

SUBTITLES IN VIRTUAL REALITY: GUIDELINES FOR THE INTEGRATION OF SUBTITLES IN 360° CONTENT

SUBTÍTULOS EN LA REALIDAD VIRTUAL: DIRECTRICES PARA LA INTEGRACIÓN DE SUBTÍTULOS EN CONTENIDO DE 360°

LES SOUS-TITRES DANS LA RÉALITÉ VIRTUELLE : DES LIGNES DIRECTRICES POUR L'INTÉGRATION DES SOUS-TITRES DANS UN CONTENU À 360°

Belén Agulló

Ph. D. in Translation and Intercultural Studies, Universitat Autònoma de Barcelona, Spain.
eLearning Director and Lead Media Researcher, Nimdzi Insights, Spain.
belen@nimdzi.com
<https://orcid.org/0000-0003-3270-1840>

Anna Matamala

Ph. D. in Applied Linguistics, Universitat Pompeu Fabra, Spain.
Associate professor, Universitat Autònoma de Barcelona, Spain.
anna.matamala@uab.cat
<https://orcid.org/0000-0002-1607-9011>

This work has been funded by the European Union's Horizon 2020 program, under agreement n° 761974 (ImAc project). Belén Agulló and Dr. Anna Matamala are members of the Trans-media Catalonia research group, an SGR research group funded by "Secretaria

ABSTRACT

Immersive content has become a popular medium for storytelling. This type of content is typically accessed via a head-mounted visual display within which the viewer is located at the center of the action with the freedom to look around and explore the scene. The criteria for subtitle position for immersive media still need to be defined. Guiding mechanisms are necessary for circumstances in which the speakers are not visible and viewers, lacking an audio cue, require visual information to guide them through the virtual scene. The aim of this reception study is to compare different subtitling strategies: always-visible position to fixed-position and arrows to radar. To do this, feedback on preferences, immersion (using the IPQ questionnaire) and head movements was gathered from 40 participants (20 hearing and 20 hard of hearing). Results show that always-visible subtitles with arrows are the preferred option. Always-visible and arrows achieved higher scores in the IPQ questionnaire than fixed-position and radar. Head-movement patterns show that participants move more freely when the subtitles are always-visible than when they are in a fixed position, meaning that with always-visible subtitles the experience is more realistic, because the viewers do not feel constrained by the implementation of subtitles.

Keywords: subtitles; virtual reality; immersive media; 360° content; hearing impairment; subtitling strategies.

RESUMEN

Los contenidos inmersivos se han convertido en un medio interesante para contar historias. A este tipo de contenido normalmente se accede con unas gafas de realidad virtual, a través de las cuales el usuario, que se encuentra en el centro de la acción, tiene la libertad de mirar y explorar la escena libremente. Los criterios para determinar la posición de los subtítulos en contenidos inmersivos todavía no se han definido. También se necesitan mecanismos de guía para indicar dónde está la persona que habla cuando está fuera del campo de visión y los usuarios no pueden contar con el audio para guiarse. El objetivo de este estudio de recepción es comparar diferentes

643

Received: 2019-08-05 / Accepted: 2020-02-27 / Published: 2020-09-17
<https://doi.org/10.17533/udea.ikala.v25n03a03>

d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya" (2017SGR113). This article is part of Belén Agulló's Ph. D. in Translation and Intercultural Studies at the Department of Translation, Interpreting and East Asian Studies (Departament de Traducció i d'Interpretació i d'Estudis de l'Àsia Oriental) of the Universitat Autònoma de Barcelona.

estrategias de subtitulación: subtítulos siempre visibles con subtítulos colocados en posiciones fijas, y flechas con radar. Para ello, se han recopilado los comentarios de 40 participantes (20 oyentes y 20 con pérdida auditiva) sobre las preferencias, la inmersión (con el cuestionario IPQ) y los patrones de movimiento de cabeza. Los resultados muestran que los subtítulos siempre visibles con flechas son la opción preferida. Además, estas soluciones obtuvieron mejores puntuaciones en el cuestionario IPQ. Los patrones de movimientos de cabeza muestran que los participantes se movían más libremente con los subtítulos siempre visibles, lo que supone que con este tipo de subtítulos la experiencia es más realista ya que los usuarios no se sienten cohibidos por la integración de los subtítulos.

Palabras clave: subtítulos; realidad virtual; medios inmersivos; contenido de 360°; pérdida auditiva; estrategias de subtitulación.

RÉSUMÉ

Les contenus immersifs sont devenus un moyen intéressant de raconter des histoires. Ce type de contenu est généralement accessible avec des lunettes de réalité virtuelle, avec lesquelles l'utilisateur est au cœur de l'action et à la liberté de regarder et d'explorer la scène librement. Les critères pour définir la position des sous-titres dans ce contenu n'a pas encore été définie. Des mécanismes de guidage sont également nécessaires pour indiquer où se trouve le locuteur lorsqu'il est hors du champ de vision et que les utilisateurs ne peuvent pas se fier à l'audio pour être guidés. L'objectif de cette étude de réception est de comparer différentes stratégies de sous-titrage : des sous-titres toujours visibles avec des sous-titres placés à des positions fixes et des flèches avec le radar. À cet égard, les commentaires de 40 participants (20 auditeurs et 20 malentendants) ont été recueillis sur les préférences, l'immersion (avec le questionnaire ipq) et les schémas de mouvements de la tête. Les résultats montrent que les sous-titres qui sont toujours visibles avec des flèches sont l'option préférée. En outre, ces solutions ont obtenu de meilleurs résultats dans le questionnaire ipq. Les modèles de mouvements de la tête montrent que les participants se sont déplacés plus librement avec les sous-titres toujours visibles, ce qui signifie qu'avec ce type de sous-titres, l'expérience est plus réaliste car les utilisateurs ne sont pas gênés par l'intégration des sous-titres.

Mots-clés: sous-titres ; réalité virtuelle ; médias immersifs ; vidéos 360 degrés ; personnes malentendantes; stratégies de sous-titrage.

Introduction

Immersive media such as virtual reality (VR) is attracting the interest of academia and the industry. Content creators are developing 360° videos that provide an immersive and engaging audiovisual experience for their viewers. 360° videos are real-world recordings, captured using special camera sets and post-edited to create a 360° sphere (MacQuarrie & Steed, 2017). The most suitable devices for accessing this type of content are head-mounted displays (HMD). With HMD, the viewer finds him/herself in a 360° sphere in the heart of the action. They have the freedom to look around and explore a virtual world, giving them the sensation that they have been physically transported to a different location. The potential for this technology in industries such as filmmaking, video games, and journalism has become evident in recent years.

The European Broadcasting Union (EBU) issued a report on the use of VR by public broadcasters in 2017. In this report, 49% of respondents stated that they offered 360° video content. Broadcasters see the potential of this medium because the audience can gain a better understanding of the content if they feel a sense of immersion, in particular with items such as the news or the narration of historical events (EBU, 2017). In that sense, *The New York Times* (NYT) and the British Broadcasting Corporation (BBC) have promoted the creation of 360° videos and launched specific platforms through which to display this content. NYT has its own app: the NYT VR (NYT VR, n.d.), accessible on any smartphone with Google Cardboard or any other VR chassis. The BBC 360° content can be accessed via their website (BBC, n.d.) or on their YouTube channel. Other broadcasters, such as CNN or ABC, have also produced immersive videos for sharing their news and events.

