Enhancing the Spectator Experience

Integrating Subtitle Display in eXtended Reality Theatres

submitted in partial fulfillment for the degree of master of science

Atanas Yonkov
15170616

Master Information Studies
information systems
Faculty of Science
University of Amsterdam

Submitted on 29.06.2024

	UvA Supervisor	UvA Examiner
Title, Name	Prof. dr. Frank Nack, PhD	Prof. dr. Hamed Seiied Alavi, PhD
Affiliation	UvA, IVI	UvA, IVI
Email	F.M.nack@uva.nl	h.alavi@uva.nl

Enhancing the Spectator Experience: Integrating Subtitle Display in eXtended Reality Theatres

Atanas Yonkov (SN: 15170616)

ABSTRACT

The rapid growth of virtual and augmented reality technologies, encapsulated by the term eXtended Reality (XR), has revolutionized the interaction with digital content, bringing new opportunities for entertainment and communication. Subtitles and closed captions are crucial in improving language learning, vocabulary acquisition, and accessibility, like understanding audiovisual content. However, little is known about integrating subtitle displays in extended reality theatre environments and their experience influence on the user. This study addresses this gap by examining subtitle placement and design attributes specific to XR settings. Building on previous research on subtitle placement, mainly in television and 360-degree videos, this project focuses on the differences between static and dynamic subtitle variants. The study uses a comprehensive literature review, Virtual Reality (VR) theatre experiment, and analytics to investigate these aspects of subtitle integration in the specific case of a VR theatrical Greek play with subtitles. The results show that the comparison between the two variants is insignificant, and both implementations produce high scores. However, thematic analysis suggests the preference for static over the dynamic variant depends heavily on the specific context and the number of speakers in the scene. Since this study focuses on a monologue theatrical play, the next step in future work would be to explore a "multi-speaker" play.

KEYWORDS

extended reality, virtual reality theatre, subtitle display, user experience, human-computer interaction

1 INTRODUCTION

The rapid growth of virtual reality (VR) and augmented reality (AR) technologies have revolutionized how we interact with virtual digital content, offering new frontiers in entertainment and communication [45]. VR is a technology that immerses users in a computer-generated environment, where images and sounds respond dynamically to their movements. AR and eXtended Reality (XR) are part of the Reality-Virtuality Continuum [5], which spans the range of experiences from the completely physical world to a fully digital environment. AR modifies the real environment by overlaying computer-generated graphics onto it. XR, an umbrella term, encompasses AR, VR, Mixed Reality (MR), and other related technologies, representing all current and future possibilities that blend physical and digital worlds [24].

The evolution in media accessibility has greatly enhanced features like subtitles[20]. Subtitles have evolved throughout history, with studies showing their positive impact and sometimes even necessity on viewer comprehension [25]. Research indicates that bilingual subtitles, combining intralingual and interlingual elements, provide linguistic support, making videos easier to comprehend and aiding learning [11]. Additionally, broadcasts with subtitled original versions have been linked to improved English proficiency worldwide, especially in listening comprehension skills [41]. Subtitles play an important role in making video content accessible [36] to diverse populations of viewers, such as non-native language speakers, the hard of hearing, and people with learning disabilities. They enable such viewers to understand the video content.

Similar to subtitles in two-dimensional spaces, subtitles in XR provide a means for understanding audiovisual content in a three-dimensional space. One such example is attending a theatrical play in a foreign language [4], [39]. However, there is a lack of guidelines on implementing subtitle or live captioning in XR and how it differs from traditional television captioning [14]. Studies have been conducted to understand user preferences for different VR captioning styles [13]. Generally, these studies tested different caption movement behaviours, such as head locked, lag, and appear captions, while participants watched live-captioned presentations in VR [13], [38].

Some areas that still require further exploration and understanding include the **position of subtitles**, speaker identification, and their **overall influence on an XR environment** [40].

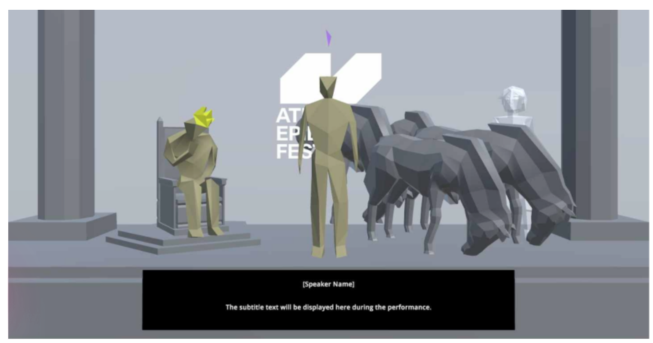

Figure 1: Static subtitles in a theatrical play scene from a participant's VR headset perspective

This thesis project aims to bridge the gap in the existing literature by exploring optimal subtitle position integration, particularly static and dynamic in the context of VR theatre environment, where **static** (Fig. 1) refers to subtitles fixed to user gaze and **dynamic** (Fig. 2) refers to subtitles fixed to objects (actors in our case) in the theatrical VR environment. That said, the problem has been narrowed down to the following research question (RQ):

To what extent can subtitles be integrated into a virtual reality theatre environment to improve the user experience for the spectator while not distracting from the main content?

Which was further split into the following sub-research questions (SRQ):

(1) *SRQ1: How do static and dynamic subtitle placements affect user engagement, and how distractive are they?*

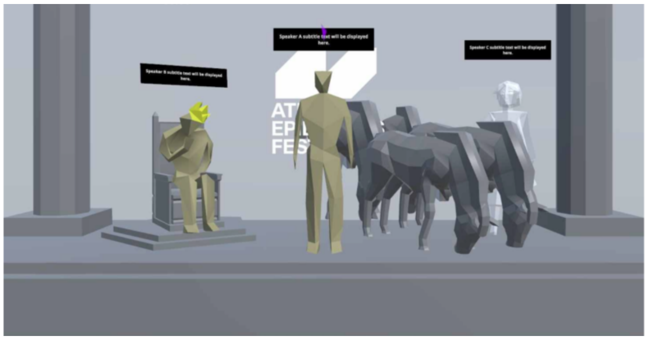

Figure 2: Dynamic subtitles in a theatrical play scene from a participant's VR headset perspective

(2) *SRQ2: How do best practices for subtitle style (colour, size, font) translate to XR environments to ensure readability and maintain immersion when VFX are present?*

(3) *SRQ3: What observations can be made about cybersickness and subtitle display?*

The thesis structure comprises a review of relevant literature (Section 2). Followed by a description of the prototype developed for the study (Section 3). An explanation of the experiment design (Section 4). An outline of the results (Section 5), a discussion (Section 6) and lastly, a conclusion (Section 7).

2 RELATED WORK

The following section outlines the development of subtitles, beginning with television, progressing through 360-degree videos, and finally observing current advancements in subtitles display in VR and possible validation measures.

2.1 Subtitles in Two-Dimensional Spaces

Subtitles in two-dimensional spaces are textual representations of spoken dialogue and audio cues that appear on flat screens, such as those on televisions, phones, tablets, and computers. These subtitles are overlaid on the video content to provide viewers with a written version of the audio, aiding in comprehension, particularly for those who are deaf or hard of hearing or who speak a different language from the audio track [30].

Brown et al. [16] discussed that the conventional practice is that subtitles are typically positioned at the centre bottom of the screen [40], a format that often results in considerable spatial gaps between text and pertinent visual elements, potentially straining viewers' eyes and causing them to miss critical visual content [33]. They compared two types of TV show subtitles: the conventional centre-bottom subtitles and dynamic subtitles that change location based on the subject of interest in the scene. They used eye-tracking and other validation measures and concluded that dynamic subtitles are less disruptive to the viewing experience than traditional ones [16].

A contemporary continuation known as speaker-following subtitles departs from this norm by placing subtitle text within speech bubbles proximate to the current speaker. To evaluate the efficacy of this approach, Kurzhals et al. [33] conducted a controlled eye-tracking laboratory investigation involving 40 participants. The study juxtaposed the traditional method of centre-bottom subtitles with implementing content-sensitive, speaker-following subtitles across various dialogue-intensive video clips. The study shows that subtitles that follow the speaker increase focus on relevant image regions and reduce eye strain by shortening saccade length.

2.2 Subtitles Position in XR Environments

The advent of VR brings the question of how this new modality of perception can be used when it comes to subtitle display in 360-degree environments and VR. The concept of dimension differs in virtual reality (VR) compared to 360-degree technology. In VR, users can experience a computer-generated environment that simulates real-world scenarios and interactions in an immersive, three-dimensional world [7]. VR usually uses head-mounted displays (HMDs), tracking systems, and input devices to create this immersive experience. On the other hand, 360-degree technology, such as 360-degree videos, allows users to view real-world footage filmed from a 360-degree camera, providing an immersive experience where users can pan around and explore the environment without the same level of depth and interaction as VR.

Adapting subtitles from two-dimensional to three-dimensional space is further explored in a study by Rothe et al. [40]. Their approach consists of exploring how dynamic subtitles are perceived compared to static in the world of XR. Nonetheless, adopting dynamic subtitles, which adjust their positioning by the underlying video content [16] and combining the possibilities of XR, presents an opportunity to enhance the coherence and immersive quality of the viewing experience. Their study compared the two main subtitle variants. Although a work-in-progress, the results indicate a preference for dynamic due to higher presence scores, less sickness, and lower workload.

Maintaining immersion, an important aspect of 360-degree experiences, necessitates minimizing disruptions caused by subtitles. Additionally, subtitles must be situated intuitively for viewers, ensuring they don't require excessive conscious effort to locate [40]. Given the prevalence of virtual reality sickness among viewers, the behaviour of subtitles should not exacerbate this issue, further complicating their design and placement. The study [40] focuses on presenting subtitles with HMDs. Three common ways of displaying dynamic subtitles are presented: evenly spaced, following the head immediately, following the head delayed [6]. Brown et al. compared the mentioned variants and found that the general preference towards following the head immediately is better than other behaviours regarding ease of locating subtitles. However, the study is limited to focusing only on subtitle behaviour and did not extensively explore other design aspects that could impact user experiences, like font size or colour.

Brescia-Zapata et al.[13] explore creative subtitle production in immersive environments, particularly focusing on VR technologies. The study introduces an experimental setup using eye-tracking to test subtitle placement in VR. The research discusses challenges in generating subtitles for 360-degree VR videos, including a subtitle editor and a system for eye movement data collection. The study's pilot phase tests three subtitling modes, combining psychophysiological and qualitative analyses. The collected eye movement data results indicate that the focus group supports the preference for subtitles fixed relative to the speaker or overhead-locked subtitles. In the terminology of this study, these are called dynamic subtitles [40].

Rzayev et al. [43] discusses the positioning and mode of textual information display in AR cases. The study investigated the impact of text position and reading mode on comprehension and workload while walking and sitting. Results indicated that displaying text in the bottom-centre position during walking yielded the highest comprehension, whereas the top-right position resulted in the lowest comprehension and highest perceived workload. Conversely, while sitting, comprehension was slightly higher when the text was centred compared to the bottom-centre position, with the top-right position yielding the lowest comprehension again. Participants reported that reading from the centre or bottom-centre position resembled reading from a computer screen or book. In contrast, continuous reading from the top-right position led to eye strain. Limitations included controlled room conditions and the use of specific smart glasses, suggesting the need for further research to explore real-world scenarios and alternative devices.

