Subtitles in a Virtual Environment

ZACH KEELER, Colorado State University, USA
TAYLOR HOLCOMB, Colorado State University, USA
JACKSON MORRIS, Colorado State University, USA
ELIJAH WEETMAN, Colorado State University, USA

This paper investigates the challenges of implementing subtitles in Virtual Reality (VR) environments, focusing on user experience, immersion, motion sickness, and readability. With a growing interest in VR content, accessibility options become crucial for diverse user groups, including hearing-impaired, non-native speakers, and those who prefer subtitles for various reasons. An experiment was conducted using the Meta Quest 2 headset, and participants were exposed to three different subtitle captioning methods in a VR environment. The methods included fixed subtitles, character/comic subtitles, and Head Mounted Display (HMD) subtitles. Participants' comprehension, readability, and preferences were assessed through quizzes and surveys. The study found that fixed subtitles were the preferred method, scoring highest in both quizzes and preference data. However, further research is needed to understand the factors contributing to this preference and explore additional subtitle presentation methods in VR environments.

ACM Reference Format:
Zach Keeler, Taylor Holcomb, Jackson Morris, and Elijah Weetman. 2023. Subtitles in a Virtual Environment. 1, 1 (May 2023), 8 pages. https://doi.org/1234567.1234567

1 INTRODUCTION

Within VR content, there is not a lot of experience integrating subtitles into the content. Including VR options within a virtual environment can be a challenge, and traditional subtitle methods that apply to 2D videos do not necessarily apply to 3D environments. In addition to hearing-impaired people, there are non-native speakers who require translated text and people that prefer to use subtitles for accessibility reasons. For developers of VR content, it's important to provide subtitles that are not distracting from the immersion of the VR environment,[22] do not cause motion sickness,[11] and are easily read and understood.[13]

Our goal with this study is to provide feedback on the user experience using these metrics.

- Within VR, subtitles should not detract from the immersion or understanding of the scene.
- Motion sickness is a major concern within VR. Subtitles should not contribute to motion sickness.
- Subtitles should always be easily visible. Obscured subtitles reduce the ability of the user to understand what is happening on screen.

Authors' addresses: Zach Keeler, keeler@colostate.edu, Colorado State University, Fort Collins, Colorado, USA, 80537; Taylor Holcomb, tjholcom@colostate.edu, Colorado State University, Fort Collins, Colorado, USA, 80537; Jackson Morris, weetmel@colostate.edu, Colorado State University, Fort Collins, Colorado, USA, 80537; Elijah Weetman, , Colorado State University, Fort Collins, Colorado, USA, 80537.

Manuscript submitted to ACM

2 RELATED WORKS

Existing research into subtitles in VR is divided between 360-degree video and VR environments. VR environments differ from 360-degree video by being more interactive since the player can also move around an environment. Still, subtitle developments in 360-degree video also influence VR environments since the two can share subtitle methods.

2.1 360 Video

To ensure optimal readability of text in VR, developers must consider several factors, such as font type, size, color, and contrast. They must also consider the placement of the text, ensuring it is in a location that is easily visible to users without causing discomfort or requiring unnatural head movements. This applies to both 360-degree videos and VR Environments. [14] [16]

Subtitles fixed to the field of view were found to be preferable in 360-degree videos in several studies. [21] [10] [14] This style of subtitle is most similar to existing 2D video and was preferred among viewers who are used to existing subtitle methods.