Fictional content is limited in the 360° filmmaking industry. The reasons for this could be due to a lack of resources and knowledge when it comes to filmmaking strategies in immersive media, along with certain factors holding back the mainstream

adoption of this technology. Some of these factors include a lack of quality content, a lack of knowledge of VR technology among the general audience, the cost of hardware and the complexity of the user experience, which is still too arduous for the average user (Catapult Digital, 2018). However, the entertainment industry is taking advantage of this technology, with an anticipated growth in video games in the VR sector (Perkins Coie LLP, 2018). The VR technology used in video games is different from the technology used in 360° content creation. In terms of creating fictional content, VR in video games is based on computer-generated images as opposed to real-life images and, therefore, offers more freedom to its content creators, without the need for expensive and complex recording equipment. In turn, video game users tend to be more familiar with adapting to new technologies, hence the success of this technology within the gaming industry.

Overall, it can be stated that immersive content is increasingly present in our society. However, it is still not accessible to all users. To address this issue, the Immersive Accessibility (n.d.) project, funded by the European Union's Horizon 2020 research and innovation program (under grant agreement number 761974), was formed. The aim of this project is to explore how access services such as subtitles for the deaf and hard of hearing (SDH), audio description, and sign language interpreting can be implemented in immersive media. The present study is focused on the implementation of subtitles in 360° videos.

Subtitles have become an intrinsic part of audiovisual content; what Díaz-Cintas (2013, 2014) has labeled as the commoditization of subtitling. Nowadays, subtitles can be easily accessed on digital televisions, video-on-demand platforms, video games and so on, with just one click. The number of subtitle consumers is also increasing and their needs and reasons for the use of the service are varied: they are learning a new language, watching television at night and do not want to disturb their children, commuting on a train and do not have

headphones, they have hearing loss and are using subtitles in order to better understand the content, or as non-native speakers, they simply prefer to watch content in its original language but with subtitles to assist.

In recent years, extensive studies on subtitling have been conducted from various perspectives (D'Ydewalle et al., 1987; Díaz-Cintas et al., 2007; Romero-Fresco, 2009; Bartoll & Martínez-Tejerina, 2010; Matamala & Orero, 2010; Perego et al., 2010; Szarkowska et al., 2011; Arnáiz-Uzquiza, 2012; Mangiron, 2013; Romero-Fresco, 2015; Szarkowska et al., 2016). However, research into subtitles in immersive content is scarce and reception studies are needed in order to find solutions for its lack of accessibility. The aim of this study is to gather feedback regarding different subtitle solutions for the main issues encountered when implementing subtitles in immersive content (i.e. position and guiding mechanisms). Along with additional data based on participants' head movements, feedback has been gathered from 40 participants regarding their preferences for the various options available and how they impact a sense of immersion. In this article, a review of previous studies related to subtitling in immersive media is presented. The methodological aspects of the study are then introduced, such as the participants, the experimental design and the materials used. Then results, along with a discussion, and conclusions.

Theoretical Framework

Research on subtitles in immersive media is relatively recent. Researchers in this field have highlighted several of the challenges being faced when designing and implementing subtitles in immersive media (Rothe et al., 2018; Brown et al., 2018). Firstly, the position of the subtitles needs to be defined. Subtitles on a traditional screen are already standardized and usually located at the bottom-center of the screen which is static. However, the field of view (FoV) in 360° content is dynamic and viewers can decide where to look at any time during

the scene. Therefore, the position of the subtitles needs to be carefully defined so as to avoid loss of content during the experience.

The subtitles need to be located in a comfortable field of view (CFoV), that is, a safe area that is guaranteed to be visible for its users. If subtitles overlay the CFoV, then they will be cropped and therefore unintelligible. In addition, a guiding mechanism needs to be included to enhance accessibility. If the speaker is outside the FoV, persons with hearing loss (for whom the audio cues are not always helpful) will need a guiding system which will indicate where to look. As 360° content aims to provide an immersive experience, subtitles must therefore be created in a way that does not disrupt this immersion. Finally, some viewers suffer from VR sickness or dizziness when watching VR content. The design of the subtitles should not worsen this negative effect, and the subtitles should be easy to read.

In the Immersive Accessibility (ImAc) project, some preliminary studies have approached this topic to gather feedback from users before developing a solution (Agulló et al., 2018; Agulló & Matamala, 2019). In a focus group carried out in Spain, participants with hearing loss were asked how they would like to receive subtitles in 360° videos. They agreed that they would like them to be as similar as possible to those shown on traditional screens. They also stated that they would like the subtitles to be bottom-center of their vantage point and always in front of them. The participants also highlighted the importance of using the current Spanish standard for SDH (AENOR, 2003). Regarding directions, participants suggested the inclusion of arrows, text in brackets (to the left and to the right), and a compass or radar to indicate where the speakers are in the scene (Agulló & Matamala, 2019). In a different preliminary study, feedback from a limited number of users was gathered regarding the CFoV and two guiding mechanisms (icons representing arrows and a compass). Results showed that the users preferred the arrows as a guiding mechanism and the largest

font in the CFoV because it was easier to read (Agulló et al., 2018).

Some reception studies have already been conducted regarding subtitles in immersive media (Rothe et al., 2018; Brown et al., 2018). The BBC Research & Development team proposed four behaviors for subtitles in 360° videos based on previous literature and design considerations. The four behaviors are:

1. 120-degree: subtitles are evenly spaced in the sphere, positioned 120° apart;
2. static-follow: subtitles are positioned in front of the viewer, responding immediately to their head movements;
3. lag-follow: the subtitles appear in front of the viewer and are fixed until the viewer rotates 30°, whereby it moves accordingly to its new position;
4. appear: the subtitles appear in front of the viewer and are fixed in that position, even if the viewer moves around the scene (Brown et al., 2017).

The team tested the four different options with 24 hearing participants, using six clips which lasted 1 to 2 minutes. The Static-Follow behavior was preferred by participants because the subtitles were considered easy to locate and gave participants the freedom to move around the scene. Some issues were highlighted, such as obstruction (a black background box was used) and VR sickness (Brown et al., 2018). Rothe et al. (2018) conducted another study with 34 hearing participants comparing two positions: (a) static subtitles, which were always visible in front of the viewer following their head movements; and (b) dynamic subtitles, which were placed near the speaker in a fixed position. Participants did not state a clear preference in this study; however, dynamic subtitles performed better regarding workload, VR sickness, and immersion.

Current solutions carried out by broadcasters such as *NYT* or *BBC* are mainly burnt-in subtitles, spaced evenly in the 360° sphere, every 120°. An example can be watched in the short documentary *The Displaced* by *NYT* (The New York Times, Within et al., 2015). In other media such as video games, subtitles that are always visible in front of the viewer are the most frequently used (Sidenmark et al., 2019). Different solutions are being tested and implemented by content creators, but a consensus on which option works best has not yet been reached.

This study aims to further clarify which subtitles are more suitable for immersive environments and for all kinds of users, including hearing participants and participants with hearing loss. With this aim, the current solutions implemented by the main broadcasters (fixed subtitles located at 120°) are compared to the solutions developed in the ImAc project (always-visible subtitles and guiding mechanisms). Please note that fixed-position subtitles in ImAc terminology are referred to as 120-Degree by Brown et al. (2017). In Rothe et al. (2018), dynamic subtitles are fixed in one position close to the speaker. Therefore, the implementation is different. Always-visible subtitles in ImAc terminology are equivalent to Static-follow by Brown et al. (2017) and to static subtitles by Rothe et al. (2018). This study contains a higher number of participants (40) than in previous research and includes a higher number of participants with hearing loss (20) in the sample. The subtitles were developed following SDH features (AENOR, 2003), and unprecedented research in the area of subtitle studies was made in the testing of guiding mechanisms in 360° videos. An additional contribution of this study is that longer content is used to better measure the preferences and immersion of its participants. In the following sections, the study and the results are presented.