Displaying translations in different positions on AR glasses could also impact comprehension and task load [42]. The study found that presenting translations in an overlay format resulted in the lowest comprehension and highest task load, as participants had difficulty viewing both foreign words and translations simultaneously. Participants preferred positions where the foreign word and translation were displayed close to each other, such as below the screen. However, statistical results did not reveal any significant differences between placements. Although the task completion time did not differ significantly across conditions, qualitative feedback and comprehension scores favoured positions where real-world and virtual information could be viewed simultaneously.

2.3 Factors Influencing XR Experience

In immersive virtual environments (IVE), user experience research often uses various measuring aspects [34] for evaluating product and service quality. This subsection outlines those factors observed in XR experiments and the validation methods utilised to get insightful results in XR studies.

We can explore several established tools and frameworks to thoroughly discuss the factors influencing subtitle experience and their validation methods. These tools comprehensively evaluate different aspects of user interaction, ensuring a holistic understanding of the subtitle experience in immersive environments.

One such tool is the System Usability Scale (SUS) [15]. The SUS is widely used to assess the perceived usability of systems, products, or services. Consisting of ten questions evaluated on a five-point Likert scale, the SUS captures a user's holistic perception of the system's effectiveness and user satisfaction by considering various usability aspects such as complexity, integration, and confidence. Its standardized format facilitates easy comparison across different systems and enables benchmarking against industry standards, making it particularly useful in VE studies.

Complementing usability assessments, the Igroup Presence Questionnaire (IPQ) measures the sense of presence or the psychological state of "being there" [37] in Immersive Virtual Environments IVEs, especially VR. Presence is important for evaluating VR applications' effectiveness, as higher levels of presence can enhance the overall experience and lead to better outcomes, such as skill transfer in professional simulators [9]. The IPQ assesses subscales like spatial Presence, involvement, and experienced Realism, providing qualitative insights into the quality of the VR experience [37], [9].

The NASA Task Load Index (NASA-TLX) becomes valuable for workload assessment when tasks involve varying cognitive and physical demands. Developed by NASA, the NASA-TLX [23] evaluates six dimensions of task load: mental demand, physical demand, temporal demand, performance, effort, and frustration. Each dimension is rated on a twenty-one-point scale, allowing users to express the intensity of workload experienced across different facets of the task. This multidimensional approach offers insights into the components contributing to the overall task load, aiding in identifying areas for optimization and improvement, particularly in IVEs [45].

In XR technologies, assessing user discomfort and adverse symptoms is essential. The Simulator Sickness Questionnaire (SSQ) [31], originally developed to test sickness symptoms in fighter jets, has been adapted for evaluating VR-induced cybersickness [28]. The SSQ covers sixteen symptoms across four categories: nausea, oculomotor disturbance, disorientation, and total score. Users rate each symptom on a four-point scale, quantitatively measuring the adverse effects experienced during VR interactions. Administered before and after the VR experience [10], the SSQ focuses on physiological responses and perceptual discrepancies induced by IVEs [17].

Lastly, a UX framework for subtitles involves analyzing the impact of subtitles on viewer engagement and satisfaction [18]. Research indicates that approximately 10% of TV audiences use subtitles, underscoring the importance of understanding subtitle users' experiences. Studies have demonstrated that subtitle positioning within videos can significantly affect user experience. For example, dynamically moving subtitles can improve viewers' engagement and personalization options [19].

2.4 Eye Tracking

Within subtitle studies, whether in two-dimensional spaces [16], [8], [18] or IVEs [13], [14], the topic of eye tracking has been included to enhance the validity of the results. It helps to show whether XR's visual perceptions and cognitive processes are as the statistical results indicate [26]. Also, it allows for interactive and controlled analysis of user behaviour [44].

Different modern devices, including the Tobii Pro, Vive Pro Eye, and Vive Focus 3, support eye tracking [26] and have been used during experiments [16], [46]. A comparative study by Hou et al. [26] indicates that while all of the tested devices yield high enough values for eye tracking, the Meta Quest Pro has a spatial accuracy and precision comparable to the other modern headsets [47]. Given the parameters intended for evaluation in this user study, the Meta Quest Pro proves to be a suitable and capable device.

3 PROTOTYPE

While VR and AR offer distinct experiences, they hold a potential for engaging theatrical experiences [2]. Based on these insights, the following section outlines a VR theatre environment prototype focusing on subtitle placement in two distinct variants. This contextualised the pilot experimental study of a Greek theatre play with English subtitles in a VR theatre environment.

3.1 Subtitles for an AR theatre

This project was conducted in scientific collaboration with the DIS (Distributed & Interactive Systems) Group at Centrum Wiskunde en Informatica (CWI), the Dutch National Research Institute for Mathematics and Informatics. One of the research interests of DIS is on XR applications and related user experiences. This project is in the research context of integrating subtitles into an AR theatre in real life. This pilot study focuses only on two representations of subtitles—static and dynamic [40]. Ideally, the final use case is a real-life AR theatre in Athens, Greece. However, scientifically informed decisions are required to enhance the general public's experiences.

3.2 Environment Recreation

Before commencing this study, for validity and accuracy repetition, the scene, the actors, and the play have been recreated in a VR theatre application developed with Unity v2022.3.22f. The theatrical play already chosen was a four-minute and thirty-second introduction to the Hippolytus tragedy by Euripides [21]. There is a project with actual stakeholders (a real theatre company in Greece) and their decision in terms of the performing arts, artistic value and development of the app have taken place before the start of this thesis. Decisions about the duration of the play and the presented objects have been made based on the script and instructions provided by them.

On top of that, for this specific project, Oculus' Integration SDK v57.0.1[1] has been utilised to integrate the eye-tracking functionality of the Meta Quest Pro [3]. A collider has been added to key environmental objects (encapsulated by the term Virtual Effects (VFX) - horses, bull, rocks, the Messenger and the King) in the scene. The VFX were directly tied to the monologue script, appearing only when specific triggering terminology was spoken. The subtitles remained present throughout the entire duration of the play, providing a consistent visual reference for the audience, whereas the objects only appeared in specific keyframes based on their relevance to the narrative.

In the context of using the Meta Quest Pro for collecting eye-tracking data during a theatrical play, the process involves utilizing the headset's embedded cameras to capture the position of the user's gaze. These cameras continuously track eye movements in real-time. As the user watches the performance, the eye-tracking data is processed to project a virtual ray corresponding to the direction of the gaze. When this ray intersects with predefined objects of interest in the scene, which have been tagged with an *ObjectCollider* property, the system monitors the duration of this intersection. If the user's gaze remains fixed on the object for more than *2 seconds*, the system records a data entry into a CSV file. This entry includes the precise position of the eyes at the moment of collision, the specific object that the gaze ray intersected with, and the exact position of the object where the collision occurred. The objects were present to maintain constant visual stimuli and mitigate the monotony of reading subtitles. This approach aimed to determine whether participants consistently read subtitles even when other attention-grabbing elements were present [47]. Furthermore, the Meta Quest Pro was chosen due to its capability to display the theatre without a wire, unlike the HTC Vive, allowing for less disruption for participants [26].

[1] https://assetstore.unity.com/packages/tools/integration/oculus-integration-deprecated-82022

3.2.1 Subtitle Variants. The two subtitle variants tested are labelled with the aliases **static** and **dynamic** [40] (Figures 1, 2). Subtitles were only available in English and the audio-visual low-poly representation of the chosen excerpt of the Hippolytus Greek play was played in modern Greek language.

The settings pane (See Appendix B) is an interactive part of the theatre performance. Participants can adjust font, background opacity (of the subtitles boxes), vertical position and reading distance (only on the static subtitles) due to decisions made beforehand during the application development. In the eye-tracking data, this game object is called "Controls".

3.2.2 Error Profiling. Additionally, a custom error profiling scene has been created for this study. Wei et al.'s study inspires this scene [47] regarding the accuracy of the Meta Quest Pro [3]. Each participant was asked to look at the centre of nine cubes for 3 seconds (See Appendix B). The cubes were permanently attached to the headset's camera to follow the gaze. That was done to ensure even small head movements would not account for precise eye tracking. The spatial position of collisions from the eye rays was recorded so that camera data from the main theatre recording could later be adjusted. The cubes are spread to the same distance as the default of the static subtitles pane. That way, it could be computed later if there is any deviation or inaccuracy with the eye-tracking rays projected from the eyes.

3.3 Apparatus

A Windows 10 computer with an Intel i7-8700K processor, NVIDIA GeForce GTX 1080 graphics card, and 32 GB of RAM was used to run a VR experience. Using a Quest Link wireless connection, the output was displayed on a Meta Quest Pro [3] device. Eye-tracking data was captured using the Quest Pro's built-in eye-tracking feature, and the headset was configured to operate at a 90Hz refresh rate.

4 EXPERIMENTAL SETUP

The following section outlines the experimental setup to investigate participants' responses to static and dynamic subtitle displays within a VR theatre IVE.

4.1 Methodology

This research combines a **within-subject study with mixed-method analysis**. The approach consists of a user study divided into two main parts: an experimental part and subsequent data analysis. The experiment combines qualitative measures: semi-structured interviews, quantitative point-scale questionnaires, and

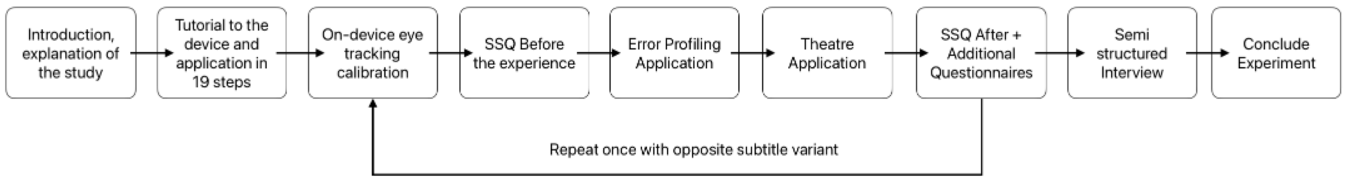

Figure 3: Study Experiment Timeline

eye-tracking vector data. A conducted power analysis with the statistical tool G*Power[2] (See Appendix D) calculated that the number of participants needed was 15 to obtain .8 power to detect a large effect size of 0.8 at the standard 0.05 alpha error probability.

4.1.1 Design of Quantitative Method. This study administered seven questionnaires to participants. Due to physiological constraints, the SSQ was the only questionnaire given to participants immediately before and after each VR session. The remaining questionnaires were administered in the following order (see Appendix E). Based on expert advice, the order of the questionnaires presented is selected for optimal experience. Consequently, the SSQ results were analyzed independently by comparing the sessions, not subtitle variants, as shown in Table 2. The cumulative results of the other questionnaires are presented in Table 1. Based on established literature, these additional questionnaires assessed various VR-related factors, such as usability, presence, and realism. It's important to note that specific formulas (see Appendices E.1, E.2, E.4 respectively) were applied to analyse the SSQ [31], NASA-TLX [23], and SUS [15]. In contrast, the General, Behavioral [48], and UX for Subtitles framework [6] were analyzed based on their average values grouped by factors they correspond to based on existing literature. The IPQ [37] and Immersion questionnaires [27], in particular, were selectively grouped according to the approach outlined in the paper by Lee et al. [34]. All formulas are listed in Appendix G.