2.2 VR Environment

In VR, users are often required to move their heads and eyes to view different parts of the environment. This can cause issues with text readability if the font size, contrast, or placement of the text is not appropriate. This is an issue not faced by 360-degree videos and is unique to VR Environments. For example, if the text is too small or placed too far away from the user, it may be difficult to read or even impossible to see. [18]

The results of the study showed that users had a preference for the fixed subtitles and subtitles attached to the user's viewpoint, but not for the subtitles fixed in the environment. Users who experienced higher levels of VR sickness tended to prefer fixed subtitles. In addition, users who preferred the attached subtitles reported higher levels of presence in the VR environment. [11]

Within a VR Environment, there was a general preference for fixed subtitles. However, users that used captions as a secondary source of information found a preference for subtitles attached to the environment. [20] It is recommended to have multiple options and settings for subtitles to suit individual users. [19] One proposed solution to the difficulty in depth tracking is better hardware that allows for eye tracking. [23] More advanced HMDs, like the are available that can track a user's gaze and can allow for subtitles that follow where the user is looking. These HMDs are just starting to be released, with more coming in the future. [24]

2.3 VR Sickness

In order to avoid VR-induced sickness interfering with results, developers have to consider hardware, content, and other factors that impact humans. The hardware must have powerful enough specifications to run the testing software as intended, without frame drops or stutters contributing to discomfort. Developers should also consider if their VR content is simulating an activity that is normally conducive to sickness in real life, such as roller coasters or a maritime situation. Developers should also screen for potential human factors including previous VR experience, nausea susceptibility, and age. These factors have all been linked to the severity and likelihood of VR sickness. [15]

Given a sufficiently provocative stimulus, nearly all people can be made motion sick. [17] Many risk factors in the experiment are impossible to avoid, which is why a questionnaire should be administered about motion sickness. Many people have never experienced VR before, but they might be aware of preexisting susceptibility to motion sickness.

It's also possible to avoid situations within VR that make people motion sick. For example, the most popular way to move around in VR is with a pointing and teleporting movement system. [12] This is a popular method because it fixes the user in place and the refresh rate of the HMD is less of an issue than it would be during movement.

3 METHODOLOGY

In this section, we outline our experiment and research design and give in-depth information on the materials, participants, subtitle presentation method, and execution.

3.1 Research Design

The experiment aimed to gather information on whether a preferred subtitle captioning method in a VR environment is related to the overall readability and comprehension of that method. Additionally, the experiment contributed to the creation of subtitle guidelines for accessibility by identifying which methods of subtitles were found easiest to use. To gather the preference for a subtitles method, readability, and comprehension, we posed this research question: Did preference for a subtitles captioning method in a VR environment relate to readability and comprehension? From this, we aimed to prove one of the following hypotheses. H0: The preference for a subtitle captioning mode had no relation to a higher level of readability or comprehension. H1: The preference for a subtitle captioning mode resulted in a higher level of readability and comprehension.

We also wanted to ensure the viability of our subtitles against the guidelines that are used for other applications. To do this we followed a number of guidelines on subtitles to ensure the words per minute and characters per minute allow enough time for the user to properly read each line [4] [1] [7]. Additionally, since our survey used the medium of VR the application of subtitles being used in video games we also took into consideration gaming guidelines. We used these guidelines to direct how our subtitles contrasted with the environment, we are also able to consider breaking a guideline such as not having information displayed by text alone. By ignoring this guideline in our experiment it increases the external validity of this application beyond video games [5]. The Oculus Developer Hub also provides some important accessibility guidelines specifically for the Quest and similar VR headsets [8]. By using a combination of information about the state of guidelines on subtitles we were able to provide a result that would follow the criteria to be shown to a large audience. Multiple factors like size, position, color, length, and timing were all considered. This also ensured the integrity of our independent variables as each scene fitting the guidelines is required.

3.2 Materials and Participants

Testing was conducted using the Colorado State University computer science equipment as well as personal equipment owned by the researchers. The VR headset used is the Meta Quest 2. [6] Additionally, student-owned laptops were used for surveys and questionnaires. We had ten participants take part in our study, nine men and one woman. The participants were given consent and VR sickness screening forms to sign and then had the experiment explained to them. The participants then viewed the scenes and filled out the corresponding questionnaire in the order they were instructed to. The order of the scenes was changed between participants to counterbalance the survey. The questions were designed to test the readability and ease of use of the subtitles, so the questions were fairly simple. At the end of the experiment, the participants were also given a preference survey about the different subtitle methods.