Method

In this section, the methodology of the test is described, including aim and conditions, design

and procedure, materials, and participants. Two pilot tests were carried out before the study to verify that the methodology was suitable for the test: a first pilot with eight participants (two participants with hearing loss and six hearing participants; Agulló et al., 2019) and a second pilot with three participants (one with hearing loss and two hearing) to confirm that the changes from the first pilot were well implemented and that they were suitable for the experiment. The pilots were particularly necessary because of the complexity of the technical setup and the newness of the medium.

Aim and Conditions

The main goal of the experiment was to test position and guiding methods for subtitles in 360° content in terms of participants' preferences, immersion, and head movement patterns, using the CVR Analyzer tool (Rothe et al., 2018). The two conditions for position were as follows: (a) fixed-position subtitles—subtitles attached to three different fixed positions in the sphere, spaced

648



Figure 1 Fixed-position subtitles attached to one position in the sphere in Episode 4 of *Holy Land*, created by Ryot.



Figure 2 Always-visible subtitles attached to the FoV in Episode 5 of *Holy Land*, created by Ryot.

evenly, 120° apart (see Figure 1); (b) always-visible subtitles—always displayed in front of the viewer, following their movements, attached to their camera or CFoV (see Figure 2). The reasons for testing these two conditions were two-fold. Firstly, the results from previous studies (Brown et al., 2018; Rothe et al., 2018) were inconclusive and included a limited number of participants with hearing loss, as they were not the main target group for the study. Secondly, a comparison between the industry developments by broadcasters like *NYT* and the *BBC* and the solution developed in the *ImAc* project was desired in order to corroborate the relevance of the project.

The conditions for the guiding methods are arrows and a radar. For the first method, intermittent arrows appear on the left or right side of the subtitle text depending on the location of the speaker. If the speaker is to the left of the viewer, the arrows appear on the left of the subtitle text; if the speaker is to the right (see Figure 3), the arrows appear on the right of the subtitle text. The arrows only appear when the speaker is outside the FoV. For the second method, a radar appears in a fixed position in the FoV (towards the bottom, on the right-hand side, near the subtitle) with information about the speaker's location (see Figure 4). With the radar, the viewer is located at the center with their gaze directed by three inverted triangles as seen in Figure 4. The speaker is represented by a small colored triangle. The color is assigned depending on the color of the subtitle for character identification purposes. The viewer needs to move their gaze



Figure 3 Subtitle with arrows in *I, Philip* created by ARTE.



Figure 4 Subtitles with radar in *I, Philip*, created by ARTE.

in order to position the small triangle (speaker) in the center of their FoV. In this part of the test, always-visible subtitles were used. These two options were tested per suggestion of participants in a focus group at the beginning of the project (Agulló & Matamala, 2019). Moreover, the radar option is used in other 360° content players, such as *NYT VR*, and the intention was to compare the current solutions with the solutions developed in the ImAc project (arrows).

Design and Procedure

The experiment was carried out in one session divided into two parts. In the first part, the position of the subtitles was tested. In the second part, the guiding methods were tested. A within-subject design was used to test the different conditions. Each participant was asked to watch an acclimation clip along with four other clips (see the Stimuli section). The clips were randomly presented with different conditions (fixed-position, always-visible, arrow and radar). Both the clips (except for the guiding methods part) and conditions were randomized among the participants to avoid a learning effect. The video used for the guiding methods part was a short science fiction movie, called *I, Philip*. The duration of this movie was around 12 minutes and was cut into two parts in order to test the two conditions (arrows and radar). However, the order of the clips was not altered, as they were contained within the narrative of the movie and the participants would not have understood the story. For this reason, only the conditions were randomized.



Figure 5 The facilitator with a participant during the test.

Twenty participants watched each clip/condition combination.

Before conducting the test, ethical clearance was obtained. The methodology went through the ethical committee at Universitat Autònoma de Barcelona, which approved the procedure and consent forms that would be signed by the participants. During the test, participants were first welcomed by the facilitator, followed by an explanation about the ImAc project and the context for the test. The approved consent form was signed by the participants, and they were then asked to complete the demographic questionnaire. Next, the acclimation clip was shown to familiarize the participants with the HMD. After that, the four clips were presented, and they were asked to complete the IPQ questionnaire (see the Questionnaires section) after each clip. Finally, after each part, the participants were asked to reply to a questionnaire on preferences.

The facilitator was present for the entire duration of the test; providing information as required, helping the participants to correctly place the HMD on their heads and assisting with the completion of the questionnaires (see Figure 5). The questionnaires were administered using Google Forms on a laptop that was provided by the facilitator.

Materials

In this section, the different materials used in the experiment, such as technical equipment, stimuli, and measuring tools, will be explained.

Technical setup

An Apache webserver was installed on a PC and was set-up for the evaluation in order to host the player resources and media assets (360° videos and subtitles). A Samsung Gear VR with Samsung Galaxy S7 was used. The videos were accessible via a URL that was directed to the server resources.

Stimuli

Five clips were shown in total. All the clips were presented without sound. The tests were conducted without sound in order to control any external variables, such as the extent of the participants' hearing loss or their level of English (the video voiceover was in English). The participants with hearing loss presented varying levels of hearing loss and different types of hearing aid, creating a condition wherein its impact on immersion could not be controlled. The level of English among the participants with hearing loss and hearing participants was unpredictable and the results could have therefore been affected. The main focus of the study was to determine which subtitles were most suitable for this type of content and for any type of user. Due to current industry requirements and the unprecedented aspect of the research taking place, emphasis was placed on position and guiding mechanisms. The ecological validity was therefore sacrificed in favor of controlling the external variables in order to receive more consistent feedback regarding the suitability of the different subtitles. In other studies involving testing of different subtitle implementations, the audio was also muted (Kurzahls et al., 2017) or manipulated (Rothe et al., 2018).

Subtitles were produced in Spanish, a language spoken and understood by all participants. The font type selected for the tests was the open-source, sans-serif Roboto, developed by Google, because studies on typefaces indicated that sans-serif fonts appear more natural and are more suitable for VR environments (Monotype, 2017). This font was used in the study carried out by the BBC as well (Brown et al., 2018). Regarding font size in VR

environments, rather than being calculated by pixels as it is for traditional screens, it is based on the CFoV (see Figure 6). The largest font was used, that is, the font that took 100% of the CFoV. The use of the largest font was selected because participants in previous pilots stated that they were more comfortable reading larger fonts, ones that encompass the largest proportion of the CFoV (Agulló et al., 2018). A background box was not included because, in previous studies, it was considered obstructive (Brown et al., 2018).

Furthermore, during the pilots, participants did not report the lack of the background box as a negative aspect of the subtitles, nor a hindrance when reading. Subtitles had a maximum of 37 characters per line and non-speech information was indicated between brackets. Colors were used for character identification when necessary, and subtitles for the narrators were in italics, as indicated in the Spanish standard for SDH (AENOR, 2003). All clips began with a 6-second, black screen with the text, "The test is about to start for (selected variable)" so that participants had enough time to accommodate the HDM before the video started.