4.1.2 Design of Qualitative Method. The interview questions were inspired by Brown's whitepaper regarding dynamic subtitles in 360-degree environments[6] and can be found in Appendix F. However, not all of them fit the IVE theme of our study, which is why a small part was chosen - 9 questions.

4.2 Participants

19 participants from the general population participated in the user study. Preferably, participants were theatre-goers, but the only strict criterion was to be a non-native Greek speaker. Recruitment aimed to ensure diversity in age, gender, and XR familiarity to remove potential sampling bias. Invitations were sent to CWI and university colleagues via email, and the organising platform used was Doodle[3].

4.3 Procedure

The study was conducted from 13.05.24 to 22.05.24 at the DIS Immersive Media Lab, CWI. The average session lasted 60 minutes, with 50 minutes being the shortest and 120 being the longest for one elderly participant, depending on the participants' familiarity and adaptability with VR headsets and controllers.

The experiment timeline (Fig. 3) began with an introductory session, during which participants were thoroughly briefed on the study's objectives, procedures, and potential risks, ensuring full comprehension before providing informed consent. All questionnaires were presented to participants on paper. Subsequently, participants were presented with a consent form delineating data processing protocols, permissions, and sharing agreements exclusively for scientific purposes, allowing them to seek clarification before affixing their signatures. Following this, participants completed a demographic questionnaire, furnishing details such as name, age, gender, and prior experience with XR technologies. Before commencing any VR experience, they were asked to note down an SSQ to establish a baseline for cybersickness.

Following this, participants were guided on navigating the VR environment to cultivate familiarity and mitigate potential discomfort during subsequent steps in the study. They were recommended to look around freely for a few minutes until they got accustomed to the VR environment and controllers; the duration varied between participants. Each session began with calibrating the eye-tracking feature from the Meta device, which participants had to do independently with thorough guidance from the experiment organisers.

Following that, before starting the theatre VR app, participants were asked to perform an error profiling scene. Participants experienced either static or dynamic subtitles in VR Theater environment in counter-measured order (even-numbered participants always began with the static variant first while odd-numbered with dynamic; this is done to account for the learning curve during the first session) and answered questionnaires and rated their experience after each session (See Appendix E).

Post-experiment, the participants completed the cybersickness questionnaire again to gauge any discomfort or fatigue experienced during their stay in VR. After the first experiment session, the adapted questionnaires were administered: SUS, IPQ, UX, Immersion, General, and Behavioural. Assessing participant perceptions of subtitle placement, readability, and overall user experience.

To reduce cybersickness symptoms, participants were provided with an intermission period between the two VR sessions. The duration of this intermission varied among participants based on age, overall health, and individual physiological conditions, ranging from none to fifteen minutes.

The second round/session mirrored the preceding one, with participants undergoing static or dynamic subtitle experiments in the counter-measured order. Following an optional short break (0-15 minutes), participants engaged in a semi-structured interview (See

[2] G*Power Statistical Tool: https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower
[3] Online meeting scheduling tool: https://doodle.com/

Section 4.1.2). This facilitated exploring their experiences, preferences, and challenges related to the subtitle's display, relation to VFX and general remarks. The interview was audio-recorded. Upon conclusion, participants received debriefing information and were granted a €10 VVV Cadeukaart[4].

5 RESULTS

To understand the user perception and preference towards static and dynamic subtitles, their usability and overall experience of the VR Theater, we compared two paired groups in questionnaires (see Appendix E) with ordinal data where normality is not assumed. Questionnaire data is analysed with Matlab_R2024a. Based on the nature of the data, the Kolmogorov-Smirnov test [1] is utilised to see if data is normally distributed within a single group separately. After that, static and dynamic subtitle results are compared to find a significant difference with a Wilcoxon Signed Rank test [1]. These statistical tests are computed using Matlab (See Appendix H) for each of the factors (see Table 1, Table 2). Eye tracking data is profiled, cleaned and analysed - Python 3 has been utilised. Qualitative data from semi-structured interviews has been transcribed and analysed in Dovetail[5].

This study's experiment included 19 participants (7 female, 12 male, mean age: 35 [min 22, max 71]). On a scale from one to five (best), the mean self-rated experience with immersive technologies was 2.8 (SD=1.22, median=3). The mean self-rated attendance in performances where string values were converted to numbers 1 (less than once a year) to 4 (more than 3 times a year) was 2.29 (SD=1.19, median=2). Of the 19 participants, eye tracking data had to be excluded for P01 (changes in development environment later on), P03, P07, P11, and P13 (either too high of a diopter or limited-eye mobility, which resulted in unsuccessful calibration on the device). 19 interview audio recordings have been transcribed and annotated for thematic analysis in Dovetail (see Section 5.2).

5.1 Quantitative

Questionnaire results are accumulated and ordered by the factor(s) each questionnaire represents (see Table 1). Questions and factors they relate to can be found in *Appendix E*. Bolded are the means which are better for the given subtitle variant. All statistical analyses yielded *p-values* of less than *0.001*, indicating strong statistical evidence to reject the null hypothesis. This suggests that the observed effect or difference is highly significant and unlikely to be due to random variation. Therefore, individual *p-values* will not be reported in the subsequent sections, only in Table 1. ***SSQ*** results are presented in table 2 as they have been analysed in another manner *(explained in section 5.1.8)*.

5.1.1 Eye Tracking Data. The interactions of each participant with various objects (Fig. 10) in the VR scene were processed to aggregate the eye-tracking data. The duration and frequency of fixations on specific objects were recorded. As mentioned in Section 3.2, VFX objects are present to break the monotony of the environment. This YouTube video shows a visual demonstration of the scene. The CSV file, containing object eye-tracking data, was read and analysed

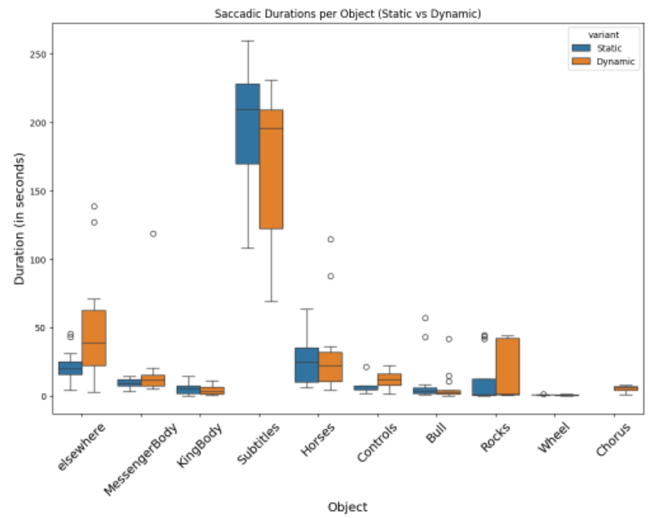

Figure 4: Saccadic Duration per Object

using a Python3 Notebook to accumulate each object's total saccadic duration[6] and the number of gaze fixations (collisions).

Figure 4 highlights the saccadic duration of attention per participant for different objects. The overall duration of attention for subtitles is the highest, with the Dynamic variant showing a slightly lower median than the Static variant but a wider range, indicating variability in individual engagement. Participants also showed substantial attention to horses, with the dynamic variant again showing a higher median and a broader range, suggesting increased and varied engagement with dynamically presented subtitles. On the other hand, engagement with the Messenger Body and King Body remains low, with minimal differences between variants, suggesting consistent but limited attention across participants. Additionally, Controls, Bull, Rocks, Wheel, and the Chorus show low engagement across participants, with the Dynamic variant occasionally showing slight increases, but overall, the values remain lower.

Figure. 5 shows how long participants fixated on different game objects. Interestingly, subtitles and horses had the longest gaze fixation duration in both play versions. The fixation duration on subtitles was longer than on any other object. The Dynamic version had a wider range of fixation duration and a higher median than the Static version. **As noted earlier, the VFX objects appeared in response to specific triggering terminology within the monologue scripts, while subtitles were consistently displayed throughout the play.** On the other hand, the Messenger Body and King Body had a relatively shorter fixation duration, with the Messenger Body seeing a slightly higher fixation duration in the Dynamic version. Meanwhile, Controls, Bull, Rocks, Wheel, and Chorus had relatively low fixation duration, with minimal differences between the two versions.

Below sections 5.1.2 to 5.1.7 refer to Table 1.

[4]Virtual gift card coupon platform: https://www.vvvcadeaukaarten.nl/

[5]Dovetail is a qualitative data analysis web application: https://dovetail.com/

[6]Saccadic duration is the time it takes for the eye to complete a saccade, typically ranging from 20 to 200 milliseconds, reflecting the rapid shift in focus from one point to another.

Table 1: Accumulated Results of questionnaires ordered by factor

Questionnaire	Factor	Static - average (mean), standart deviation (*SD*) and distribution (*p*)				Dynamic - average (mean), standart deviation (*SD*) and distribution (*p*)				Mean both variants, *SD*, significance (p) and effect size (r)			
		mean	*SD*	*M*	*p*	*mean*	*SD*	*M*	*p*	*mean*	*SD*	*p*	*r*
General (Appendix E.7)	Interface	**3.89**	0.875	4	≤ .001	3.84	0.834	4	≤ .001	3.87	0.035	1	0
	Experience Rating	**4.05**	0.848	4	≤ .001	3.84	0.834	4	≤ .001	3.95	0.148	0.157	0.229
Behavioural (Appendix E.7)	Behavioural Intention	4.65	0.613	5	≤ .001	**4.74**	0.516	5	≤ .001	4.70	0.064	0.236	0.192
UX for Subtitles (Appendix E.3)	UX Rating	**5.14**	1.232	5	≤ .001	4.97	1.177	5	≤ .001	5.06	0.120	0.717	0.059
SUS (Appendix E.4)	Usability	83.42	10.145	85	≤ .001	**83.82**	8.223	85	≤ .001	83.62	0.283	0.795	0.042
IPQ and Immersion (App. E.5 and E.6)	Immersion and Presence	**4.21**	0.875	4	≤ .001	4.1	0.681	4	≤ .001	4.16	0.078	0.584	0.089
	Realism	**3.45**	1.153	4	≤ .001	2.99	1.165	3	≤ .001	3.22	0.325	**0.033**	0.347
NASA-TLX (Appendix E.2)	Mental	**27.90**	24.513	20	≤ .001	30.789	26.261	25	≤ .001	29.34	2.046	0.7462	0.05
	Physical	**15.00**	17.873	5	≤ .001	16.579	19.152	5	≤ .001	15.79	1.117	0.6308	0.08
	Temporal	30.26	29.034	20	≤ .001	**29.474**	27.482	25	≤ .001	29.87	0.558	0.9641	0.01
	Performance	**20.26**	19.894	10	≤ .001	26.842	26.782	20	≤ .001	23.55	4.652	0.5663	0.09
	Effort	26.58	20.953	15	≤ .001	**22.105**	17.742	20	≤ .001	24.34	3.164	0.5644	0.09
	Frustration	**12.90**	16.693	5	≤ .001	17.105	19.742	5	≤ .001	15.00	2.977	0.4473	0.12

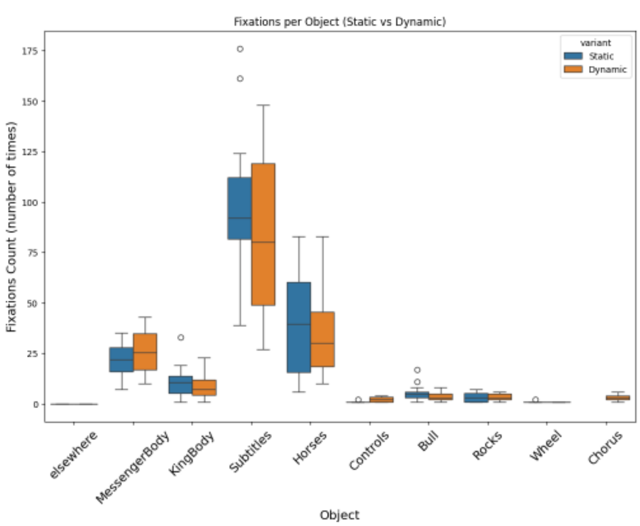

Figure 5: Gaze Fixations per Object

5.1.2 General. The general questionnaire observes 2 **self-reliant** factors - interface and experience rating on a scale of 5 with values ranging from Bad to Excellent. This test observes the participants' self-rated likeness and general rating towards the subtitle interface and experience accordingly. The medians (M) for both factors are 4, with no statistically significant difference between subtitle variants.

