

ISSN: 1579-9794

Production of access services in immersive content: understanding the needs of audio describers

ANITA FIDYKA,
ANNA MATAMALA
anita.fidyka@uab.cat
anna.matamala@uab.cat
Universitat Autònoma de Barcelona

Fecha de recepción: 20 de febrero de 2019

Fecha de aceptación: 26 de julio de 2019

Resumen: Los contenidos en 360°, ofrecidos cada vez con más frecuencia por diversas emisoras e instituciones culturales, deberían satisfacer las necesidades de todos los miembros de nuestra sociedad, incluidas las personas con pérdida de visión. Sin embargo, hasta ahora se ha investigado poco sobre la audiodescripción (AD) de dichos contenidos. Este artículo presenta los resultados de un estudio de usabilidad del prototipo de editor de AD que se ha desarrollado en el proyecto Immersive Accessibility (ImAc). También pone de relieve las necesidades de los audiodescriptores profesionales cuando trabajan con contenidos de 360°. El editor es una herramienta en línea que permite a los audiodescriptores elegir diferentes tipos de sonido para la AD y colocar los segmentos de la AD en la esfera de 360°. El estudio se llevó a cabo en línea y los datos se recopilaban a través de un prequestionario demográfico y un posquestionario que incluía el SUS (System Usability Scale) y preguntas sobre preferencias. Los resultados obtenidos proporcionan información valiosa sobre cómo mejorar las funcionalidades de la herramienta. También detectan la necesidad de disponer de directrices sobre la selección de información en este formato, lo que sugiere que la AD en contenido inmersivo podría integrarse en estándares o recomendaciones de AD o en cursos de formación. Los resultados de esta investigación son solo un punto de partida en el campo de la accesibilidad inmersiva, de ahí la recomendación de seguir investigando sobre accesibilidad en este formato.

Palabras clave: Traducción audiovisual, Accesibilidad en los medios, audiodescripción, vídeos de 360°, usabilidad

Abstract: 360° content, offered more and more frequently by various broadcasters and culture institutions, should cater for the needs of all members of our society, including persons with sight loss. So far, however,

the question of providing audio description (AD) in such content has been under researched. This study aims to report the results of the usability study of the prototype AD editor developed within the Immersive Accessibility (ImAc) project, which allowed us to gain insights into the needs of professional audio describers when working with 360° content. The editor is an online tool which allows describers to choose an appropriate sound type for AD, and place AD segments in the 360° sphere. The study was conducted online and data was collected by means of a demographic pre-questionnaire and a post-questionnaire, consisting of a System Usability Scale and additional preference questions. The results obtained provide valuable feedback on how to improve the functionality of the tool to meet the needs of its users. They also indicate the need for guidance when selecting content to be described in this media format, which suggests that AD in immersive content could be integrated into AD guidelines or specific courses offered by training institutions. The results of this study are just a starting point in the field of immersive accessibility, hence the recommendation for further research on the subject of accessibility in this media format.

Keywords: Audiovisual Translation, Media Accessibility, audio description, 360° videos, usability

INTRODUCTION

Audio description (AD) is an access service that is considered as a way of retelling a story: it translates the visual channel to a verbal mode (Maszerowska, Matamala, & Orero, 2014; Fryer, 2016; Snyder, 2008). It is used widely to render standard, two-dimensional audiovisual (AVT) products, such as films and TV programs, accessible. It is also used to describe, for instance, artworks in museums or live events, rendering cultural property accessible for those who cannot access the visuals, and to other groups at risk of social exclusion (Greco, 2016). So far, AD has been mainly provided for two-dimensional content (Fidyka and Matamala, 2018a). Although more interactive forms of AD exist, such as AD in the theatrical environment (Fryer and Freeman, 15) or AD in planetariums – in which content is displayed in a dome, surrounding the viewer –, research on AD in immersive media is practically non-existent nowadays (Fidyka and Matamala, 2018a). The few studies conducted to date include audio description in 3D cinema (Greening, 2011), or the integration of haptics in audio description (Viswanathan, McDaniel, Krishna, & Panchanathan, 2010). Regarding other access services, the implementation of subtitles has been researched (Agulló, Matamala, & Orero, 2018; Agulló, Matamala, & Orero, In press; Rothe, Tran, & Hußmann, 2018) and some solutions on the implementation have been proposed by

major distributors (Brown, 2017). However, the issue of implementing AD in 360° content has not been addressed so far.

All new technologies appearing on the market should be accessible to ensure that all members of our society have access to culture and arts, as specified in documents on Human Rights, such as the Universal Declaration of Human Rights and the UN Convention on the Rights of Persons with Disabilities. To ensure that such highly visual environments are accessible for persons with sight loss, devices used to consume immersive content (eg. glasses) should be made accessible and access services such as audio description should be provided.