An acclimation clip was prepared so that participants would become familiar with the HMD and 360° videos. The first video was 1-minute long and did not contain subtitles. It was assumed that most participants did not have extensive experience with the use of HMD and VR content and that the familiarization step was therefore necessary. The replies in the demographic questionnaire confirmed this assumption.

Two clips were used for the first condition (position). The videos were part of the series, *Holy Land*, created by Ryot in collaboration with Jaunt VR (Graham, 2015). Specifically, episodes 4 (4:13) and 5 (04:58) were used. Permission was given by the creators to use the videos in the study. The episodes depict Middle Eastern territories such as Jerusalem or Palestine. A narrator (voice-over) explains historical facts and the socio-cultural political situation. In the videos, different

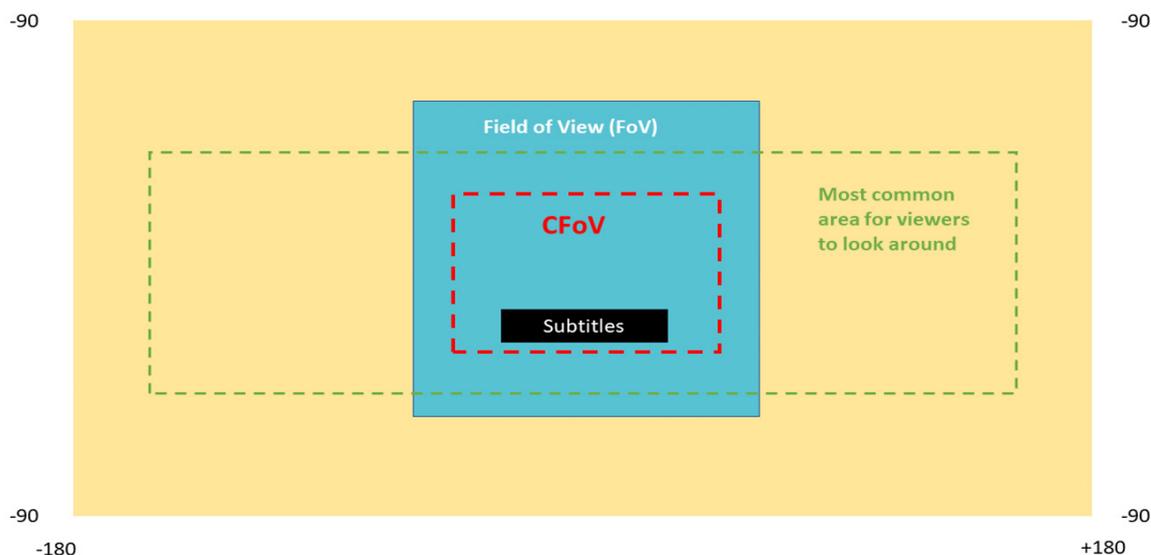


Figure 6 Diagram of 360° video view and its different areas.

locations and landscapes are presented, inviting the viewers to explore the scenario. The hostess appears several times in the scene. However, her presence is not relevant to the narrative of the video. The clips were considered suitable for testing the first condition (position) due to the lack of multiple speakers and a range of landscapes, highlighting the exploratory nature of 360° videos. In this sense, participants could focus on exploring the scenarios in the clips and read the subtitles, without being required to make an effort to understand different speakers or narrative complexities.

For the second condition (guiding mechanisms), the award-winning short movie, *I, Philip*, created by ARTE (Cayrol & Zandrowicz, 2016), was used. The movie runtime was 12 minutes and 26 seconds (excluding the credits) and was split into two parts to test both variables. In this case, permission to use the clip in the study was also granted. The movie narrates the story of David Hanson, who has developed the first-ever android, and is based on the personality and memories of the famous science fiction writer, Philip K. Dick. It is a first-hand experience because the viewer sees the story through the eyes of Phil the robot. This clip was desirable for the test because it involved different

speakers appearing in a variety of locations. This urged the participants to locate the speakers, allowing the different guiding methods to be tested.

Questionnaires

In order to measure the impact of conditions on immersion, the IPQ questionnaire was used. A review of presence questionnaires, such as Slater-Usoh-Steed Presence Questionnaire (Slater & Usoh, 1993), Presence Questionnaire (Witmer & Singer, 1998), ICT-SOPI (Lessiter et al., 2001) and IPQ Questionnaire (Schubert, 2003) was conducted in order to select the one most suitable for the study. IPQ Questionnaire (igroup presence questionnaire, n.d.) was chosen because it is a 14-item questionnaire based on a 7-point Likert scale and includes different subscales that can be analyzed separately if necessary. First, it includes an independent item not belonging to any subscale that reports on general presence. Then, there are three subscales: spatial presence (that is, how much the viewer is feeling physically present in the virtual world), involvement (that is, how much attention the viewer is paying to the virtual world), and realness (that is, to what extent the viewer feels that the virtual world is real). The questionnaire has also been validated in previous studies with similar technology (i.e.

desktop VR, 3D games, VR or CAVE-like systems, etc.). Other questionnaires such as the Presence Questionnaire by Witmer and Singer (1998) include questions on physical interaction with the virtual world, and this was not suitable for our 360° content because the interaction is non-existent. The IPQ questionnaire was translated into Spanish for the purpose of this study.

For each part of the test, an ad-hoc questionnaire in Spanish was created. In regard to position, participants were asked about their preferences, any difficulties faced when reading or locating the subtitles, and whether or not they were obstructive or distracting. Regarding guiding mechanisms, participants were also asked about their preferences, any difficulties experienced when locating the speaker, and whether or not the guiding mechanisms were distracting. Closed and open questions were used, as well as 7-point Likert-scales. To ensure that the wording of the questions was suitable, all questionnaires were reviewed by a professional educational psychologist who specializes in teaching oral language to persons with hearing loss.

Head tracking-CVR Analyzer

Rothe et al. (2018) developed a tracking tool for head and eye movements in 360° content. According to the authors,

[...] the CVR-Analyzer, which can be used for inspecting head pose and eye tracking data of viewers experiencing CVR movies. The visualized data are displayed on a flattened projection of the movie as flexibly controlled augmenting annotations, such as tracks or heatmaps, synchronously with the time code of the movie and allow inspecting and comparing the users' viewing behavior in different use cases. (Rothe et al., 2018, p. 127).

This technology was used in the study to track the head position/orientation of the participants, with the aim of verifying whether a movement pattern could be linked to each specific variable in order to triangulate results with the data from the questionnaires. The results will be discussed in the corresponding section.

Participants

Forty participants (14 male and 26 female), aged 18 to 70 ($M = 37.4$, $Mdn = 37$; $SD = 14.8$), took part in the test. It was decided that the sample should be balanced, including participants ranging from 18 to 60 years of age or more. Participants from different generations are likely to have different levels of technical knowledge as well as different habits when it comes to consuming subtitled content. Therefore, representation from different age groups was obtained.

Twenty-seven participants defined themselves as hearing, six as hearing-impaired and seven as deaf. Twenty participants were identified as persons with hearing loss, which can be confirmed by the responses to the questionnaire item, "Age in which your disability began" (answered by 20 participants with hearing loss). However, the question about defining themselves was more subjective; some people with hearing aids considered themselves to be hearing. In two cases, the users' disability started from birth, for one user it began from 0 to 4 years old; for five users, from 5 to 12 years old; for five users, from 13 to 20 years old; for five users, from 21 to 40 years old and for two users, from 41 to 60 years old. The hearing participants were recruited via personal contacts, and the participants with hearing loss were recruited via the association APANAH (n.d.). The decision was made to include both hearing users and users with hearing loss because, as explained in the theoretical framework, subtitles are beneficial for a wide range of viewers (persons with hearing loss, non-native speakers, persons who are in noisy environments, etc.). In this regard, if part of the spectrum of subtitled content users had been excluded, the study would have been incomplete. Concerning language, all of the participants were Spanish native speakers. It was decided that participants with hearing loss as well as hearing participants should be users of the Spanish oral language in order to have a sample that was as homogeneous as possible in terms of language proficiency.