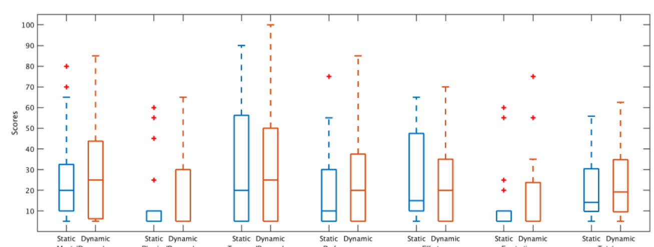

Figure 6: Comparison of static and dynamic subtitle variants for NASA-TLX Factors (lower scores are better)

5.1.3 Behavioural. The Behavioural Intention factor observes 3 questions on a scale of 5. Higher values are better. The results yield M-values of 5 for both variants, with the dynamic variant yielding a slightly higher mean score. No statistical significance is found when comparing both variants.

5.1.4 UX Framework for Subtitles. When analyzing the UX Framework for Subtitles using a Likert scale of 7, both variants yield

Table 2: Accumulated results of SSQ by session (See Appendix E.1)

	ABAA (AB - AA)	AABB (AA - BB)	BBBA (BB - BA)
TS	163.83 (140.96 - 186.71) p=0.147, r=0.224	162.52 (186.71 - 138.34) **p=0.007**, r=0.413	132.99 (138.34 - 127.63) p=0.326, r=0.152
w_N	10.90 (9.99 - 11.81) p=0.377, r=0.136	10.22 (11.81 - 8.63) **p=0.020**, r=0.360	8.40 (8.63 - 8.18) p=0.705, r=0.058
w_O	17.33 (14.44 - 20.21) p=0.120, r=0.240	17.33 (20.21 - 14.44) **p=0.032**, r=0.331	13.90 (14.44 - 13.36) p=0.465, r=0.113
w_D	15.58 (13.26 - 17.90) p=0.100, r=0.254	15.91 (17.90 - 13.92) **p=0.034**, r=0.327	13.26 (13.92 - 12.59) p=0.680, r=0.064

M-values = 5. The mean value for static subtitles is slightly higher, at 5.14. No statistical difference is found when comparing the variants.

5.1.5 SUS. The results from the SUS scaled between 0 and 100, showed medians of 85 for both variants, with a slightly higher mean score for the static variant. No statistical significance was found between them. According to Lewis et al., both interfaces can be considered equally usable with a SUS score of "A" for usability [35].

5.1.6 IPQ and Immersion. The Immersion and Presence factor (scale of 5) showed medians of 4 for both subtitle variants, with a slightly higher mean score for the static subtitles = 4.21. No statistical significance was found when comparing the variants.

However, the Realism (scale of 5) indicated a statistical significance benefitting the static variant (M = 4), whereas the dynamic yielded an M-value of 3.

5.1.7 NASA-TLX. The NASA-TLX test (see Fig. 6) showed that the Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, Frustration [23] factors were not significantly different between the static and dynamic subtitle variants. The scores for both variants were generally low, showing a low demand in all categories. None of the median scores (ranges 5-25) surpassed 30 points on a scale of 100, with a maximum score at the 75-percentile of 56.25 for the Temporal Demand factor, further supporting the observed low demand.

5.1.8 SSQ. The SSQ is administered before and after the VR experience at each of the two sessions (Session A and Session B) - one observing static subtitles and the other - dynamic; thus, a total of four SSQ scores are in Table 2 (AB, AA, BB and BA). It measures simulator sickness in three areas: Nausea (w_N), Oculomotor (w_O), and Disorientation (w_D), along with a Total Score (TS). The only significant difference was observed in the comparison between the SSQs administered at session A After (AA) and session B Before (BB), the Total Score decreased significantly from 186.71 to 138.34, with a notable reduction in simulator sickness symptoms. It is notable that this reduction was observed following an up to 15-minute break between sessions.

5.2 Qualitative

The interviews conducted for this study were analyzed using the thematic analysis approach as outlined by Braun and Clarke [12]. Individual participants are labeled P1-P19. The number of participants who agreed with the given statement is indicated in parentheses.

5.2.1 General Opinion. Overall, opinion was good (15) - (*P12: "...The theatre experience was quite interesting. I felt like I was in the film or the scene. It was extremely good..."*) Regarding the **interface** participants (13) stated they liked and enjoyed the subtitle display interface: (*P13: "...So I found the interface easy to use, and I found the subtitles interface also easy to use and clear..."*), (*"P3: "...The interface, like the settings, was really intuitive and very straightforward. You see it, and you see exactly what's happening directly..."*). However, some participants (6) did not show the same acceptance: (*P12: "...So the subtitle interface, I did not like neither of them nor the 2D nor the 3D..."*). Some participants (15) were **aware of the environment**, expecting a low-poly version of the environment: (*P9: "...I felt immersive even in the in the prototype environment..."*), while others (4) were not aware and experienced a negative moment: (*P4: "...it's not consistent with the real world..."*).

5.2.2 Static vs Dynamic Subtitle Placements. **For this particular play**, 8 participants preferred dynamic (*P9: "...I prefer the dynamic subtitles because I can enjoy more the theatre so I can see the whole scene..."*), while 10 - static (*P8: "...because with the static, it was moving along with where I was looking, whereas in the dynamic, I had to look at the character speaking..."*). If it **was a longer play**, sometimes participants (7) changed their opinion to the more conventional [16] type subtitles (*P9: "...in this case, I would prefer static because it follows my head movement and after a long time, I will start to, to move my head and change my position ... I think the static position, it will be better. Especially if I use this on my home to lay down, I can see it..."*), while others (11) kept their opinion (*P2: "...still dynamic, I would still prefer it, especially if there were multiple actors and speakers..."*).

5.2.3 Effect of VFX on subtitles. **The VFX**, most participants (10) **found useful** (*P16: "...I thought they were great. I thought they made the experience more immersive. It was quite stunning to see the big figures..."*). Nevertheless, they were **considered as a necessary part** of the experience (7) (*P13: "...I think the subtitles should be secondary to those ... subtitles should be a support..."*) and **sometimes even too minimal** (*P2: "...they were like a picture, they were not very detailed..."*) and **even a little distracting** (4) (*P19: "...sometimes the scale was very different. So like there's a big bull that comes. So I have to be like "oh", and then I'm distracted and then, and I missed what he (Messenger) said..."*).

5.2.4 Suggestion Improvements. Regarding **technical errors**, participants (15) reported issues affecting their experience: (*P17: "...when the water rose it covered the subtitles. So, I was like, oh, I have to look up now..."*), (*P1: "...the subtitles were not synchronized with the audio..."*). Participants (17) also suggested **usability improvements** to enhance their experience: (*P13: "...I need to then adjust the distance in such a way that it's not disturbing me in while I'm looking around at everything..."*), (*P2: "...the static ones? I wanted them a little bit lower, but even in the settings, they were already like, it couldn't go* *any lower..."*). Participants (14) were allowed to make open suggestions, which did not negatively affect their experience but **were rather "could-haves"**: (*P9: "...the colour, because I was reading the text colour in white and maybe to put in yellow or something yellow and black, yellow or black, the three colours, black, yellow and white..."*)

6 DISCUSSION

The aim of this study was to explore the integration of subtitles in a VR theatre and assess their impact on user experience. The findings focused on balancing improved spectator engagement with minimizing distractions. This investigation was guided by three specific SRQs (Section 1).

6.1 Static over Dynamic: Mixed Method Analysis

6.1.1 Eye-Tracking Data. Participants were observed to have a more consistent fixation pattern on static subtitles than on dynamic subtitles. The findings support the interpretation that the conventional habit of reading subtitles at the lower part of the screen contributes to this preference. Additionally, the static nature of the subtitles likely imposes less physical strain on the eye muscles, as it minimizes the need for frequent eye movement between the subtitles and the objects or scenes behind them. This suggests that static subtitles provide a more stable viewing experience, reducing the cognitive load (CL) [32] to constantly switch between subtitles and the context.

6.1.2 Questionnaire Data. These suggestions are further validated by the results of the questionnaires, which show the same preference - 9 out of 13 observed factors ranking in favour of static subtitles (bolded in Table 1). The results from the administered questionnaires further corroborate the superiority of static subtitles in the VR theatre context. Five of the seven evaluated factors—interface, experience rating, UX rating, immersion and presence, and realism—indicated a significant preference for static subtitles over dynamic ones. Participants consistently rated their experience with static subtitles higher, suggesting that these subtitles better support the overall immersive experience in a VR environment. Four out of six factors favoured static subtitles in the cognitive task load questionnaire, which assessed mental demand, physical demand, performance, and frustration. Participants reported lower mental and physical demands, better performance, and less frustration when using static subtitles. These results highlight the ease and efficiency of processing static subtitles, which are perceived as less cognitively and physically demanding.

6.1.3 Qualitative Data. What equalises the results is the qualitative data, which suggests the choice for subtitles is very personalised, subjective, and highly context-dependent. For example, participants who preferred static subtitles explained their preference was due to their familiarity and comfort with traditional viewing environments [16], such as two-dimensional platforms like YouTube, Netflix, and TV. On the other hand, dynamic subtitles were favoured as they could be overlaid directly on top of the speakers, preventing them from being obscured by VFX and maintaining a fixed position within the scene. Furthermore, some participants preferred static subtitles due to the greater degree of freedom they provided. With static subtitles, users felt more at ease moving around and exploring the VR environment without fearing missing out on context, as the subtitles remained fixed.