Although immersive media are still emerging on the market, they already have a wide array of creative formats (Allen and Tucker, 2018). One of them is 360° videos, belonging to Virtual Reality (VR). These videos are typically between 5 and 15 minutes long, and they are «the most tightly authored among all VR formats» (Allen and Tucker, 2018:17) that have particular market potential. It means that the story – which is told with a central protagonist's «rise, fall and resolution» (ibid.) – is linear, often driven by dialogue and pre-scripted by a director or content creator (ibid.).

In 360° videos, also referred to as omnidirectional or spherical videos, the main challenge lies in storytelling techniques, which are still being defined. Sighted users can access 360° content by means of head-mounted displays, i.e. special glasses to access the Virtual Reality content, which allow them to feel as if they are inside a sphere, while the linear story pre-scripted by a director unfolds all around them (Fidyka and Matamala, 2018b; Dooley, 2017). As different events can occur at different angles of the visual scene, guiding the sighted users inside the virtual world towards the main action is considered challenging (Rothe, Hußmann, & Allary, 2017; Jones, 2017; Syrett, Calvi, & van Gisbergen, 2017), and various storytelling techniques are being researched (Gödde, Gabler, Siegmund, & Braun, 2018; Jerald, 2015). In any case, sighted audiences can look around the sphere, following the main story or ignoring it.

The media format is mostly visually-driven, with images being triggered by head movements. Finding a way to guide persons with sight loss effectively inside the sphere increases the challenge (Fidyka and Matamala, 2018a, 2018b). The technology of spatial sound, which gives audiences a three-dimensional soundscape of the elements presented in the content may prove of value, as well as different scripting styles. The implementation of spatial sound, including ambisonics (Johansson, 2019) in AD provided for 360° videos is being researched within the Immersive Accessibility (ImAc) project.

Its application in AD is also being investigated by other researchers (López, Kearney, & Hofstädter, 2016; Portillo, 2018).

ImAc is an H2020 project based on a user-centred methodology (Suojanen, Koskinen & Tuominen, 2015). Thanks to this methodology, end users and professional audio describers collaborate with project partners at every stage of the project. In its early stages, feedback from users was gathered through a series of studies based on focus groups (Fidyka and Matamala 2018a, 2018b). These qualitative studies allowed researchers to define the implementation of access services and editing tools, and their results suggest that spatial sound can serve as a tool to facilitate orientation within the sphere. As far as the production of AD in immersive content is concerned, the development of an editing tool for this access service is one of the aims of the project.

To respond to the challenge of guiding persons with sight loss in this new medium, a new method of producing audio description needed to be proposed prior to the development of the AD web editor. Based on the early feedback from end users, it was decided that when providing AD by means of the web editor, audio describers would be able to choose between three different audio description modes, associated with specific sound and scripting features, as discussed in detail in the next section. This choice will allow describers to choose the most appropriate style for each content (Allen & Tucker, 2018). This new approach also needed to be tested with professional describers, as it impacts on their current workflow.

This article will discuss the methodology and results of the usability tests of the AD web editor in its first prototype version, which allowed us to understand the needs of professional audio describers when producing AD in this media format. To contextualize the study, the main features of the prototype editor, developed by Anglatècnic, including the functions specific to immersive environments, will be described in the next section. Section 2 explains the methodology of the test, and section 3 offers the discussion of its results. Although the article reports on the results of a usability test, the interest lies in the interesting insights provided by professionals on the specific challenges the immersive content poses when audio describing, and how this can impact their current practices.

1. FEATURES OF THE WEB EDITING TOOL

The following section describes different features of the first version of the AD editing tool. Firstly, standard features that can also be found in other existing editors are discussed. In the second part, features specific to the medium under discussion are explained.

1.1. Standard features

The AD editor is an online tool, and is comprised of different sections. In the central section of the editor, audio describers can edit the AD script, add timecodes and preview the video with AD, as shown on Figure 1.

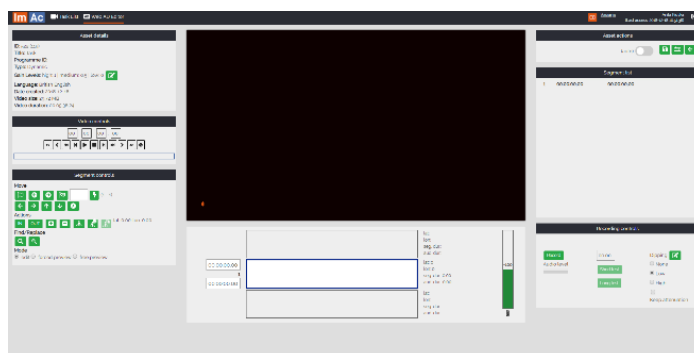


Figure 1 General view of the AD web editor

Another section, «Asset details», displays the basic information about the video, such as its name, size and language. Also, specific messages appear in that section in case of errors. Figure 2 shows the «Video controls», which allow audio describers to play, pause or stop the video, and navigate through it. Video controls have their own shortcuts, which were pre-set for the test, but will be customizable in future versions of the AD editor.