Regarding use of technology, the devices found to be used most frequently on a daily basis were mobile phone (38), TV (31), laptop (20), PC (13), tablet (11), game console (4) and radio (1). Most users (30) had never watched VR content before. The reasons behind a lack of prior experience with VR content were: they are not interested (3), they have not had the chance to use it (23), it is not accessible (3), or other reasons (4). Ten participants had previously experienced VR content. When directly asked if they were interested in VR content some users (14) were strongly interested, some (14) were interested, some (11) were indifferent, and one was not interested. Five participants owned a device to access VR content (Plastic VR Headsets for smartphone and PlayStation VR).

In terms of content preferences, most users liked news (29), fiction (36), and documentaries (33); while some also liked talk shows (14), sports (17), and cartoons (18). The users activate subtitles depending on the content and their individual needs. Thirteen of them stated that they do not need subtitles, 15 of them stated that it depends on the content, and five of them stated that they always activate subtitles. Seven participants stated other reasons for not using subtitles: bad quality, fast reading speed, issues activating subtitles, subtitles are distracting, and difficulty reading the subtitles and simultaneously viewing the images. One participant also pointed out that they only use subtitles when the hearing aid is not available. Regarding how many hours a day they watch subtitled content, 13 participants indicated zero; 11, less than one hour; seven, from one to two hours; eight, from two to three hours, and one participant indicated four hours or more. When asked why they used subtitles (more than one reply was permitted), 22 indicated that they helped them understand, three participants indicated that it is the only way to access the dialogue, 11 participants said that they use them for language learning, and 11 participants stated that they never use subtitles.

Results

The results from the test are reported in this section, which will discuss the results found for demographics, preferences, presence, and head tracking patterns.

Demographics

In this section, some correlations found in the demographic questionnaire are reported. The Spearman test was used for calculating correlations.

Different correlations were found with age and the use of certain devices. Younger participants used the laptop ($r = -.360$, $p = .023$) and the game console ($r = -.350$, $p = .027$) more than older participants. From the four participants that played games on a daily basis, three had previous experience using VR which could indicate that VR content is mainly consumed in gaming platforms. Regarding the previous experience with VR contents, it was found that younger participants ($M = 27.8$) had previous experience ($r = -.503$, $p = .001$). A correlation was found regarding interest in VR content and the type of participant. Participants with hearing loss were more interested in VR content ($M = 4.3$) than hearing participants ($M = 3.75$) ($r = -.314$, $p = .049$). This was not related to age (participants with hearing loss age $M = 42.55$; hearing participants age $M = 37.55$).

Of the 27 participants consuming subtitled content on a daily basis, 12 had hearing loss and 15 were hearing. Of the 13 participants that did not consume subtitles on a daily basis, eight were persons with hearing loss and five were hearing. A correlation was found between age and the number of hours of subtitled content watched by participants. The younger the participant, the more subtitled content they watched on a daily basis ($r = -.592$, $p = .000$). Participants who watched subtitled content every day presented a mean of 34.2 years old and participants who never watched subtitled content had a mean of 52.2. Also, there was a correlation between the use of a laptop and the subtitled content that participants

watched on a daily basis ($r = .374, p = .018$). From the 20 participants using a laptop every day, 17 watched subtitled content on a daily basis. From the 20 participants who indicated that they did not use the laptop every day, 10 stated that they never watched subtitled content.

Preferences

In this section, results from the preferences questionnaires in the two parts of the test are reported. Statistical comparisons were calculated using a Wilcoxon signed-rank test. Correlations were calculated using the Spearman test and are reported when statistically relevant.

Always-Visible vs. Fixed-Position

Thirty-three participants preferred the always-visible subtitles and seven participants preferred the fixed-position subtitles. According to the participants, with always-visible subtitles they felt more freedom to explore and could comfortably access the content of the subtitles and video scenes without missing details. Participants in favor of fixed-position subtitles stated that they could read them more easily when comparing them to always-visible subtitles. Moreover, they attested that the dizziness effect was minimized with fixed-position subtitles. Some users suggested making the always-visible subtitles more static and less bumpy because the movement hindered their readability and was sometimes distracting.

A difference was found regarding the ease of locating the subtitles in the videos based on a 7-point Likert scale (7 being “the easiest”, 1 being “the most difficult”). Always-visible subtitles ($M = 6.32; SD = 1.5$) were considered easier to find than fixed-position subtitles ($M = 4.25; SD = 1.61$). This difference is statistically significant ($Z = -3.986, p = .000, ties = 5$). A correlation was found with the demographic profile: participants with hearing loss ($M = 4.85$) considered the fixed-position subtitles easier to find than the hearing participants ($M = 3.65$) ($r = -.368, p = .019$). Again, a

difference was encountered when asking about the ease of reading the subtitles in the videos based on a 7-point Likert scale (7 being “the easiest”, 1 being “the most difficult”). Always-visible subtitles ($M = 5.72; SD = 1.88$) were considered easier to read than fixed-position subtitles ($M = 4.77; SD = 1.91$). However, this difference is not statistically significant ($Z = -1.919, p = .055, ties = 9$).

When asked whether always-visible subtitles were obstructing important parts of the image based on a 7-point Likert scale (7 being “not at all”, 1 being “yes, very much”), participants felt that fixed-position subtitles ($M = 5.87; SD = 1.28$) were slightly less obstructive than the always-visible subtitles ($M = 5.5; SD = 1.9$). However, the difference is not statistically significant ($Z = -1.123, p = .261, ties = 23$). Two correlations were found. Firstly, participants with hearing loss ($M = 6.2$) considered always-visible subtitles to be less obstructive than hearing participants ($M = 4.8$) ($r = -.452, p = .003$). Secondly, participants with hearing loss ($M = 6.5$) also considered fixed-position subtitles less obstructive than the hearing participants ($M = 5.25$) ($r = -.552, p = .000$). A difference was found when participants were asked if the subtitles were distracting them from what was happening in the video based on a 7-point Likert scale (7 being “not at all”, 1 being “yes, very much”). The always-visible subtitles ($M = 5.4; SD = 1.98$) were considered to be less distracting than the fixed-position subtitles ($M = 3.77; SD = 2.33$). This difference is statistically significant ($Z = -2.696, p = .007, ties = 13$). Another correlation was found here: Participants with hearing loss ($M = 4.8$) considered the fixed-position subtitles to be less distracting than the hearing participants ($M = 2.75$) ($r = -.397, p = .011$). In Table 1, a summary of the results for preference regarding always-visible and fixed-position subtitles can be found.