6.2 Hybrid Approach and Accessibility Improvements

As mentioned in the previous section, most participants indicated that their choice would depend on the situation, suggesting a potential for a hybrid approach. Only two participants explicitly mentioned this approach, which involves dynamically switching subtitle variants based on the viewer's gaze: dynamic subtitles when looking at the scene and static subtitles when looking elsewhere, ensuring that users do not miss the context.

Additionally, some users proposed simplifying the main controls, similar to remote control, for essential functions such as increasing font size, minimal positional changes or changing the colour. They recommended providing a single tutorial at the beginning of the theatre experience, allowing users to navigate the most necessary settings easily while reserving more detailed adjustments for deeper settings menus.

6.3 Limitations

One primary limitation is the sample size of 19 participants, which, although calculated to be adequate through a power analysis, remains relatively small for generalizing the findings to a broader population. The research study on eye-tracking during a theatrical performance exhibited notable limitations, particularly concerning internal validity.

Specifically, there was a lack of pre-screening to determine participants' familiarity with the play. This omission is important as prior knowledge of the narrative and characters could influence engagement and eye movement patterns, potentially confounding the results. Familiarity with the play might lead to more focused attention on key elements or characters. At the same time, a lack of prior knowledge could result in more exploratory or less engaged viewing behaviour. Initially, pre-screening was not conducted due to logistical constraints, and the focus was on other research design aspects. However, in retrospect, including pre-screening would have enhanced the study by ensuring more consistent and interpretable data. Future studies should address this aspect to improve internal validity and avoid potential confounding factors related to participants' familiarity with the play.

Due to technical errors and the limitations of the machine, the error profiling scene could not be conducted. Error profiling is essential in eye-tracking studies as it helps identify and correct inaccuracies in data collection. Consequently, the absence of an initial round of error profiling before the commencement of the play may compromise the reliability of the eye-tracking data. Without this important step, the analysis relies on raw data that may contain uncorrected measurement errors, potentially reducing the precision and validity of the findings.

Additionally, the choice of a monologue from Euripides' "Hippolytus" as the central focus of the study posed another limitation, largely due to the app developers' decisions before this study. Most participants found this segment to be rather static and disengaging, as evidenced by the observation that participants, on average, diverted their gaze elsewhere for periods ranging from 30 to 50 seconds within the 4:30-minute duration of the play. The static nature of a monologue, as opposed to a multi-speaker interaction, likely contributed to decreased engagement and varying attention spans among the audience. Future research could benefit from including more dynamic scenes with multiple speaker dialogues to maintain participant engagement and provide a more comprehensive analysis of eye-tracking behaviour in response to different theatrical elements.

7 CONCLUSION AND FUTURE WORK

This thesis presents a user study (conducted in a VR theatre showing a Greek play) regarding subtitles (English) in IVEs to give an insight towards subtitle placements between two variants of subtitles - static and dynamic, particularly in a VR theatre. For this study, it was required to have integrated the OVR eye-tracking functionality, validated by a set of literature-backed questionnaires, have interview questions prepared, conduct the experiment, and analyse the results. The study was carried out with participants (N = 19) from various backgrounds related to theatre and immersive technologies. A mixed-method approach was used to analyze the user experience.

Regarding *SRQ1*, static subtitles were generally preferred over dynamic ones due to their stability and ease of reading, reducing cognitive load and physical strain, despite dynamic subtitles sometimes enhancing engagement by integrating into the scene. For *SRQ2*, best practices for subtitle style in XR environments suggest that static subtitles are more effective for maintaining readability and immersion, as they provide a consistent focal point in the presence of VFX. Lastly, for *SRQ3*, the display of subtitles did not contribute to cybersickness; rather, the overall VR experience was more influential, especially for those unfamiliar with VR technology. This preliminary study has also highlighted areas for future research:

(1) As participants suggested during interviews, future studies should incorporate a multi-speaker layout. This adjustment will allow for a better evaluation of the potential benefits of dynamic subtitles, particularly in scenarios involving multiple speakers and more complex dialogues.

(2) Participants also preferred having options between different subtitle styles, such as yellow on black or blue on white. Integrating these customisable subtitle styles into the application will be important for the next study to accommodate individual preferences and enhance user experience.

(3) Additionally, several participants noted the low-poly nature of the virtual environment, which, despite its simplicity, did not fully meet their expectations for emotional depth and theatrical presence. Future iterations of the study should aim to enhance the visual and emotional fidelity of the VR scene better to replicate the immersive qualities of a physical theatre. This improvement will likely increase participant engagement and provide a more accurate assessment of subtitle efficacy in a highly immersive setting.

REFERENCES

[1] [n. d.]. hcistats:start [Koji Yatani's Course Webpage]. https://yatani.jp/teaching/doku.php?id=hcistats:start

[2] 2021. How VR and AR are Changing the World of Immersive Theater. https://amt-lab.org/blog/2021/8/how-vr-and-ar-are-changing-the-world-of-immersive-theater

[3] 2022. Meta Quest Pro hands-on: The $1,500 headset that 'will enable the metaverse'. https://www.engadget.com/meta-quest-pro-hands-on-specs-price-the-headset-that-will-enable-the-metaverse-173253437-173253029.html

[4] 2023. sychology of Aesthetics, Creativity, and the Arts Immersive-360° theatre: user experience in the virtual auditorium and platform efficacy for current and underserved audiences. https://doi.org/10.31234/osf.io/2fyv7

[5] 2024. What Is the Virtuality Continuum? — updated 2024. https://www.interaction-design.org/literature/topics/virtuality-continuum

[6] A. Brown, J. Turner, J. Patterson, A. Schmitz, M. Armstrong, and M. Glancy. 2018. Exploring Subtitle Behaviour for 360° Video. BRITISH BROADCASTIND CORPORATION, 40. https://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP330.pdf

[7] Belén Agulló and Anna Matamala. 2020. Subtitles in virtual reality: Guidelines for the integration of subtitles in 360º content. *Íkala, Revista de Lenguaje y Cultura* 25, 3 (Sept. 2020), 643–661. https://doi.org/10.17533/udea.ikala.v25n03a03 Number: 3.

[8] Mike Armstrong, Andy Brown, Michael Crabb, Chris J. Hughes, Rhianne Jones, and James Sandford. 2016. Understanding the Diverse Needs of Subtitle Users in a Rapidly Evolving Media Landscape. *SMPTE Motion Imaging Journal* 125, 9 (Nov. 2016), 33–41. https://doi.org/10.5594/JMI.2016.2614919

[9] Allison Bayro, Ting Dai, and Heejin Jeong. 2022. Does Gender Influence Presence in Virtual Environments? An Analysis of Open-Source Igroup Presence Questionnaire Data. In *2022 IEEE 3rd International Conference on Human-Machine Systems (ICHMS)*. 1–6. https://doi.org/10.1109/ICHMS56717.2022.9980589

[10] Pauline Bimberg, Tim Weissker, and Alexander Kulik. 2020. On the Usage of the Simulator Sickness Questionnaire for Virtual Reality Research. IEEE Computer Society, 464–467. https://doi.org/10.1109/VRW50115.2020.00098

[11] Sharon Black. 2022. Subtitles as a Tool to Boost Language Learning and Intercultural Awareness? : Children's Views and Experiences of Watching Films and Television Programmes in Other Languages with Interlingual Subtitles. *Journal of Audiovisual Translation* 5, 1 (April 2022), 73–93. https://doi.org/10.47476/jat.v5i1.2022.157 Number: 1.

[12] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* (Jan. 2006). https://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa Publisher: Taylor & Francis Group.

[13] M. Brescia-Zapata, K. Krejtz, A. Duchowski, and C. Hughes. 2022. VR 360º subtitles: designing a test suite with eye-tracking technology. *Journal of Audiovisual Translation* 6, 2 (Dec. 2022). https://salford-repository.worktribe.com/output/1324830/vr-360o-subtitles-designing-a-test-suite-with-eye-tracking-technology

[14] Marta Brescia-Zapata, Krzysztof Krejtz, Andrew T. Duchowski, Christopher J. Hughes, and Pilar Orero. 2023. Eye-tracked Evaluation of Subtitles in Immersive VR 360\circ Video. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 769–770. https://doi.org/10.1109/VRW58643.2023.00227

[15] john Brooke. 1996. SUS: A 'Quick and Dirty' Usability Scale. In *Usability Evaluation In Industry*. CRC Press. Num Pages: 6.

[16] Andy Brown, Rhia Jones, Mike Crabb, James Sandford, Matthew Brooks, Mike Armstrong, and Caroline Jay. 2015. Dynamic Subtitles: The User Experience. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '15)*. Association for Computing Machinery, New York, NY, USA, 103–112. https://doi.org/10.1145/2745197.2745204

[17] Jean-Rémy Chardonnet, Mohammad Ali Mirzaei, and Frédéric Mérienne. 2017. Features of the Postural Sway Signal as Indicators to Estimate and Predict Visually Induced Motion Sickness in Virtual Reality. *International Journal of Human–Computer Interaction* 33, 10 (Oct. 2017), 771–785. https://doi.org/10.1080/10447318.2017.1286767 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447318.2017.1286767.

[18] Michael Crabb, Rhianne Jones, and Mike Armstrong. 2015. The Development of a Framework for Understanding the UX of Subtitles. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility (ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 347–348. https://doi.org/10.1145/2700648.2811372

[19] Sagi Eden, Amit Livne, Oren Sar Shalom, Bracha Shapira, and Dietmar Jannach. 2022. Investigating the Value of Subtitles for Improved Movie Recommendations. In *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '22)*. Association for Computing Machinery, New York, NY, USA, 99–109. https://doi.org/10.1145/3503252.3531291

[20] Deborah I. Fels. 2002. Accessible Digital Media. In *Computers Helping People with Special Needs*, Klaus Miesenberger, Joachim Klaus, and Wolfgang Zagler (Eds.). Springer, Berlin, Heidelberg, 282–283. https://doi.org/10.1007/3-540-45491-8_59

[21] W. S. Hadley. 2012. *The Hippolytus of Euripides*. Cambridge University Press.

[22] Sandra G. Hart. 1986. NASA Task Load Index (TLX). https://ntrs.nasa.gov/citations/20000021487 NTRS Author Affiliations: NASA Ames Research Center NTRS Document ID: 20000021487 NTRS Research Center: Ames Research Center (ARC).

[23] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (Oct. 2006), 904–908. https://doi.org/10.1177/154193120605000909 Publisher: SAGE Publications Inc.

[24] Marc Helmold. 2023. Extended Reality (XR) in QM. In *Virtual and Innovative Quality Management Across the Value Chain: Industry Insights, Case Studies and Best Practices*, Marc Helmold (Ed.). Springer International Publishing, Cham, 21–26. https://doi.org/10.1007/978-3-031-30089-9_3

[25] Melinda Hestiana and Anita Anita. 2022. THE ROLE OF MOVIE SUBTITLES TO IMPROVE STUDENTS' VOCABULARY. *Journal of English Language Teaching and Learning* 3, 1 (July 2022), 46–53. https://doi.org/10.33365/jeltl.v3i1.1715 Number: 1.