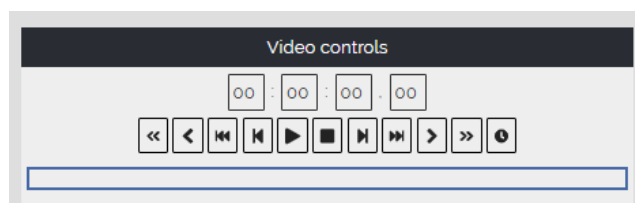


Figure 2 Video controls

The next section, «Segment controls», is related to audio description. It allows audio describers to add, remove and navigate through AD segments as well as add or remove timecodes (Figure 3). It also allows audio describers to set the angle, as described in section 1.2. Similarly to «Video controls», «Segment controls» also have their own shortcuts.

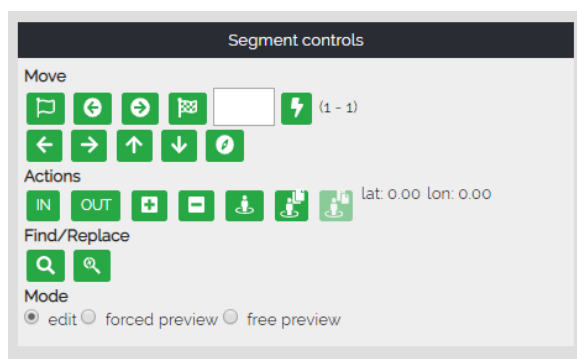


Figure 3 Segment controls section

As already mentioned, audio describers edit the script below the video player (Figure 4). To the left of the script, audio describers can set timecodes for a given AD segment. The number of the segment is also displayed in this section. To the right of the script editing area, the longitude and latitude of the current segment can be found. Also, the duration of the segment is displayed below. In the corner of the right side of this section, reading speed is displayed. When the reading speed is appropriate, the colour is green. When the number of characters per second is too high, it becomes red.

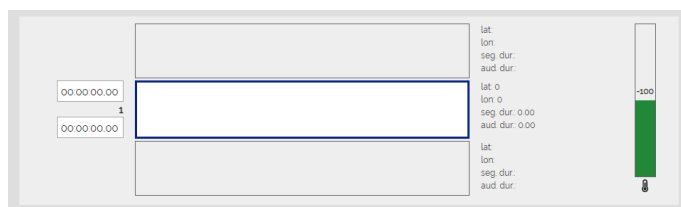


Figure 4 Script editing area

On the right side of the editor, there are three sections, named respectively: «Asset action», «Segment list» and «Recording controls» (Figure 5). The section «Asset action» allows to save audio files, and go back to the main page of the editing interface. «Segment list» contains the AD script with timecodes and a segment number. When AD for a given segment is recorded, the colour of this segment changes to green. When all segments are green, it means that all segments are recorded correctly. AD segments can be recorded by pressing the «Record» button. A countdown is provided for the recording to show audio describers how much time is left according to the timecodes set by them. Below the «Record» button, the audio level of the recording can be checked. After the AD is recorded, the recording can be

previewed in two tests: one starting two seconds before the timecode, and another one starting 5 seconds before the timecode.

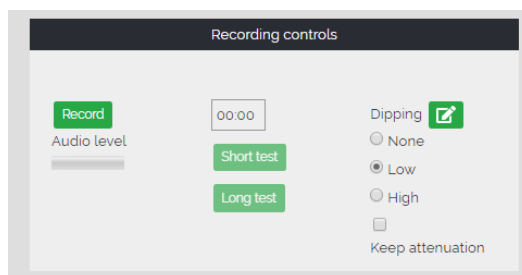


Figure 5 Recording controls

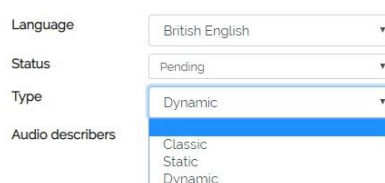
Additionally, fading can be chosen by audio describers. It refers to the decrease in the volume of the video when the AD is playing. In this regard, audio describers can choose between: none, low or high attenuation. The «None» option does not reduce the volume of the video. Conversely, when the option «High» is chosen, it will result in the significant reduction of the volume of the video. When the box «Keep fading» is checked, the volume of the main audio will be lowered until the next AD segment.

1.2. Features specific to immersive content

The AD web editor allows professional users to produce, preview and record audio description for 360° videos. As such, it includes functions which are specific to immersive content, which happens all around the 360° sphere.