3.3 Subtitle Presentation Method

There are three methods of subtitle captioning that we used in the experiment.

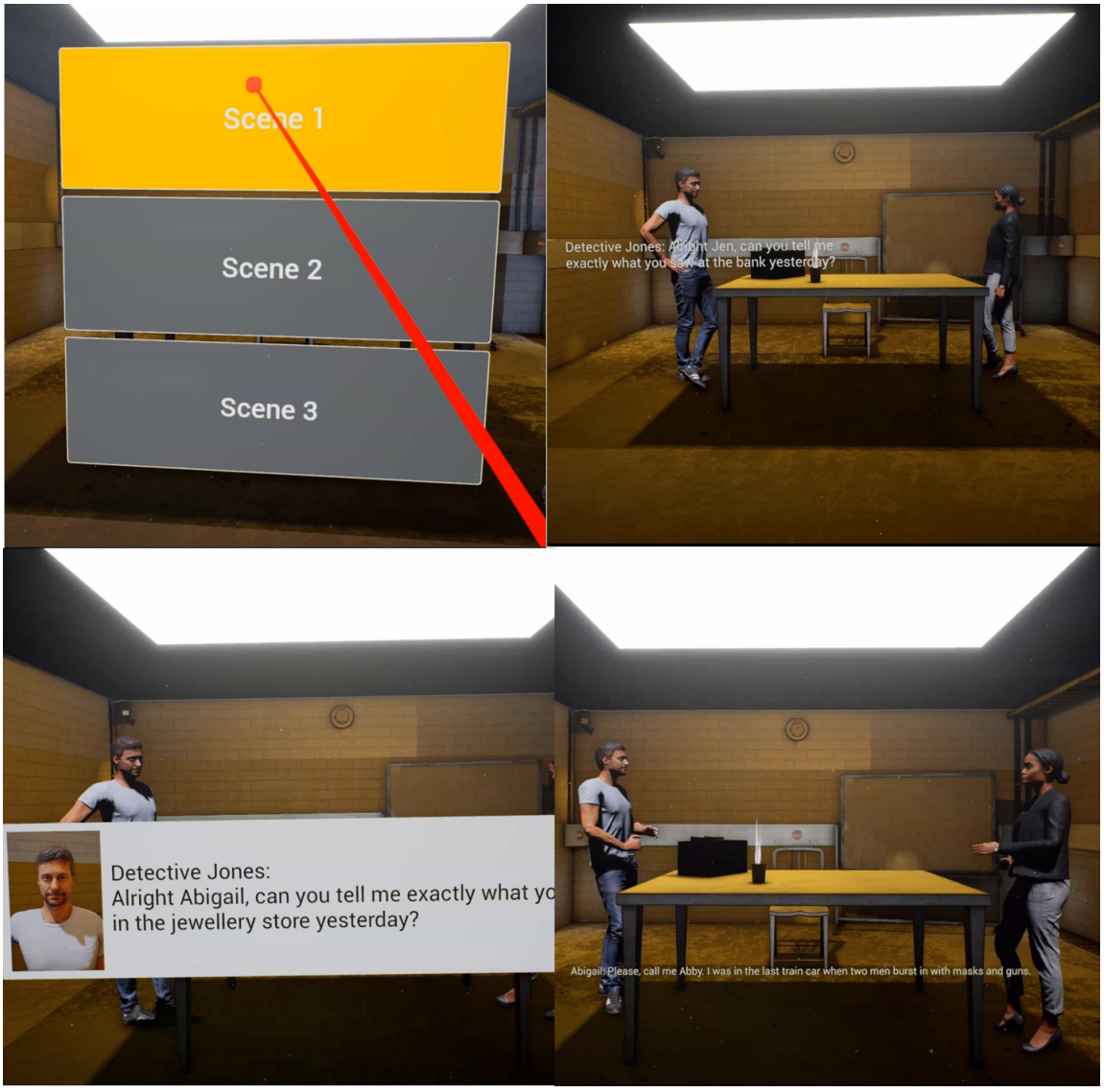

Fig. 1. Visual examples of the four scenes. From top left to bottom right, Scene Select, Fixed, Character/Comic, HMD

3.3.1 Fixed. With this method, subtitles are placed in front of the characters. These subtitles appear as floating text within the environment. They are placed near the person speaking so that they can be viewed in addition to the scene. As the viewer looks around the virtual environment, the subtitles are fixed in their position in the world. If the participant looks away from the characters and scene, the captioning in place will also go out of view.