Arrow vs. Radar

Thirty-three participants preferred arrows and seven participants preferred the radar as a guiding

Table 1 Summary of Results (Mean) in Closed Questions for the Preference Questionnaire Comparing Always-Visible and Fixed-Position Subtitles

	Easy to find (7 "the easiest", 1 "the most difficult")	Easy to read (7 "the easiest", 1 "the most difficult")	Level of obstruction (7 "not at all", 1 "yes, very much")	Level of distraction (7 "not at all", 1 "yes, very much")
Always-visible	6.32	5.72	5.5	5.4
Fixed-position	4.25	4.77	5.87	3.77

method. More participants with hearing loss (six out of seven) preferred the radar as a guiding mechanism. Participants who favored the arrows argued that this guiding method is more intuitive, direct, comfortable, less invasive, and less distracting. Furthermore, during the test, several participants reported that they did not understand the radar. Participants who preferred the radar argued that it provides more spatially accurate information. One participant suggested that a short explanation about how the radar works before using it could improve the user perception of this mechanism. Two participants suggested moving the radar from the right-hand side to the center (maybe directly below the subtitle) in order to improve usability. Four participants, despite having chosen one method or the other, stated that they would prefer to not use guiding mechanisms and that the color identification was sufficient.

A difference in the results was found when asking about how easy it was to find the speaker using the arrow guiding method based on a 7-point Likert scale (7 being "the easiest", 1 being "the most difficult"; see Table 2). Participants felt that the speaker was easier to find with the arrows ($M = 6.12$; $SD = 1.26$) than with the radar ($M = 3.9$; $SD = 2.24$). The difference is statistically

Table 2 Summary of Results (Mean) in Closed Questions for the Preference Questionnaire Comparing Subtitles with Arrows and Radar

	Easy to find the speaker (7 "the easiest", 1 "the most difficult")	Level of distraction (7 "not at all", 1 "yes, very much")
Arrows	6.12	6.42
Radar	3.9	3.9

significant ($Z = -4.166$, $p = .000$, ties = 10). A difference was also found regarding the level of distraction with each method based on a 7-point Likert scale (7 being "not at all", 1 being "yes, very much"). Arrows ($M = 6.42$; $SD = 1.03$) were considered to be less distracting than radar ($M = 3.9$; $SD = 2.46$). The difference is statistically significant ($Z = 4.125$, $p = .000$, ties = 12). A correlation was found here: The four participants that played video games on a daily basis found the radar less distracting ($M = 1$) than the rest of the participants ($M = 4.22$) ($r = -.460$, $p = .003$).

Presence

In this section, results from the IPQ questionnaires for each video are analyzed and compared. Statistical comparisons were calculated using a Wilcoxon signed-rank test.

Always-visible vs fixed-position

A comparison between always-visible and fixed-position subtitles was made in the IPQ, and the results are as follows (see Table 3): spatial scale ($Z = -1.791$, $p = .073$, ties = 7), involvement scale ($Z = -1.229$, $p = .219$, ties = 8), realness scale ($Z = -.064$, $p = .949$, ties = 14). The test indicated that the differences between the results were not statistically significant. However, for the general presence item, the test indicated that the difference between the results is statistically significant ($Z = -2.694$, $p = .007$, ties = 17). This means that the fixed-position subtitles had a negative impact on the presence of the participants.

Arrow vs radar

A comparison between arrow and radar methods was made in the IPQ and the results are as follows

Table 3 IPQ Results (Mean) Comparing Always-Visible and Fixed-Position Subtitles

	General Presence	Spatial Presence	Involvement	Experienced Realism
Always-visible	4.7	3.74	3.31	2.36
Fixed-position	3.95	3.43	3.38	2.49

(see Table 4): general presence item ($Z=-1.852$, $p=.064$, ties = 21), spatial scale ($Z=-1.000$, $p=.317$, ties = 13), realness scale ($Z=-1.430$, $p=.153$, ties = 7). The test indicated that the differences between the results were not statistically significant. However, for the involvement scale, the test indicated that the difference between the results is statistically significant ($Z=-2.138$, $p=.033$, ties = 12). In this case, the radar had a negative impact on the involvement of the participants.

Head tracking

656

The qualitative results from the CVR Analyzer shed some light on the participants' viewing patterns. The data gathered from this tool is visual, using heat maps to represent the participants' head movements. A difference in viewing patterns was observed between the always-visible and fixed-position subtitles. With the always-visible subtitles, the head movements were more scattered, showing that participants explored the scenes with more freedom (see Figures 8 and 10). In the case of fixed-position subtitles, it can be observed that the participants were more concentrated on one specific point, that is, the central subtitle (see Figures 7 and 9).

Significant differences were not found when comparing viewing patterns for arrows and the radar. It was observed that participants directed

Table 4 IPQ Results (Mean) Comparing Subtitles with Arrows and Radar

	General Presence	Spatial Presence	Involvement	Experienced Realism
Arrow	4.6	3.66	3.5	2.49
Radar	4.27	3.49	3.27	2.47



Figure 7 The pattern of head movements from *Holy Land* with fixed-position subtitles.



Figure 8 The pattern of head movements from *Holy Land* with always-visible subtitles (exactly the same scene as Figure 7).



Figure 9 The pattern of head movements from *Holy Land* with fixed-position subtitles.



Figure 10 The pattern of head movements from *Holy Land* with always-visible subtitles (exactly the same scene as Figure 9).

their attention to the speakers and their faces (see Figures 11 and 12), as is usual in 2D content (Kruger, 2012). No significant differences were observed between hearing participants and participants with hearing loss for any of the variables.



Figure 11 I, *Philip* scene with arrows.



Figure 12 I, *Philip* scene with radar.

Discussion

Always-visible subtitles are the preferred option by 82.5% of participants. Subtitles in this position were considered to be easier to read and less distracting than fixed-position subtitles. Participants stated that with these subtitles, they had more freedom to look around the scenes without feeling restricted. In general, always-visible subtitles performed better in the IPQ questionnaire whereas fixed-position subtitles had a negative impact on presence. According to the comments made by participants in the open questions, this could be because they felt less free to explore the 360° scene and claimed to have missed parts of the subtitle content. Moreover, as reported above, participants encountered more difficulty when locating and reading the subtitles

in this mode, and also considered them to be more distracting. This extra effort could therefore have caused a negative impact on the feeling of being present in the scene.

Head tracking data also supported the qualitative data, with the participants showing more exploratory behavior in the always-visible condition compared to the fixed-position condition. Results from the head-tracking tool were similar to the results for the study by Rothe et al (2018). They also found that with always-visible subtitles, the head movements are more scattered. In this study, unlike for others before it (Brown et al., 2018), font with an outline was used instead of a background box which resulted in less obstructive subtitles. However, the readability of the subtitles can still be hindered, due to, and depending on, the video background. It would be interesting to carry out further research using artificial intelligence technology that could potentially add a background box automatically when the contrast is poor, as explained by Sidenmark et al. (2019). Also, always-visible subtitles were considered slightly more obstructive than fixed-position. Further studies could be carried out to test different font sizes or different positions (for example, lower in the FoV, or above the FoV).

The arrows were the preferred option for guiding mechanisms with 82.5% of participants favoring it over radar. Most people considered them to be more intuitive and comfortable than the radar, which required an extra cognitive effort (understanding how it works), something that was not called for in the arrows condition. The radar was also considered more distracting than the arrows. In this regard, it can be asserted that in 360° content, the simpler solution is preferred because moving and interacting with the virtual world is already complex for the viewers and the cognitive workload should therefore be reduced. The sample of participants who preferred the radar happened to also be those who played video games on a daily basis (four participants, three of whom were participants with hearing loss). Previous experience with

video games and the use of a radar for spatial orientation might have an impact on the users' preferences in terms of guiding mechanisms for VR content.