[26] Baosheng James Hou, Yasmeen Abdrabou, Florian Weidner, and Hans Gellersen. 2024. Unveiling Variations: A Comparative Study of VR Headsets Regarding Eye Tracking Volume, Gaze Accuracy, and Precision. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 650–655. https://doi.org/10.1109/VRW62533.2024.00127

[27] Sarah Hudson, Sheila Matson-Barkat, Nico Pallamin, and Guillaume Jegou. 2019. With or without you? Interaction and immersion in a virtual reality experience. *Journal of Business Research* 100 (July 2019), 459–468. https://doi.org/10.1016/j.jbusres.2018.10.062

[28] Claire L. Hughes, Cali Fidopiastis, Kay M. Stanney, Peyton S. Bailey, and Ernesto Ruiz. 2020. The Psychometrics of Cybersickness in Augmented Reality. *Frontiers in Virtual Reality* 1 (2020). https://www.frontiersin.org/articles/10.3389/frvir.2020.602954

[29] R. Jain. 2004. Quality of Experience. *IEEE MultiMedia* 11, 1 (Jan. 2004), 96–95.

[30] Jeroen van Paridon, and Bill Thompson. 2021. subs2vec: Word embeddings from subtitles in 55 languages. *Behavior Research Methods* 53, 2 (April 2021), 629–655. https://doi.org/10.3758/S13428-020-01406-3 Publisher: Springer US.

[31] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology* 3, 3 (July 1993), 203–220. https://doi.org/10.1207/s15327108ijap0303_3 Publisher: Taylor & Francis _eprint: https://doi.org/10.1207/s15327108ijap0303_3.

[32] Sixin Liao Kruger, Jan-Louis. 2023. Cognitive processing of subtitles: Charting the future by mapping the past. In *The Routledge Handbook of Translation, Interpreting and Bilingualism*. Routledge. Num Pages: 16.

[33] Kuno Kurzhals, Emine Cetinkaya, Yongtao Hu, Wenping Wang, and Daniel Weiskopf. 2017. Close to the Action: Eye-Tracking Evaluation of Speaker-Following Subtitles. 6559–6568. https://doi.org/10.1145/3025453.3025772

[34] Sueyoon Lee, Irene Viola, Silvia Rossi, Zhirui Guo, Ignacio Reimat, Kinga Ławicka, Alina Striner, and Pablo Cesar. 2024. Designing and Evaluating a VR Lobby for a Socially Enriching Remote Opera Watching Experience. *IEEE Transactions on Visualization and Computer Graphics* 30, 5 (May 2024), 2055–2065. https://doi.org/10.1109/TVCG.2024.3372081 Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[35] James Lewis. 2018. The System Usability Scale: Past, Present, and Future. *International Journal of Human-Computer Interaction* (March 2018), 1–14. https://doi.org/10.1080/10447318.2018.1455307

[36] Shadi R. Masadeh and Saba A. Soub. 2020. Educational Video Classification Using Fuzzy Logic Classifier Based on Arabic Closed Captions. In *Intelligent Computing Paradigm and Cutting-edge Technologies (Learning and Analytics in Intelligent Systems)*, Lakhmi C. Jain, Sheng-Lung Peng, Basim Alhadidi, and Souvik Pal (Eds.). Springer International Publishing, Cham, 133–138. https://doi.org/10.1007/978-3-030-38501-9_13

[37] Miguel Melo, Guilherme Gonçalves, josé Vasconcelos-Raposo, and Maximino Bessa. 2023. How Much Presence is Enough? Qualitative Scales for Interpreting the Igroup Presence Questionnaire Score. *IEEE Access* 11 (2023), 24675–24685. https://doi.org/10.1109/ACCESS.2023.3254892 Conference Name: IEEE Access.

[38] Pranav Pidathala, Dawson Franz, James Waller, Raja Kushalnagar, and Christian Vogler. 2022. Live Captions in Virtual Reality (VR). https://doi.org/10.48550/arXiv.2210.15072 arXiv:2210.15072 [cs].

[39] Krzysztof Pietroszek, Manuel Rebol, and Becky Lake. 2022. Dill Pickle: Interactive Theatre Play in Virtual Reality. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology (VRST '22)*. Association for Computing Machinery, New York, NY, USA, 1–2. https://doi.org/10.1145/3562939.3565678

[40] Sylvia Rothe, Kim Tran, and Heinrich Hußmann. 2018. Dynamic Subtitles in Cinematic Virtual Reality. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '18)*. Association for Computing Machinery, New York, NY, USA, 209–214. https://doi.org/10.1145/3210825.3213556

[41] Augusto Rupérez Micola, Ainoa Aparicio Fenoll, Albert Banal-Estañol, and Arturo Bris. 2019. TV or not TV? The impact of subtitling on English skills. *Journal of Economic Behavior & Organization* 158 (Feb. 2019), 487–499. https://doi.org/10.1016/j.jebo.2018.12.019

[42] Rufat Rzayev, Sabrina Hartl, Vera Wittmann, Valentin Schwind, and Niels Henze. 2020. Effects of position of real-time translation on AR glasses. In *Proceedings of Mensch und Computer 2020 (MuC '20)*. Association for Computing Machinery, New York, NY, USA, 251–257. https://doi.org/10.1145/3404983.3405523

[43] Rufat Rzayev, Paweł W. Woźniak, Tilman Dingler, and Niels Henze. 2018. Reading on Smart Glasses: The Effect of Text Position, Presentation Type and Walking. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3173574.3173619

[44] Lena Stubbemann, Dominik Dürrschnabel, and Robert Refflinghaus. 2021. Neural Networks for Semantic Gaze Analysis in XR Settings. In *ACM Symposium on Eye Tracking Research and Applications (ETRA '21 Full Papers)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3448017.3457380

[45] Yobbahim J. Vite and Yaoping Hu. 2023. The Role of Attention and Cognitive Workload in Measuring Levels of Task Complexity within Virtual Environments. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 651–652. https://doi.org/10.1109/VRW58643.2023.00168

[46] B. Vivekanandam. 2023. Recent development in Extended Reality technologies. *rrrj* 2, 1 (June 2023), 135–144. https://irojournals.com/rrrj/articles/view/2/1/11

[47] Shu Wei, Desmond Bloemers, and Aitor Rovira. 2023. A Preliminary Study of the Eye Tracker in the Meta Quest Pro. In *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences (IMX '23)*. Association for Computing Machinery, New York, NY, USA, 216–221. https://doi.org/10.1145/3573381.3596467

[48] Aleksandra Zheleva, Anne Roos Smink, Paul Hendriks Vettehen, and Paul Ketelaar. 2021. Modifying the Technology Acceptance Model to Investigate Behavioural Intention to Use Augmented Reality. In *Augmented Reality and Virtual Reality*, M. Claudia tom Dieck, Timothy H. Jung, and Sandra M. C. Loureiro (Eds.). Springer International Publishing, Cham, 125–137. https://doi.org/10.1007/978-3-030-68086-2_10

Appendix A LIST OF ABBREVIATIONS AND ACRONYMS

- AA - Session A After experience (regarding SSQ)
- AB - Session A Before experience (regarding SSQ)
- AR - Augmented Reality
- BA - Session B After experience (regarding SSQ)
- BB - Session B Before experience (regarding SSQ)
- BI - Behavioral Intention
- CL - Cognitive Load
- CSV - Comma-Separated Values
- HMD - Head-Mounted Display
- IPQ - Igroup Presence Questionnaire
- IVE - Immersive Virtual Environment
- MR - Mixed Reality
- NASA-TLX - NASA Task Load Index
- OVR - Oculus Virtual Reality
- RAM - Random Access Memory
- RQ - Research Question
- SD - Standard Deviation
- SDK - Software Development Kit
- SRQ - Sub-Research Question
- SSQ - Simulator Sickness Questionnaire
- SUS - System Usability Scale
- TV - Television
- UX - User Experience
- VFX - Visual Effects
- VR - Virtual Reality
- XR - Extended Reality

Appendix B PHOTOS FROM THE APPLICATION

Figure 7: Settings Pane with both Subtitle Variants Enabled from a different POV

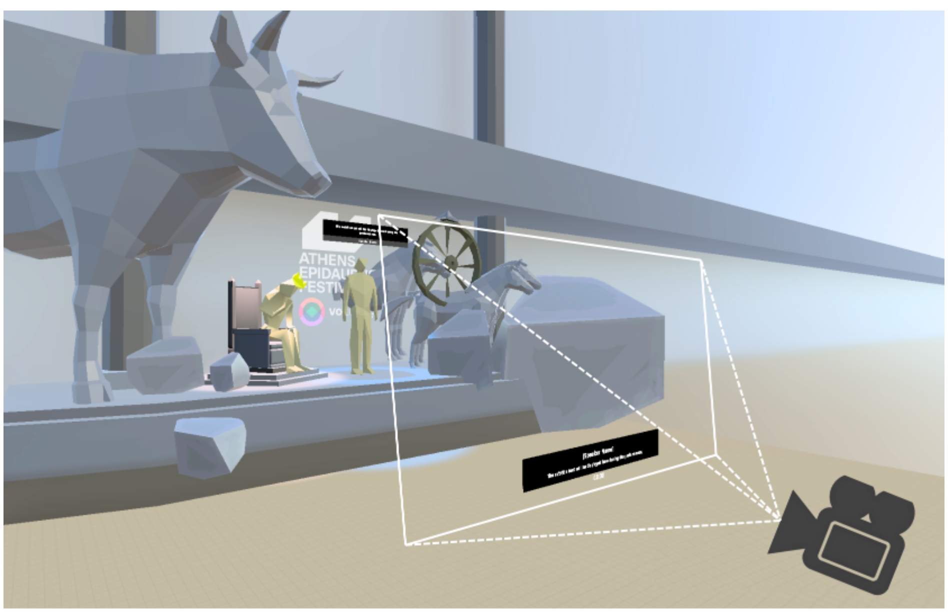

Figure 8: Theatre Scene Layout from a different POV

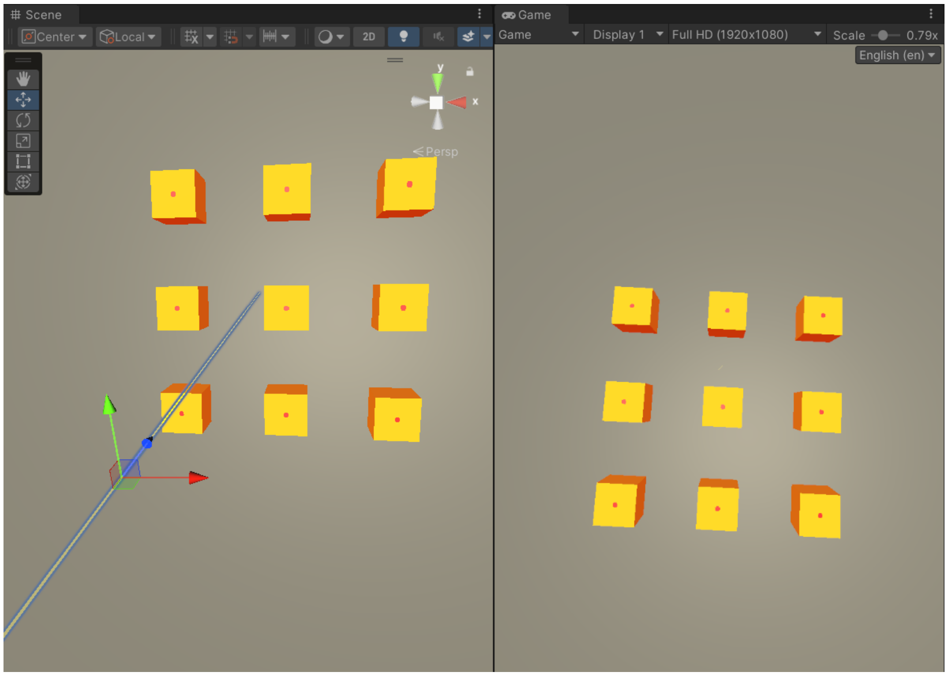

Figure 9: Error Profiling Scene (on the left is the scene in development, and on the right is a display of what the participant sees through the headset)