3.3.2 Character/Comic. With this method, subtitles are fixed to the character speaking, similar to the Fixed subtitles in the previous section. However, this subtitle method differs in a few ways. The subtitles are placed onto a 2-dimensional

floating plane that acts as a backdrop to the text for better readability. It also places a picture of the character speaking on the plane next to the text. The subtitles and image of the speaking character provide extra information to discern who is speaking.

3.3.3 HMD. The last subtitle method follows the headset display around as the viewer's head moves. The subtitles are fixed to the viewpoint of the viewer and are independent of the location of the speaker. The subtitles are visible at all times. This method is closest to existing subtitle methods that are employed in 360-degree videos and also subtitles available in 2D media like movies and TV shows.

3.4 Environment

We created our VR environment using Unreal Engine 5 [9]. The engine provided the basic framework to create our VR environment and port it onto the VR headset. The viewer is placed into the same interrogation environment with three unique scripts. Each script is designed to follow similar story points with small details changed to the story to challenge the user on the overall readability of that captioning mode and their comprehension. We imported an interrogation room asset [2] and two high resolution character models [3] to the environment to allow more immersion into the scene and allow the user to focus more on the captions and overall scene. The environment does not provide any audio to the user. This was done to keep the participant focused on the content within the subtitles using only the subtitle method. The environment also contains a scene selection menu that allows us to easily change the order in which scenes are shown with a button selection from the Quest 2 controller. Controllers are not used in any other environment except for scene selection.

3.5 Execution

Each participant was introduced to the experiment with an explanation of the survey and what they are expected to contribute. They are then administered their consent form and VR Sickness questionnaire. Before allowing them to put on the headset, we assure the participant that if they feel uncomfortable they can take off the headset or stop at any time. The participant then selects the predetermined scene and watches each scene with their respective subtitle captioning method. Each participant had a different order to view the scenes in order to avoid any unwanted learning and allow counterbalancing of our results. Between each scene, they take a small questionnaire about details in the scene to test their understanding of the subtitles in the scene. They also filled out the preference survey for just that scene before starting the next one. We also allowed a break for participants at this time. After every scene and questionnaire was completed, a final preference survey was administered. This survey asked five questions about which subtitle method was preferred on a 1 to 5 scale. At the end of the survey, participants were debriefed.

4 DATA AND RESULTS

In the upcoming section, we will discuss both the results from the scene questionnaires and preference survey questions.

4.1 Questionnaire Results

The results of the scene quizzes were an easy way to check how effective the methods were at relaying information. Since we ran participants through different permutations of the scenes, we eliminated most of the learning that would have skewed results. The average correct answers from the three methods are as follows (out of 8): Scene 1: 6.9, Scene 2: 6.0, Scene 3: 5.9

These quiz results show us that method 1's subtitle style was favored for comprehension and readability, while methods 2 and 3 were about the same and a little less effective. However, a further distinction between methods 2 and 3 will occur when we discuss the preference surveys.

4.2 Preference Survey Results

The preference surveys were helpful as a secondary metric to judge a method's efficacy. For example, a subtitle method that filled the screen with text would definitely be readable and help to answer the quiz questions, but would not be an enjoyable user experience. For this reason, we put a lot of weight on these survey results.