It was found that the radar had a negative impact on the involvement of the participants. Participants reported that the radar was difficult to understand, and some of them felt frustrated by it when watching the stimuli at the beginning. At certain points during the story, the participants were busy trying to understand how it worked as opposed to following the plot, which, therefore, led to a loss of content. As stated before, one of the prerequisites for implementing subtitles in immersive content is that they do not disrupt the immersive experience. In this case, the complexity of the radar interfered with users' immersion. Some participants claimed that they would prefer no-guiding mechanism at all. If 360° videos are post-produced to a high standard from a filmmaking perspective, it is possible that there will be no need for guiding mechanisms because when changing from one scene to another, speakers will be located in the center of the FoV. Therefore, non-intrusive mechanisms, such as arrows, might be more advisable for that type of content. Guiding methods in this medium such as arrows and radar could be used to improve not only subtitles but also narrative. For example, radar or arrows could indicate that the action or point of interest is located in a specific location outside the FoV. In general, the most optimal solution can vary depending on the type of content.

Despite the fact that it was not the aim of the study, replies from the demographic questionnaire also shed some light on the profile of subtitled content users. Hearing participants (15) consumed more subtitled content than participants with hearing loss (12) on a daily basis. Moreover, younger participants ($M = 34.2$) watched subtitled content every day, and a correlation was found between the use of a laptop and the consumption of subtitled content ($r = .374$, $p = .018$). Younger viewers (hearing or with hearing loss), who use advanced technology and devices on a daily basis are more active users of subtitles than older users, who

mainly consume audiovisual content on TV sets. This indicates that subtitle consumption habits are evolving. Consumption habits do not necessarily depend on the level of hearing loss but on age or familiarity with the technology (Ulanoff, 2019). This might be for consideration when profiling participants for future research in subtitle studies.

Conclusions

In this article, the potential of immersive content and the need to make it more accessible with subtitles has been highlighted. A review of the main studies regarding subtitles in immersive content has been presented, showing that further research is needed. The study has been presented as well as the results. The primary aim was to compare different subtitling solutions in two main areas of 360° content: position and guiding mechanisms. Feedback on preferences, immersion and head movements from 40 participants was gathered to clarify the open questions. Results have shown that always-visible subtitles, that is, subtitles that are always in front of the viewers, are the preferred option and, in turn, more immersive. Regarding guiding mechanisms, arrows were preferred over radar due to their simple and efficient design that was easier to understand.

The present study also had some limitations regarding the type of content. For each variable, a specific type of content was used: a documentary and a short science-fiction movie. This might have had an impact on the results and a replication of this study with different content is encouraged. However, the agreement of participants on the preferred option was clear. It can be concluded from the results that the subtitles that best perform in terms of usability and immersion are always-visible subtitles with arrows. From a production point of view, this type of subtitling is more scalable and easier to produce than burnt-in subtitles created manually. A 360° subtitle editor is being developed in the ImAc project for this purpose (Agulló, in press).

Always-visible subtitles with arrows could be considered as good practice guidelines for subtitles in 360° videos, especially for exploratory content such as documentaries or fast action and interactive content such as video games. However, results from the present study are not final and further research is encouraged. Depending on the content, subtitles located close to the speaker in a fixed position could perform better. They could also be considered more aesthetically suitable and better integrated within the scenes, as reported by Rothe et al. (2018). Fixed-position subtitles could be created manually in the creative part of the production of the film, as suggested in the accessible filmmaking theory by Romero-Fresco (2019). Even a combination of both methods (always-visible and fixed-position subtitles) could work, as suggested by Rothe et al. (2018).

For exploratory content or fast action, fixed-position subtitles might not be the most suitable because fixed-position subtitles could hinder the immersive experience. Finally, studies into the reading speed of subtitles in immersive content are required as user interaction with content in this new medium is different from that of traditional screens, and the current rules, therefore, do not necessarily apply. The processing of subtitles may be affected when exploring VR scenes with an HMD, as it may entail a heavier cognitive workload compared to traditional media. Further studies with eye-tracking could clarify current and future questions.

Subtitles are an important part of audiovisual media and beneficial to a wide range of viewers. Virtual and augmented reality applications in our society are numerous, so further studies are necessary if user experience and access to immersive content are to be improved. In this study, it has been proven that subtitles can be implemented in this type of content, and the preferred solution (always-visible with arrows) has been provided.

Acknowledgments

Authors would like to thank ImAc project partners, especially i2Cat members, Mario Montagud

and Isaac Fraile, for their technical contributions; to the content creators, Ryot and ARTE, for giving us permission to use their fantastic 360° content; also, to APANAH for their inestimable help carrying out the tests (thanks to Anto, Rafa, Isabel, Graciela, Vicente, and everyone); thanks to Sylvia Rothe for sharing her expertise and for the CVR tool; thanks to Irene Tor-Carroggio and José Ramón Calvo-Ferrer for their help with the statistical analysis; and, especially, to all participants, without their involvement, this study would not have been possible.

References

- Agulló, B. (in press). Technology for subtitling: a 360-degree turn. *Hermeneus*, 22.
- Agulló, B., Matamala, A., & Orero, P. (2018). From disabilities to capabilities: testing subtitles in immersive environments with end users. *HIKMA*, 17, 195-220. <https://doi.org/10.21071/hikma.v17i0.11167>
- Agulló, B., & Matamala, A. (2019). Subtitling for the deaf and hard-of-hearing in immersive environments: results from a focus group. *The Journal of Specialised Translation*, 32, 217-235. Retrieved from http://www.jostrans.org/issue32/art_agullo.pdf
- Agulló, B., Montagud, M., & Fraile, I. (2019). Making interaction with virtual reality accessible: rendering and guiding methods for subtitles. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 33(4), 416-428. <https://doi.org/10.1017/S0890060419000362>
- APANAH (n.d.). Retrieved from <https://www.apanah.com/>
- Arnáiz-Uzquiza, V. (2012). *Subtitling for the deaf and the hard-of-hearing: Some parameters and their evaluation* [Doctoral thesis, Universitat Autònoma de Barcelona, Barcelona, Spain]. <https://www.tdx.cat/handle/10803/117528>
- Asociación Española de Normalización y Certificación (AENOR). (2003). *Norma UNE 153010: Subtitulado para personas sordas y personas con discapacidad auditiva. Subtitulado a través del teletexto*. <https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma/?Tipo=N&c=N0029761>
- Bartoll, E., & Martínez-Tejerina, A. (2010). The positioning of subtitles for the deaf and hard of hearing. In A. Matamala, & P. Orero (Eds.), *Listening to subtitles:*