Because of the characteristics of this media format, three types of AD can be created for each 360° video clip when working with the editor. They differ regarding the placement of spatial sound, and it is also recommended that audio describers use a different way of scripting for each of them. For the purposes of this project, these three AD types are provisionally referred to as: «Classic», «Static» and «Dynamic» (Figure 6). When professional users access the web editor, they choose one of them. The first sound option allows users to hear the AD as if it was coming from above their heads. It is the one that more closely mimics what is usually provided in standard, 2D, audio described content, and such way of scripting is recommended. The second one, «Static», is heard from the left or right side, as if someone was standing beside the user. As the AD is heard as if coming from a short distance from the user's location, it is recommended to write the AD in a non-standard way, with a stronger involvement of the describer. The third AD type, «Dynamic» allows audio describers to place the descriptions at different angles of the sphere, and it can be used to locate a character, object or event. In other

words, when a user plays a video with this type of AD, they will hear the AD coming from a specific point in the space. We believe that this option can guide users effectively within this highly visual medium, and will also allow them to feel more present in the content. This is an innovation in the field of audio description, as it can provide additional information on the location of the visuals being described by means of audio cues. As far as the scripting is concerned, a minimum audio description is recommended to allow users to be guided by the sound of the video.



The image shows a portion of a web editor interface. It features four dropdown menus arranged vertically. The first menu is labeled 'Language' and has 'British English' selected. The second menu is labeled 'Status' and has 'Pending' selected. The third menu is labeled 'Type' and has 'Dynamic' selected. The fourth menu is labeled 'Audio describers' and has three options: 'Classic', 'Static', and 'Dynamic', with 'Dynamic' currently selected and highlighted in blue.

Figure 6 AD types

Although many of the AD web editor's features are similar to the features of others subtitling or audio description editors, there are some new features that needed to be implemented, taking into account the specificity of the environment. One of such functions is «Set angle», a function implemented within the «Dynamic» AD type. This function allows users to assign a given AD segment to a specific angle within the sphere, specified by latitude and longitude. It means that the user, when watching the content with AD, will hear an AD of a given object, or event from the direction set by an audio describer, as it stays 'tied' to that part of the 360° sphere. This function, which can help users of access services to orient themselves inside the sphere, was implemented based on requirements defined in the focus groups (Fidyka and Matamala, 2018a). When an audio describer wants to create a «Dynamic» AD type, they need to look inside the sphere for a desired angle, and set it by clicking on a special button (Figure 7), or by a combination of keyboard shortcuts (customizable in the future version of the editor).

This new option will change the current workflow of producing AD. So far, audio describers needed to write AD instances between the dialogues, and mark their timecodes. Those audio describers whose task was also to voice the AD, needed to record it and preview it. When producing AD with a «Dynamic» AD type, besides following all the aforementioned steps, audio describers need to set an angle for every AD instance. In other words, they need to decide from which angle the linear AD will be heard by the end users.

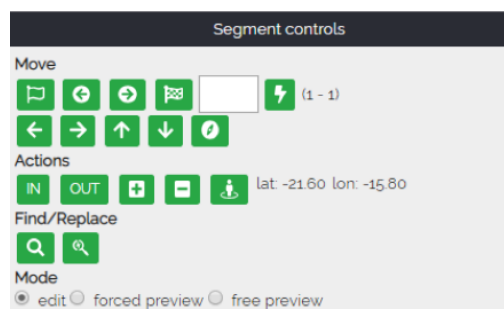


Figure 7 Set angle function and preview modes

As far as the preview of AD is concerned, two modes can be used for verification in the «Dynamic» AD type, namely «Forced preview» and «Free preview». In the first verification mode, AD and angle are bound to the video, which means that the video changes the angle during playback in the AD web editor. In other words, audio describers do not need to navigate through the video to find the angles they have set, as the software does it automatically. Thanks to this preview mode, audio describers are able to check whether the angles they set are correct. In the «Free preview», however, only the AD segments are bound to the timecodes. It means that audio describers can freely navigate the video during the playback, as during the editing.

As the functionalities described above are new, further testing is required to ascertain that they meet the needs of actual audio describers. This is why usability tests were conducted on the «Dynamic» AD type, in which they are implemented. The overall goal of our study was to assess the progress made in the development of the audio description editor. In particular, we wanted to learn whether the tool meets the needs of its actual users, so that its future version can cater for them more effectively. Additionally, our aim was to gather feedback on the work of the describer in relation to immersive environments and learn how such content will impact on their workflow.

2. METHOD

The usability test of the AD editor was performed online from the 24th of September to the 19th of October 2018. It aimed at receiving feedback from audio describers from different countries. The test instructions were given in English, and participants were asked to fill in the questionnaires in the same language. They were, however, requested to provide audio description in their native languages, which allowed us to gather feedback from users from

diverse countries, with different AD traditions. The response rate was 70.59%. The study was conducted voluntarily by professionals in accordance with ethical procedures approved by the Ethics Committee at Universitat Autònoma de Barcelona (UAB). Participants were informed about the aim and context of the study, and gave their consent before the test. Data are confidential, and the privacy of participants is ensured.

2.1. Participants

24 participants completed the test, 15 females and 8 males, plus one participant who preferred not to reply to this question. Their ages ranged from 25 to 64 years. The mean age was 36.71 years old. Their main languages were Catalan (1 respondent), Spanish (6 respondents), both languages (2 respondents), Bosnian (1 respondent), English (6 respondents), Dutch (2 respondents), Polish (3 respondents), German (2 respondents), and Swedish (1 respondent). They were mainly AVT translators, audio describers, AD and Media Accessibility supervisors, researchers and project managers. 21 participants had completed university studies, one participant had further education, and two participants, had secondary education.