Descriptives

Descriptives

	Scene 1	Scene 2	Scene 3
N	5	5	5
Missing	0	0	0
Mean	4.58	4.20	2.32
Median	4.60	4.10	2.40
Standard deviation	0.192	0.141	0.444
Minimum	4.30	4.10	1.60
Maximum	4.80	4.40	2.80

Fig. 2. Results of Survey

The results show that even though method 1 was clearly the most effective choice for the quiz questions, it was closer in preference to method 2 than the quiz results would show. Additionally, the preference survey showed that although Method 3's quiz scores were similar to Method 2, people did not like it nearly as much. In fact, the question "Which interface did you find the most frustrating to use?" was answered unanimously with method 3. A box at the bottom had people state why, and the majority of complaints were either that the text was too small or that the text moving with the head was detrimental to ease of use. Overall, the preference and feedback surveys confirmed what the quiz scores showed us. The results shed some light on important specifics that had to be taken into consideration for the final result decision.

4.3 Final Results Analysis

The results of the surveys and scene quizzes both offer a decisive leader in our hunt for the optimal subtitle method out of the methods we had tested. Scoring highest on both quizzes and in preference data was method 1: fixed subtitles. Fixed subtitles are the closest to how 2D subtitles function, so it makes sense that this sense of familiarity would cause better performance and also receive glowing commendations. Second place in this experiment was method 2: character/comic subtitles. We believe that the size and clarity allowed participants to read and absorb information

quickly, but the large text box and unorthodox style may have led them to be slightly overwhelmed when trying to absorb the information for the quizzes. That being said, people responded well to seeing the character portraits and colorful text, which placed it close behind method 1 in the surveys although it may have lagged in quiz performance. In clear last place was method 3: HMD. These subtitles were last in both quiz and preference. While the quiz scores were a close third, the preference survey proved that this method was the least favored. A combination of small text, awkward position of text, and the fact that it moved with the user's head all contributed to a low preference score. In conclusion, considering the results of the preference surveys and questionnaires, we can confirm the [H1] hypothesis that a preference for a subtitle captioning method will result in a higher level of readability and comprehension.

5 CONCLUSION

In conclusion, this study aimed to show the most effective and preferred subtitle method within a VR environment. Participants were shown Fixed, Character/Comic, and HMD subtitle methods. The results from the scene quizzes and preference surveys revealed that fixed subtitles were the most effective and preferred method, followed by Character/Comic subtitles, and lastly HMD subtitles. It is our belief that this study's findings can provide insight into what types of subtitles are preferred in VR content.

REFERENCES

[1] 2018. Web content accessibility guidelines (WCAG) 2.1. https://www.w3.org/TR/WCAG21/

[2] 2019. The interrogation in environments - UE marketplace. https://www.unrealengine.com/marketplace/en-US/product/the-interrogation

[3] 2019. Scanned 3D people pack in characters - UE marketplace. https://www.unrealengine.com/marketplace/en-US/product/9c3fab270dfe468a9a920da0c10fa2ad

[4] 2022. BBC subtitle guidelines. https://www.bbc.co.uk/accessibility/forproducts/guides/subtitles/

[5] 2022. Game Accessibility Guidelines. https://gameaccessibilityguidelines.com/full-list/

[6] 2022. Get started developing for the Oculus Quest Platform. https://developer.oculus.com/quest/

[7] 2022. Timed text style guide: Subtitle timing guidelines. https://partnerhelp.netflixstudios.com/hc/en-us/articles/360051554394-Timed-Text-Style-Guide-Subtitle-Timing-Guidelines

[8] 2022. VR accessibility design: Captions and subtitles | oculus developers. https://developer.oculus.com/resources/design-accessible-vr-captions/

[9] 2023. Unreal Engine. https://www.unrealengine.com/en-US

[10] Belén Agulló, Mario Montagud, and Isaac Fraile. 2019. Making interaction with virtual reality accessible: rendering and guiding methods for subtitles. *AI EDAM* 33, 4 (2019), 416–428. https://doi.org/10.1017/S0890060419000362

