. *Korean Journal of English Language and Linguistics*, 19(4), 688–712.
- Leppink, J., Paas, F., Van Gog, T., van Der Vleuten, C. P., & Van Merriënboer, J. J. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, 30, 32–42.
- Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 316–344). Cambridge University Press.

- Papin, K., & Kaplan-Rakowski, R. (2022). A study on vocabulary learning using immersive 360° pictures. *Computer Assisted Language Learning*, 35. <https://doi.org/10.1080/09588221.2022.2068613>
- Pinto, D., Peixoto, B., Krassmann, A., Melo, M., Cabral, L., & Bessa, M. (2019, April). Virtual reality in education: Learning a foreign language. In Á. Rocha, H. Adeli, L. Reis, & S. Costanzo (Eds.), *World Conference on Information Systems and Technologies* (pp. 589–597). Springer.
- Tai, T.-Y., & Chen, H. H.-J. (2021). The impact of immersive virtual reality on EFL learners' listening comprehension. *Journal of Educational Computing Research*, 59(7), 1272–1293. <https://doi.org/10.1177/0735633121994291>
- Ye, Y., & Kaplan-Rakowski, R. (2023). Practicing listening comprehension skills in high-immersion virtual reality. *SSRN*, <https://ssrn.com/abstract=4335690>