Appendix C THEATRE SCENE LAYOUT FROM THE PARTICIPANT CAMERA POINT OF VIEW (POV)

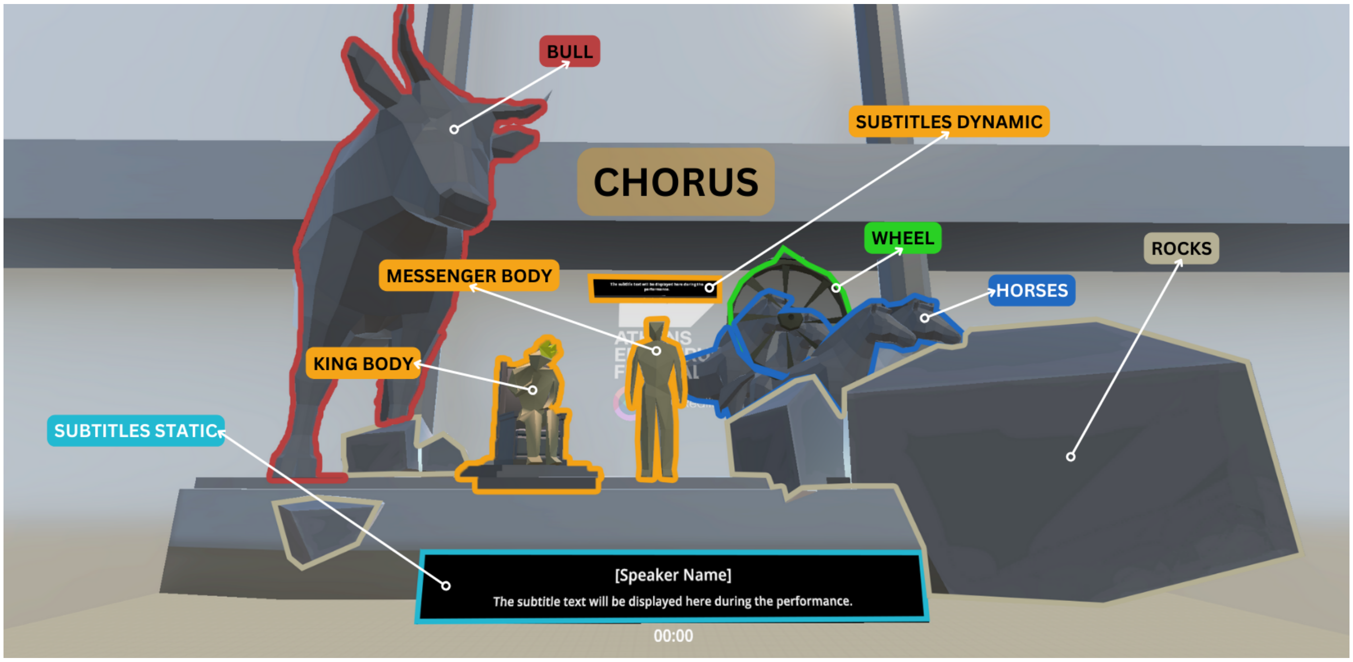

Figure 10: Theatre scene layout from the participant camera point of view (POV). For demonstrative purposes, in this screenshot of the Unity VR Theatre application, both subtitle variants are enabled along with all VFX game objects (highlighted and colour-coded) to which a trigger eye gaze collider has been added

Appendix D G*POWER 3.1 STATISTICAL TEST

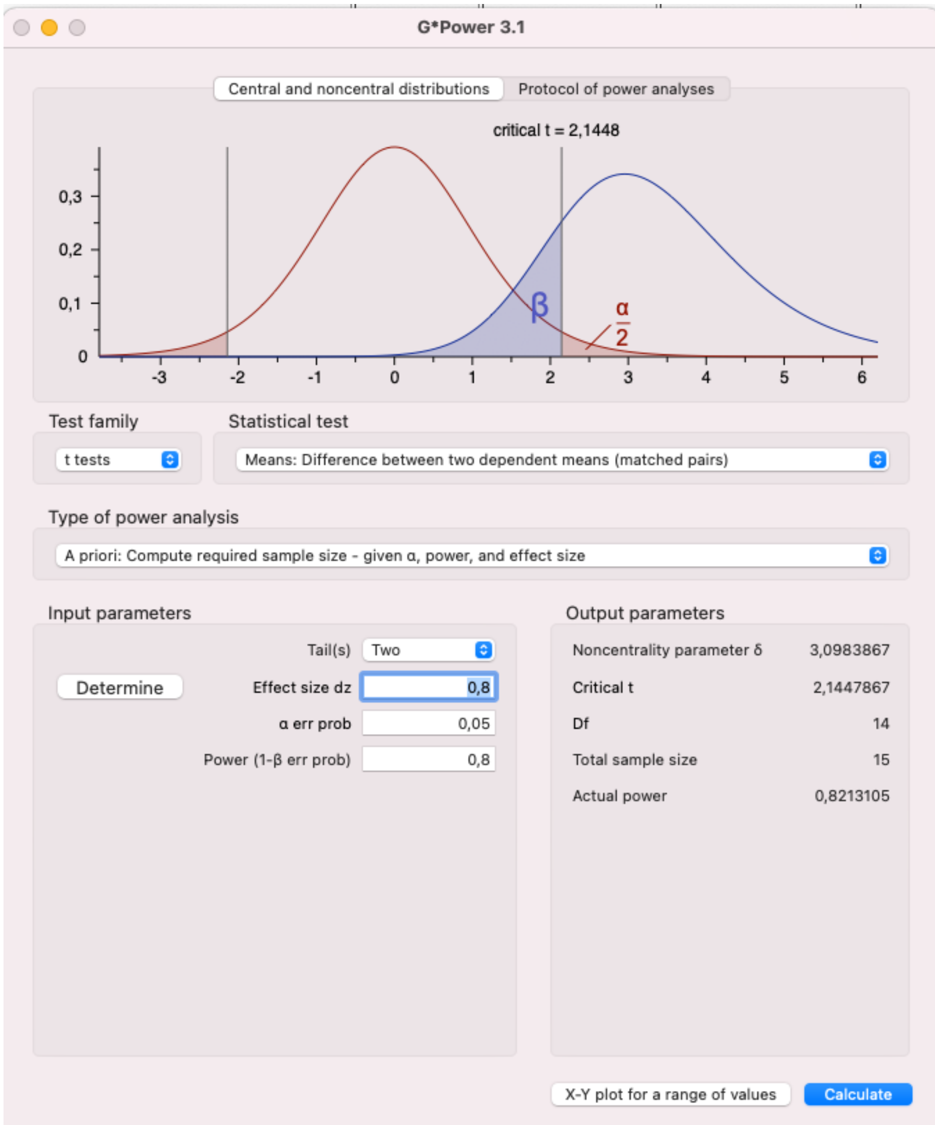

Figure 11: G*Power 3.1 Statistical Test

Appendix E QUESTIONNAIRES USED

E.1 SSQ Questionnaire

Participant ID: ________________________________ Session: _______________

Simulator Sickness Questionnaire

How much each symptom below is affecting you right now ***(Before experience)***

		None	Slight	Moderate	Severe
01	General discomfort	◯	◯	◯	◯
02	Fatigue	◯	◯	◯	◯
03	Headache	◯	◯	◯	◯
04	Eye strain	◯	◯	◯	◯
05	Difficulty focusing	◯	◯	◯	◯
06	Salivation increasing	◯	◯	◯	◯
07	Sweating	◯	◯	◯	◯
08	Nausea	◯	◯	◯	◯
09	Difficulty concentrating	◯	◯	◯	◯
10	Fullness of head	◯	◯	◯	◯
11	Blurred vision	◯	◯	◯	◯
12	Dizziness with eyes open	◯	◯	◯	◯
13	Dizziness with eyes closed	◯	◯	◯	◯
14	*Vertigo	◯	◯	◯	◯
15	**Stomach awareness	◯	◯	◯	◯
16	Burping	◯	◯	◯	◯

* Vertigo is experienced as loss of orientation with respect to vertical upright.
** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

VR Theater User Study
CWI, Amsterdam

Page 1 of 8

Participant ID: ______________________________ Session: ______________

Simulator Sickness Questionnaire

How much each symptom below is affecting you right now ***(After experience)***

		None	Slight	Moderate	Severe
01	General discomfort	○	○	○	○
02	Fatigue	○	○	○	○
03	Headache	○	○	○	○
04	Eye strain	○	○	○	○
05	Difficulty focusing	○	○	○	○
06	Salivation increasing	○	○	○	○
07	Sweating	○	○	○	○
08	Nausea	○	○	○	○
09	Difficulty concentrating	○	○	○	○
10	Fullness of head	○	○	○	○
11	Blurred vision	○	○	○	○
12	Dizziness with eyes open	○	○	○	○
13	Dizziness with eyes closed	○	○	○	○
14	*Vertigo	○	○	○	○
15	**Stomach awareness	○	○	○	○
16	Burping	○	○	○	○

* Vertigo is experienced as loss of orientation with respect to vertical upright.
** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

E.2 NASA-TLX

Participant ID: ________________________________ Session: _______________

NASA Task Load Index (TLX)

We are interested in the workload you experienced while completing this task. As workload can be caused by several different factors, we ask you to rate several of the factors individually on the scales provided.

Note: Performance goes from good on the left to bad on the right.

Mental Demand: How mentally demanding was the task?

Very Low | Very High

Physical Demand: How physically demanding was the task?

Very Low | Very High

Temporal Demand: How hurried or rushed was the pace of the task?

Very Low | Very High

Performance: How successful were you in accomplishing what you were asked to do?

Perfect | Failure

Effort: How hard did you have to work to accomplish your level of performance?