2.2. Materials

Measures included usability and preferences. To obtain demographic data, two online questionnaires were prepared: a pre-questionnaire, and a post-questionnaire. They were sent to participants through separate e-mails upon agreeing to participation, as explained in section 2.3.

The post-questionnaire consisted of two parts: the System Usability Scale (SUS) questionnaire (Sauro & Lewis, 2016) and a preference questionnaire, with both closed and open questions which aimed at gathering additional user feedback. The SUS questionnaire chosen as a measure of usability contains 10 items, each with five steps: from strongly disagree to strongly agree (Sauro & Lewis, 2016, p. 198). It was available in English.

Regarding preferences, a specific questionnaire containing nine items was developed. It asked what participants appreciated or disliked about the editor, the possible improvements and missing functionalities, the level of difficulty related to the «Set angle» function, and the usefulness of preview modes. It also gathered the opinions of participants regarding whether or not describing 360° content is more time-consuming than that of standard content, and whether 360° videos will impact on their work in the coming years. The last question provided space for additional comments.

Each participant was assigned with one clip to be audio described with a «Dynamic» AD type. The video chosen was an initial 1-minute excerpt of a

fictional story suitable for audio description: there was enough space to produce audio description between the music and other sounds and, most importantly, the story developed at different angles of the sphere, which allowed us to test the «Set angle» function. The video was played in low resolution (720s) to avoid overloading the server, and to make the audio describing task smoother.

Other materials used in the study included a document with detailed instructions on how to perform the tasks, and a step-by-step user guide with screenshots on how to use the web editor. They were provided to participants by e-mail, as discussed in the next section.

2.3. Procedure

Prior to the study, a small-scale pilot test was performed with two users to evaluate the experimental protocol to be used on a larger scale. The participants met the criteria for inclusion in the sample, as they were professional audio describers. However, no changes were made to the final methodology because no problems arose in the pilot test development.

The study was developed online, and consisted of the subsequent steps. Participants were contacted through social media and personal contacts. Upon agreeing to participation, they received separate e-mails from the researchers, outlining the broader context and the procedure of the study. These e-mails included: (1) access to the editor, (2) a link to instructions on how to complete the test, (2) a link to a user guide, (4) links to the online questionnaires. Participants were informed about the exact order in which they should complete the tasks and fill in the questionnaires.

Participants completed the study in the following order. First they gave their informed consent to take part in the study. Then, the information that their data would be kept confidential was provided by the researchers. They were then asked to fill in the demographic questionnaire, perform a series of tasks in the editor, and fill in the post-questionnaire.

Participants were asked to complete the following tasks: (1) log in to the editor and open the assigned video, (2) audio describe the video excerpt in the user's native language, (3) preview the video in forced preview mode, (4) preview the video in free preview mode, and (5) save the AD and go back to the main window. Task 2 mentioned above («Audio describe the video excerpt») consisted of a series of sub-tasks: (1) add AD instances with correct timecodes, (2) set the angle for each AD instance, (3) record the AD segments produced, (4) insert one AD segment between two existing ones, and (5) delete two AD segments.

Once the tasks were completed, participants were asked to fill in the post-questionnaire, consisting of SUS and preference questions. At the end

of the study, participants were thanked, and information on how to obtain feedback was provided.

3. RESULTS AND DISCUSSION

Below we present results obtained from pre- and post-questionnaire as well as their discussion.

3.1. Results from pre-questionnaire: user profile

The pre-questionnaire focused on gathering socio-demographic information. Regarding the previous experience of participants in the AD field, most of the participants had received training in AD (91.67%). Nine participants (37.5%) had produced more than 300 hours of AD content, four participants (16.67%) had produced between 151 and 300 hours of AD content, four participants (16.67%) had produced between 51 and 150 hours, and seven participants (29.2%) had produced less than 50 hours of audio description. However, as little as four participants had had previous experience in audio describing a 360° video.

When asked about the usage of the AD editors, results show that not all audio describers use specific editing tools on a daily basis, with 10 of the participants reporting using only text processors and video players when producing AD.

The pre-questionnaire also aimed at determining cyber potential of research participants. When asked which devices they use on a daily basis, almost all participants (23) confirmed using mobile phones (95.83%); 21 participants use laptops (87.5%); 15 participants use TVs (62.5%), 14 participants use PCs (58.33%), eight of them use tablets (33.33%), one uses HMD (4.17%), and one participant chose the option «Other» (4.17%).

Unsurprisingly, most of participants do not use VR on a daily basis. To establish how frequently participants use VR, we asked them: 'How often do you watch Virtual Reality content (for instance, 360° videos)?' The results show that most of the participants (21, 87.5%) have never watched such content on a smartphone plugged into a head-mounted display or in a head-mounted display (20, 87.5%). Only one participant declared occasionally consuming Virtual Reality content on their smartphone, one participant occasionally uses a tablet to consume VR content and, regarding PCs, two participants (8.33%) use this device occasionally.