- Subtitles for the deaf and hard of hearing* (pp. 69-86). Peter Lang.
- BBC (n.d.). Reel World. 360° Videos from the BBC. Retrieved from <https://www.bbc.com/reel/playlist/reel-world-360-videos-from-the-bbc>
- Brown, A., Turner, J., Patterson, J., Schmitz, A., Armstrong, M., & Glancy, M. (2017). Subtitles in 360-degree video. *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video, Hilversum, Netherlands*, 3-8. <https://doi.org/10.1145/3084289.3089915>
- Brown, A., Turner, J., Patterson, J., Schmitz, A., Armstrong, M., & Glancy, M. (2018). *Exploring subtitle behaviour for 360° video* (White Paper WHP 330) <https://www.bbc.co.uk/rd/publications/whitepaper330>
- Catapult Digital. (2018). *Growing VR/AR companies in the UK: A business and legal handbook*. <https://www.pwc.co.uk/intelligent-digital/vr/growing-vr-ar-companies-in-the-uk.pdf>
- Cayrol, A. (Producer), & Zandrowicz, P. (Director). (2016). *I, Philip* [Video file] Retrieved from <https://www.artetv/sites/webproductions/en/i-philip/>
- D'Ydewalle, G., Pollet, J., & van Rensbergen, J. (1987). Reading a message when the same message is available auditorily in another language: the case of subtitling. In J. K. O'Regan, & A. Lévy-Schoen (Eds.), *Eye movements: From physiology to cognition*. (pp. 313-321). Elsevier Science Publishers. <https://doi.org/10.1016/B978-0-444-70113-8.50047-3>
- Díaz-Cintas, J. (2013). The technology turn in subtitling. In M. Thelen, & B. Lewandowska-Tomaszczyk (Eds.), *Translation and meaning: Part 9* (pp. 119-132). Zuyd University of Applied Sciences.
- Díaz-Cintas, J. (2014). Technological strides in subtitling. In Chan Sin-wai (Ed.), *The Routledge encyclopedia of translation technology*, (pp. 632-643). Routledge.
- Díaz-Cintas, J., Orero, P., & Remael, A. (Eds.), (2007). *Media for all: Subtitling for the deaf, audio description, and sign language*. Rodopi. <https://doi.org/10.1163/9789401209564>
- European Broadcasting Union (EBU). (2017). *Virtual reality: How are public broadcasters using it?* <https://www.ebu.ch/publications/virtual-reality-how-are-public-broadcasters-using-it>
- Graham (2015). *Jaunt And RYOT Announce VR Documentary Series Holy Land*. Retrieved from <https://www.vrfocus.com/2015/12/jaunt-and-ryot-announce-vr-documentary-series-holy-land/>
- Igroup presence questionnaire (n.d.). Igroup Presence questionnaire overview [html]. Retrieved from <http://www.igroup.org/pq/ipq/index.php>
- Immersive Accessibility website (n.d.). Presentation [html]. Retrieved from <http://www.imac-project.eu/>
- Kruger, J.L. (2012). Making meaning in AVT: eye tracking and viewer construction of narrative. *Perspectives: Studies in Translation Theory and Practice*, 20(1), 67-86. <https://doi.org/10.1080/0907676X.2011.632688>
- Kurzahls, K., Cetinkaya, E., Hu, Y., Wang, W., & Weiskopf, D. (2017). Close to the action: Eye-Tracking evaluation of speaker-following subtitles. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, USA* (pp. 6559-6568). <https://doi.org/10.1145/3025453.3025772>
- Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators, and Virtual Environments* 10(3), 282-297. <https://doi.org/10.1162/105474601300343612>
- MacQuarrie, A., & Steed, A. (2017). Cinematic virtual reality: Evaluating the effect of display type on the viewing experience for panoramic video. *Proceedings of 2017 IEEE Virtual Reality (VR), USA* (pp. 45-54). <https://doi.org/10.1109/VR.2017.7892230>
- Mangiron, C. (2013). Subtitling in game localisation: a descriptive study. *Perspectives: Studies in Translation Theory and Practice*, 21(1), 42-56. <https://doi.org/10.1080/0907676X.2012.722653>
- Matamala, A., & Orero, P. (Eds.) (2010). *Listening to subtitles: Subtitles for the deaf and hard of hearing*. Peter Lang.
- Monotype (2017). *The virtual frontier* [blog entry]. Retrieved from <https://medium.com/@monotype/the-virtual-frontier-8f05bf20e92d>
- NYT VR (n.d.). Virtual reality [blog entry]. Retrieved from <http://www.nytimes.com/marketing/nytvr/>
- NYT VR Player (n.d.). Virtual reality player [blog entry]. Retrieved from <https://www.nytimes.com/video/360-video>
- Perego, E., Del Missier, F., Porta, M., & Mosconi, M. (2010). The cognitive effectiveness of subtitle processing. *Media Psychology*, 13(3), 243-272. <https://doi.org/10.1080/15213269.2010.502873>
- Perkins Coie LLP (2018). *2018 augmented and virtual reality survey report: Insights into the future of AR/VR*. <https://www.perkinscoie.com/images/content/1/8/v2/187785/2018-VR-AR-Survey-Digital.pdf>

- Romero-Fresco, P. (2009). More haste less speed: edited vs. Verbatim respeaking. *Vigo International Journal of Applied Linguistics*, (6), 109-133. <https://dialnet.unirioja.es/servlet/articulo?codigo=3207248>
- Romero-Fresco, P. (Ed.). (2015). *The reception of subtitles for the deaf and hard of hearing in Europe*. Peter Lang. <https://doi.org/10.3726/978-3-0351-0888-0>
- Romero-Fresco, P. (2019). *Accessible filmmaking: Integrating translation and accessibility into the filmmaking process*. Routledge. <https://doi.org/10.4324/9780429053771>
- Rothe, S., Höllerer, T., & Hussmann, H. (2018). CVR-analyzer: a tool for analyzing cinematic virtual reality viewing patterns. *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia, Cairo, Egypt*, 127-137. <https://doi.org/10.1145/3267782.3274688>
- Rothe, S., Tran, K., & Hussmann, H. (2018). Dynamic subtitles in cinematic virtual reality. *ACM TVX 2018, Seoul, Republic of Korea*, 209-214. <https://doi.org/10.1145/3210825.3213556>
- Schubert, T. (2003). The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Zeitschrift für Medienpsychologie*, 15(2), 69-71. <https://doi.org/10.1026//1617-6383.15.2.69>
- Sidenmark, L., Kiefer, N., & Gellersen, H. (2019). Subtitles in interactive virtual reality: Using gaze to address depth conflicts. *Proceedings of Workshop on Emerging Novel Input Devices and Interaction Techniques*. Osaka, Japan. <https://eprints.lancs.ac.uk/id/eprint/132411/>
- Slater, M., & Usoh, M. (1993). Representations systems, perceptual position, and presence in immersive virtual environments. *Presence* 2(3), 221-233. <https://doi.org/10.1162/pres.1993.2.3.221>
- Szarkowska, A., Krejtz, I., Kłyszczko, Z., & Wiczorek, A. (2011). Verbatim, standard, or edited? Reading patterns of different captioning styles among deaf, hard of hearing, and hearing viewers. *American Annals of the Deaf*, 156(4), 363-378. <https://doi.org/10.1353/aad.2011.0039>
- Szarkowska, A., Krejtz, I., Pilipczuk, Dutka, L., & Kruger, J. (2016). The effects of text editing and subtitle presentation rate on the comprehension and reading patterns of interlingual and intralingual subtitles among deaf, hard of hearing and hearing viewers. *Across Languages and Cultures*, 17(2), 183-204. <https://doi.org/10.1556/084.2016.17.2.3>
- The New York Times & Within (Producers), & Solomon, B. C. & Ismail, I. (2015). *The Displaced* [Video file]. Retrieved from: <https://www.nytimes.com/video/magazine/100000005005806/the-displaced.html>
- Ulanoff, L. (2019, January 17). Why Gen Z loves closed captioning [Web log post]. <https://onezero.medium.com/why-gen-z-loves-closed-captioning-ec4e44b8d02f>
- Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence Teleoperators & Virtual Environments*, 7(3), 225-240. <https://doi.org/10.1162/105474698565686>

How to cite this article: Agulló, B., and Matamala, A. (2020). Subtitles in virtual reality: Guidelines for the integration of subtitles in 360° content. *Íkala, Revista de Lenguaje y Cultura*, 25(3), 643-661. <https://doi.org/10.17533/udea.ikala.v25n03a03>