Very Low | Very High

Frustration: How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low | Very High

VR Theater User Study
CWI, Amsterdam

Page 3 of 8

E.3 UX Framework For Subtitles

Participant ID: ______________________________ Session: ______________

UX (User Experience) framework for subtitles

1. I enjoyed the clip
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Strongly agree

2. I found it easy to locate the subtitles
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Strongly agree

3. I found it easy to read the subtitles
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Strongly agree

4. I am confident that I followed all the speech
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Strongly agree

5. I found it easy to follow the visual action
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Strongly agree

6. I feel that I understand what was happening in the vr theater experience
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Strongly agree

7. I felt immersed in the scene, like I was there
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Strongly agree

E.4 SUS

Participant ID: ______________________________ Session: ______________

System Usability Scale (SUS)

1. I think that I would like to use this subtitle interface frequently.
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree
2. I found the subtitle interface unnecessarily complex.
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree
3. I thought the subtitles interface was easy to use.
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree
4. I think that I would need the support of a technical person to be able to use this subtitles interface.
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree
5. I found the various functions in this subtitles interface were well integrated.
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree
6. I thought there was too much inconsistency in this subtitles interface.
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree
7. I would imagine that most people would learn to use this subtitles interface very quickly.
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree
8. I found the subtitles interface very cumbersome to use.
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree
9. I felt very confident using the subtitles interface.
   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree
10. I needed to learn a lot of things before I could get going with this subtitles interface.
    Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree

E.5 IPQ

Participant ID: ________________________________ Session: _______________

Presence (IPQ)

1. In the computer generated world, I had a sense of "being there"

   Not at all ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Very much

2. Somehow I felt that the virtual world surrounded me.

   Fully disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Fully agree

3. I felt like I was just perceiving pictures.

   Fully disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Fully agree

4. I did not feel present in the virtual space.

   Did not feel ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Felt present

5. I had a sense of acting in the virtual space, rather than operating something from outside.

   Fully disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Fully agree

6. I felt present in the virtual space.

   Fully disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Fully agree

7. How aware were you of the real-world surroundings while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)?

   Extremely aware ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Not aware at all

8. I was not aware of my real environment.

   Fully disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Fully agree

9. I still paid attention to the real environment.

   Fully disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Fully agree

10. I was completely captivated by the virtual world.

    Fully disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Fully agree

11. How real did the virtual world seem to you?

    Completely real ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Not real at all

12. How much did your experience in the virtual environment seem consistent with your real-world experience?

    Not consistent ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Very consistent

13. How real did the virtual world seem to you?

    As real as an imagined world ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Indistinguishable from the real world

14. The virtual world seemed more realistic than the real world.

    Fully disagree ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 Fully agree

E.6 Immersion

Participant ID: ______________________________ Session: ______________

Immersion in VR

Please answer the following questions about your VR experience

1. I felt free to interact with the environment as if I was in the real world.
   Not at all ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 A lot

2. I felt detached from the outside world.
   Not at all ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 A lot

3. I felt completely immersed that I forgot about my everyday concerns.
   Not at all ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 A lot

4. I felt free to explore and look where I wanted.
   Not at all ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 A lot

5. I felt free to move around in the VR experiment.
   Not at all ◯1 ◯2 ◯3 ◯4 ◯5 ◯6 ◯7 A lot

E.7 Behavioural Intention and General

Participant ID: ______________________________ Session: ______________

Behavioral Intention

1. I think subtitles are useful for watching theater plays in foreign language

   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree

2. I am very interested in using subtitles

   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree

3. I would like to use subtitles to help watching theater plays in foreign languages

   Strongly Disagree ◯1 ◯2 ◯3 ◯4 ◯5 Strongly agree

General Question

How do you like the subtitles interface?

◯Bad ◯Poor ◯Fair ◯Good ◯Excellent

How would you rate your experience with VR Theater experience?

◯Bad ◯Poor ◯Fair ◯Good ◯Excellent

Appendix F INTERVIEW QUESTIONS

Interview Questions (10min)

VR Theater Experiment dashboard_VR_Theater

---

1. What do you think of the theater experience in itself?
    a. Did you enjoy it? Was it easy to follow?

2. What did you think of the subtitles presentation?
    a. Would you actually use them? Or switch them off?

3. Among 2D and 3D subtitles, which ones were easy to read?

4. Imagine the theater play being longer, which subtitle would you prefer for the whole duration of that longer play, 2D or 3D? Why?

5. If possible, would you change the position of subtitles on the screen?
    a. Any preference for specific subtitle placements or position?

6. Did you have any problems identifying which person was speaking when VFX were active?

7. What do you think about the effect of visuals on subtitles?

8. Is there anything you would have changed, or would like to be able to adjust?

9. Any feedback, comments for us.

Appendix G FORMULAS USED

G.1 SSQ

Based on Kennedy et al.[31]

```
function [w_N, w_O, w_D, TS] = calculateRowWiseSum(Table)
    % Calculate row-wise sum for w_Nausea
    w_N = sum(Table{:, ["GeneralDiscomfort", "SalivationIncreasing", "Sweating", ...
                     "Nausea", "DifficultyConcentrating", "x__StomachAwareness", ...
                     "Burping"]}, 2) * 9.54;

    % Calculate row-wise sum for w_Occulomotor
    w_O = sum(Table{:, ["GeneralDiscomfort", "Fatigue", "Headache", ...
                     "EyeStrain", "DifficultyFocusing", "DifficultyConcentrating", ...
                     "BlurredVision"]}, 2) * 7.58;

    % Calculate row-wise sum for w_Disorientation
    w_D = sum(Table{:, ["DifficultyFocusing", "Nausea", "FullnessOfHead", ...
                     "BlurredVision", "DizzinessWithEyesOpen", ...
                     "DizzinessWithEyesClosed", "x_Vertigo"]}, 2) * 13.92;

    % Calculate the total sum TS Total Score
    TS = (w_N + w_O + w_D) * 3.74;
end
```

G.2 NASA-TLX

Based on Hart et al.[22]

```
NASA_2D = 5*([T1{2:2:end, 2:7};T2{1:2:end, 2:7}]-1);
NASA_3D = 5*([T2{2:2:end, 2:7};T1{1:2:end, 2:7}]-1);
```

G.3 UX for Subtitles

Based on Crabb et al.[18]

```
UX_2D = mean([ T1{2:2:end, 2:8} ; T2{1:2:end, 2:8} ], 2);
UX_3D = mean([ T2{2:2:end, 2:8} ; T1{1:2:end, 2:8} ], 2);
```

G.4 SUS

Based on Brooke et al.[15]

```
SUS_2D = 2.5 * (20 + sum(Odd_Questions_2D,2) - sum(Even_Questions_2D,2));

SUS_3D = 2.5 * (20 + sum(Odd_Questions_3D,2) - sum(Even_Questions_3D,2));
```

G.5 Immersion and Presence, and Realism

Based on Lee et al.[34]

```
T1_IPQ(:, [5, 8, 11]) = [];
T2_IPQ(:, [5, 8, 11]) = [];

T1_Immersion(:, 6) = [];
T2_Immersion(:, 6) = [];

%%
IP_Factor_2D = mean(horzcat([ T1_IPQ{2:2:end,3:9} ; T2_IPQ{1:2:end,3:9} ] ,
[T1_Immersion{2:2:end,3:5} ; T2_Immersion{1:2:end,3:5}]), 2);
IP_Factor_3D = mean(horzcat([ T2_IPQ{2:2:end,3:9} ; T1_IPQ{1:2:end,3:9} ] ,
[T2_Immersion{2:2:end,3:5} ; T1_Immersion{1:2:end,3:5}]), 2);
```

```
R_Factor_2D = mean(horzcat([ T1_IPQ{2:2:end,10:12} ; T2_IPQ{1:2:end,10:12} ] ,
[T1_Immersion{2:2:end, 2} ; T2_Immersion{1:2:end, 2}]) , 2);
R_Factor_3D = mean(horzcat([ T2_IPQ{2:2:end,10:12} ; T1_IPQ{1:2:end,10:12} ] ,
[T2_Immersion{2:2:end, 2} ; T1_Immersion{1:2:end, 2}]) , 2);
```

G.6 Behavioural Intention

Based on Zheleva et al.[48]

```
Q1_2D = [sum(T1{2:2:end,2},2);sum(T2{1:2:end,2},2)];
Q1_3D = [sum(T2{2:2:end,2},2);sum(T1{1:2:end,2},2)];

Q2_2D = [sum(T1{2:2:end,3},2);sum(T2{1:2:end,3},2)];
Q2_3D = [sum(T2{2:2:end,3},2);sum(T1{1:2:end,3},2)];

Q3_2D = [sum(T1{2:2:end,4},2);sum(T2{1:2:end,4},2)];
Q3_3D = [sum(T2{2:2:end,4},2);sum(T1{1:2:end,4},2)];

printResults(mean(horzcat(Q1_2D, Q2_2D, Q3_2D),2), mean(horzcat(Q1_3D, Q2_3D, Q3_3D),2),'Overall')
```

G.7 General

Based on Quality of Experience (QoE) [29]

```
Q1_2D = [sum(T1{2:2:end,2},2);sum(T2{1:2:end,2},2)];
Q1_3D = [sum(T2{2:2:end,2},2);sum(T1{1:2:end,2},2)];

Q2_2D = [sum(T1{2:2:end,3},2);sum(T2{1:2:end,3},2)];
Q2_3D = [sum(T2{2:2:end,3},2);sum(T1{1:2:end,3},2)];
```

Appendix H CODE FOR THE STATISTICAL COMPARISON FUNCTION

```
function printResults(VarA, VarB, factor)

    %first question: is the data normally distributed?
    [H1, p1] = kstest(VarA);
    [H2, p2] = kstest(VarB);

    sprintf('Factor %s : normally distributed? \n 2D: p = %s\n 3D: p = %s\n', factor, p1, p2)

    [p,h,stats_ssq] = signrank(VarA, VarB, 'method', 'approximate');
    r = abs(stats_ssq.zval./sqrt(size(VarB, 1) + size(VarA, 1)));

    mean_2D = mean(VarA);
    std_2D = std(VarA);
    median_2D = median(VarA);

    mean_3D = mean(VarB);
    std_3D = std(VarB);
    median_3D = median(VarB);

    mean_2Dvs3D = mean([VarA; VarB]);
    std_2Dvs3D = std([VarA; VarB]);
    median_2Dvs3D = median([VarA; VarB]);

    sprintf('Score for static (mean = %.2f, SD = %.3f, median = %.f)', mean_2D, std_2D, median_2D)
    sprintf('Score for dynamic (mean = %.2f, SD = %.3f, median = %.f)', mean_3D, std_3D, median_3D)
    sprintf('Score for static vs dynamic (mean = %.2f, SD = %.3f, median = %.f, p = %.3f, r = %.3f)', ...
    mean_2Dvs3D, std_2Dvs3D, median_2Dvs3D, p, r)

end
```

```
printResults(SUS_2D, SUS_3D, 'SUS Score (out of 100)')
```

Where:

(1) ***VarA*** is a vector with data for static subtitles data for a given factor
(2) ***VarB*** is a vector with data for dynamic subtitles data for a given factor
(3) ***factor*** is a string label

and results are presented in this format (in this example, the results are for the SUS [15] test)

```
ans =
    'Factor SUS Score (out of 100): normally distributed?
      2D: p = 3.620218e-18
      3D: p = 3.620218e-18
     '
ans =
    'Score for static (mean = 83.42, SD = 10.145, median = 85)'
ans =
    'Score for dynamic (mean = 83.82, SD = 8.223, median = 85)'
ans =
    'Score for static vs dynamic (mean = 83.62, SD = 9.111, median = 85, p = 0.795, r = 0.042)'
```