The next question asked about the reasons behind not using VR content or using it only occasionally. In this question, six participants (25%)

replied that they are not interested, three participants (12.5%) replied that it is not accessible, 12 participants (50%) replied that they have not had the chance to use it, two participants (8.33%) chose the option «Other reasons», and one participant (4.17%) did not provide any answer to this question. One of the participants provided an additional comment: «I don't normally access this content, I thought there were just a few, although I was surprised when accessing the project.»

Research participants were also asked to state their level of agreement with the statement «I am interested in Virtual Reality content (such as 360° videos)». The results show that four participants strongly agree with the statement (16.67%), eight participants agree (33.33%), nine participants neither agree nor disagree (37.5%), one participant disagrees (4.17%), and two participants strongly disagree (8.33%). Finally, when asked if they own any device with which to access Virtual Reality content, 10 participants replied that they do not (41.67%), five replied that they do not know or prefer not to reply (20.83%), and eight replied that they do (33.33%).

3.2. *Results on usability*

Regarding the results on usability, the score obtained in the SUS questionnaire is 55.9, which is considered below average, with a score of 68 or more considered as average. The obtained score corresponds to the percentile rank of 19%, and when converted to the letter grades, the obtained mark is D (Sauro & Lewis, 2016, pp. 203–204). This shows that the prototype web AD tool still has a lot of potential for improvement, and makes user testing at this stage even more relevant, as users are contributing to the definition of requirements as the tool is being developed.

The second part of the post-questionnaire focused on gathering data on users' preferences, and the results will be discussed question by question in the following section.

3.3. *Results from preference questions*

As far as the first question is concerned («What did you like the most about the AD editor?»), two of the participants (8.33%) highlighted the fact that the whole process of producing AD takes place in one piece of software. Four (16.66%) of the comments pointed to the fact that the most important functions (video, AD segments and recording) are displayed on one page, which facilitates the production of AD. In this regard, one of the participants commented: «It is quite easy, it has shortcuts and everything is visible and easily accessible on one page». Nine participants (37.5%) commented positively on the interface and its layout. They used the following words to describe it: «very clear», «simple», and «easy to understand». One of the

participants (4.17%) also positively assessed that the software is available online.

In the second and third questions, we asked participants which functions are the least useful, and how they could be improved. Nine audio describers (37.5%) pointed to the second question («What did you like less about the AD editor?») the problems encountered in the recording and preview modules. Also in the third question («What do you think could be improved, and how?»), problems with the video and buttons which would freeze or play with delay, were mentioned by five describers (20.83%). One comment (4.17%) in the third question suggested that a better video quality would be needed in order to describe all details. In this case, this was due to the testing conditions, as a higher quality video could overload the editor, and slow down the AD production process. Additionally, five (20.83%) participants suggested in both questions that they would prefer a different, more intuitive configuration of the shortcuts – or that they would prefer to customize the shortcuts themselves. They reported that this change is essential, as using the shortcuts they are accustomed to would allow them to work more efficiently. As already mentioned, although the shortcuts were not customizable in the prototype version of the editor, audio describers will be able to adapt them to their needs in the final version. Importantly, one comment (4.17%) in the second question pointed out that it would be necessary to preview the produced AD not only in the web editor, but also in a head-mounted display. We also believe that this option would be worth implementing, as it would allow audio describers to preview their work in the way in which end users will be consuming the content.

The following question asked participants about other functionalities that could be implemented. In this regard, participants suggested many improvements that are technically feasible. Among them, a waveform that would indicate music and other sounds, was suggested by two participants (8.33%) as a way to time audio description more precisely. This question was put forward in one comment (4.17%) as follows: «Perhaps one [function] in which you could see the lines corresponding to sound (...) is very useful to ensure that AD does not interfere with dialogues, or other sounds in the film.» Regarding combining AD with the other sounds of the video, one of the participants suggested that more options for fading the audio of the video would be needed, as there are currently only three options which are possible: «none», «low» and «high fading». This comment seems particularly relevant, as the audio volume of the AD track is a critical aspect when ensuring that descriptions can be heard clearly throughout the video or film.

An important improvement was suggested by another participant, who considers that a synchrony between AD segments and video would be needed

to facilitate the work. It means that if someone clicks on an AD segment on the «Segment list», the video should move to its timecode.

Other suggested improvements focused on both standard functionalities, and functions specific to the medium under discussion. They included: moving back or forward 5–10 frames at a time, a map in which all the locations with ADs assigned to them would be located, being able to merge or separate segments, setting a minimum separation between AD segments and dialogues or other sounds, and seeing the actual reading speed of the AD. Another participant suggested that using earphones during recording would be needed.