[11] Lindsey Bouwels. 2022. *Accessible Subtitles in 6DoF 3D Virtual Reality: How Preference of Subtitle Presentation Modes relates to Presence and VR Sickness.* Ph. D. Dissertation. https://www.researchgate.net/publication/362703439_Accessible_Subtitles_in_6DoF_3D_Virtual_Reality_How_Preference_of_Subtitle_Presentation_Modes_relates_to_Presence_and_VR_Sickness

[12] Evren Bozgeyikli, Andrew Raij, Srinivas Katkoori, and Rajiv Dubey. 2016. Point amp; Teleport Locomotion Technique for Virtual Reality: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in play. https://dl.acm.org/doi/abs/10.1145/2967934.2968105

[13] Marta Brescia-Zapata, Krzysztof Krejtz, Pilar Orero, Andrew Duchowski, and Chris Hughes. 2022. VR 360º subtitles: Designing a test suite with eye-tracking technology. https://doi.org/10.47476/jat.v5i2.2022.184

[14] Andy Brown, Jayson Turner, Jake Patterson, Anastasia Schmitz, Michael Armstrong, and Maxine Glancy. 2018. Exploring Subtitle Behaviour for 360° Video. https://www.bbc.co.uk/rd/publications/whitepaper330

[15] Eunhee Chang, Hyun Taek Kim, and Byounghyun Yoo. 2020. Virtual reality sickness: A review of causes and Measurements. *International Journal of Human–Computer Interaction* 36, 17 (2020), 1658–1682. https://doi.org/10.1080/10447318.2020.1778351

[16] Mario Montagud Climent, Olga Soler-Vilageliu, Isaac Fraile Vila, and Sergi Fernández Langa. 2021. VR360 Subtitling: Requirements, Technology and User Experience. *IEEE Access* 9 (2021), 2819–2838. https://doi.org/10.1109/ACCESS.2020.3047377

[17] John F Golding. 2006. Motion sickness susceptibility. https://www.sciencedirect.com/science/article/pii/S1566070206002128

[18] Jacek Jankowski, Krystian Samp, Izabela Irzynska, Marek Jozwowicz, and Stefan Decker. 2010. Integrating Text with Video and 3D Graphics: The Effects of Text Drawing Styles on Text Readability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 1321–1330. https://doi.org/10.1145/1753326.1753524

[19] Mario Montagud, Cristian Hurtado, Juan Antonio De Rus, and Sergi Fernández. 2021. Subtitling 3D VR Content with Limited 6DoF: Presentation Modes and Guiding Methods. *Applied Sciences* 11, 16 (2021). https://doi.org/10.3390/app11167472

[20] Pranav Pidathala, Dawson Franz, James Waller, Raja Kushalnagar, and Christian Vogler. 2022. Live Captions in Virtual Reality (VR). arXiv:2210.15072 [cs.HC] https://arxiv.org/abs/2210.15072

[21] Sylvia Rothe, Kim Tran, and Heinrich Hussmann. 2018. Positioning of Subtitles in Cinematic Virtual Reality. In *ICAT-EGVE 2018 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, Gerd Bruder, Shunsuke Yoshimoto, and Sue Cobb (Eds.). The Eurographics Association. https://doi.org/10.2312/egve.20181307

[22] Yusuke Shimizu, Ayumi Ohnishi, Tsutomu Terada, and Masahiko Tsukamoto. 2021. Gaze-Adaptive Subtitles Considering the Balance among Vertical/Horizontal and Depth of Eye Movement. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 127–132. https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00035

[23] Ludwig Sidenmark, Nicolas Kiefer, and Hans Gellersen. 2019. Subtitles in Interactive Virtual Reality: Using Gaze to Address Depth Conflicts. https://eprints.lancs.ac.uk/id/eprint/132411/

[24] Lili Wang, Xuehuai Shi, and Yi Liu. 2023. Foveated rendering: A state-of-the-art survey - computational visual media. https://link.springer.com/article/10.1007/s41095-022-0306-4