Importantly, an option to export the script to a text file for a professional recording was suggested in one comment (4.17%), as not all audio describers record AD themselves. For example, in Poland, audio describers send their script to a professional responsible for recording (Chmiel & Mazur, 2014). Audio describers who do not record the AD themselves, mark parts of the dialogue in the script or sounds that proceed, or follow a given AD segment. They also indicate in brackets how the given instance should be read, for example slowly or rapidly. The need to mark the part of dialogue in the script intended for recording by a professional was put forward in the comment of one of the participants: «I am used to using sound and dialogue cues rather than 'In times' only. I find it helpful to see the dialogue cue that leads into a description».

Regarding the innovative «Set angle» function, the results suggest that most of the participants (75%) found it easy to use, despite one participant reporting technical problems with their laptop. Another response suggests that one participant would like to set the angle for only some events, and not for all AD segments: «(...) I would like to have more freedom. The tutorial tells me we need an angle for each segment. I would like to have an angle only for very important situations». Therefore, it would be worth exploring a combination of the so-called «Dynamic» and «Classic» audio descriptions. Moreover, it would be worth exploring other spatial sound possibilities that allow one not only to set the angle from which the audio description will be heard, but also a specific position, ideally linked to an object.

As far as the preview modes – crucial to check if the set angles are correct – are concerned, five participants (20.83%) did not encounter any problems while using them, e.g. «(...) one allows you to move, the other one makes you see your fixed angles», but for 11 participants (approximately 45.8%) one or both preview modes did not work properly, and delay and visualisation problems were reported in the prototype version. Five participants (20.83%) did not provide an answer to this question.

Beyond the specific questions linked to the editor, it was interesting to gather general feed-back on the experience of audio describing 360° content. When asked about whether it takes longer to describe 360° videos than standard content, most participants (79.2%) replied positively, adding that setting an angle for each AD segment needs additional time.

Although not entirely within the scope of this article, some participants addressed the question of content selection in this emerging medium, the question discussed in depth in Fidyka and Matamala (2018b). Similarly to the results of the focus groups organized to provide a first insight into this question, it was deemed challenging, as explained in one of the comments: «technically, I consider it easy to use, the problem is which angle is the most important to describe.» It was added in other comments that audio describers will need special guidance when deciding which elements are the most relevant and should be audio described. This comment is similar to the comments put forward in the focus group organized in Poland, during which participants suggested that cooperation with content creators should be sought when audio describing 360° videos (Ibid.). In this regard, guidelines that would provide audio describers with guidance on what to describe, when to describe it and how would prove particularly useful, but they have not yet been created.

The last question in the post-questionnaire asked participants whether they consider that 360° videos would impact on their AD practice in the coming years. In this regard, participants expressed varying opinions. 58.3% of the participants answered affirmatively to this question. One of the comments further explained that the application for this medium is vast, and as more content is created, it should be provided with access services. Another comment pointed to the fact that this medium changes the current approach for producing AD, which means that specific training, or guidelines should be offered to audio describers producing AD in this medium. As explained by one of the participants: «I wouldn't feel comfortable taking on a job like this without proper instructions from the client and/or relevant training». This opens an opportunity for training institutions, which should accept the challenge of training professionals not only in the traditional modes, but also in more innovative media.

Additionally, one of the participants provided a detailed report after completing the test. This participant's overall assessment was positive, but they suggested some technical changes. Most of the comments pointed to improvements that could be made in the standard functions of the prototype editor (such as including a more responsive video control bar, improving the navigation by timecode, moving between AD segments), but one comment focused on the functions specific to immersive environments: audio describers

producing AD in this medium would need a globe («Mercator map projection») on which all the set locations could be seen on one map. Moreover, echoing the opinions voiced in the focus group in Kraków, discussed in Fidyka and Matamala (2018b), this participant suggested that there should be the opportunity to read the audio description text via speech synthesis during the preview. Importantly, previewing in a full-screen mode should be possible.

All in all, the results of both parts of the questionnaire have provided valuable feed-back to improve the preliminary version of the tool, thanks to a user-centric methodology in which users are asked for feedback during the development process. Although quantitative data regarding usability in these stages generally falls short of what is expected from a market tool, it provides a benchmark for comparison in future releases, and contributes to qualitative feedback.

CONCLUSIONS

The usability test of the AD web editor, developed so as to respond to the characteristics of 360° content, allowed us not only to assess the progress made in the development of the tool, but also to better understand the habits and needs of professional audio describers.

The results of the usability test show that although many of the features of the AD web editor have been assessed positively by the respondents taking part in the study, there is still room for considerable improvement in order to meet the needs of professional users. Thanks to participants' suggestions, we found several aspects of the editor which could be improved. These include, for instance, customizing the shortcuts, previewing AD with speech synthesis or previewing it by means of HMD. The technical feasibility of some other proposals, such as a map on which AD segments could be seen, needs to be explored.

Additionally, the results confirm the findings of the previous studies conducted in the form of focus groups (Fidyka and Matamala, 2018a, 2018b): content selection in this media format is considered challenging by audio describers who need guidance on how to describe. Beyond the specific analysis of the tool, the test has shown some attitudes of current audio describers towards new immersive media: in this regard, they state that this new production system will be more time-consuming, and request specific guidelines and training.

The limitations of this study need to be acknowledged. One limitation is the sample size of 24 participants. The future usability study testing the final version of the tool could be conducted with a higher number of participants to

obtain more reliable results. Secondly, the fact that the 360° video used as a stimulus in the study was watched on the flat screens of laptops or PCs with which the audio describers worked might have had impact on the results. Although this was due to the fact that the study was conducted online to reach professional audio describers from different backgrounds, watching the stimulus first in the HMD would be beneficial for audio describers to better understand the presented content. Another limitation could be related to the audio describers' lack of previous experience with AD software, which could have an impact on the obtained results.

Regarding future studies, more research will be needed on the implementation of spatial audio in AD in this media format to better respond to the needs of the end users of the access services. Additionally, it should be explored how spatial sound could enhance user experience when used in audio subtitles (AST), which serve audiences who cannot access the written subtitles and cannot understand the language of the original (Braun & Orero, 2010; Reviere & Remael, 2015). Also, more studies would be needed on AD in more complex VR types. Finally, an additional set of empirical tests will need to be conducted with end-users in order to assess whether the access services produced cater to their needs, allowing them to understand, enjoy, orient themselves, and feel present in the 360° content.

ACKNOWLEDGEMENTS

ImAC has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the grant agreement 761974. The authors are members of TransMedia Catalonia, an SGR research group funded by «Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya» (2017SGR113). This article is part of Anita Fidyka's PhD 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.

REFERENCES

- Agulló, B., Matamala, A., & Orero, P. (2018). From Disabilities to Capabilities: Testing Subtitles in Immersive Environments with End Users. *Hikma*, 17, pp. 195–220. doi: 10.21071/hikma.v17i0.11167

- Agulló, B., Matamala, A., & Orero, P. (In press). The Challenge of Subtitling for the Deaf and Hard-of-hearing in Immersive Environments: Results from a Focus Group. *Journal of Specialised Translation*, 32.
- Allen, C., & Tucker, D. (2018). Immersive Content Formats for Future Audiences. Retrieved from Digital Catapult on February 14, 2019: https://assets.ctfassets.net/nubxhjiwc091/5WypbBxFEQqS816WaKUQgg/ca25eb800e9f88a740079dbca623ca2c/Immersive_Content_Formats_for_Future_Audiences.pdf
- Braun, S., & Orero, P. (2010). Audio Description with Audio Subtitling: An Emergent Modality of Audiovisual Localization. *Studies in Translatology*, 18(3), pp. 173–188. doi: <https://doi.org/10.1080/0907676x.2010.485687>
- Brown, A. (2017). User Testing Subtitles for 360° Content. Retrieved from BBC (Research & Development) on February 14, 2019: <https://www.bbc.co.uk/rd/blog/2017-10-subtitles-360-video-virtual-reality-vr>
- Chmiel, A. & Mazur, I. (2014). *Audiodeskrypcja [Audio description]*. Poznań: Wydział Anglistyki UAM.
- Dooley, K. (2017). Storytelling with Virtual Reality in 360-degrees: A New Screen Grammar. *Studies in Australasian Cinema*, 11(3), pp. 161–171. doi: 10.1080/17503175.2017.1387357
- Fidyka, A., & Matamala, A. (2018a). Audio description in 360° videos: Results from Focus Groups in Barcelona and Kraków. *Translation Spaces*, 7(2), pp. 285–303. doi:10.1075/ts.18018.fid
- Fidyka, A., & Matamala, A. (2018b). Retelling Narrative in 360° Videos: Implications for Audio Description. Manuscript submitted for publication.
- Fryer, L. (2016). *An Introduction to Audio Description. A Practical Guide*. London: Routledge.
- Fryer, L., & Johnatan, F. (2013). Cinematic Language and the Description of Film: Keeping AD Users in the Frame. *Perspectives: Studies in Translatology*, 21(3), pp. 1–15. doi: 10.1080/0907676X.2012.693108
- Gödde, M., Gabler, F., Siegmund, D., & Braun, A. (2018). Cinematic Narration in VR – Rethinking Film Conventions for 360 Degrees. *Virtual, Augmented and Mixed Reality. 10th International Conference, VAMR 2018*. (pp. 184–201). Las Vegas: Springer. doi: 10.1007/978-3-319-91584-5_15

- Greco, G. M. (2016). On Accessibility as a Human Right, with an Application to Media Accessibility. In A. Matamala & P. Orero (Eds.), *Researching Audio Description: New Approaches* (pp. 11–34). London: Palgrave.
- Greening, J. (2011). Do we Need 3D Audio Description Guidelines: Recommendations from Focus Group Study Report. Retrieved from (February 20, 2019): http://audiodescription.co.uk/uploads/general/3D_film_and_tv_report_2.pdf
- Jerald, J. (2015). *The VR Book: Human-centered Design for Virtual Reality*. Sant Rafael, CA: Morgan & Claypool. doi: 10.1145/2792790.2792827
- Johansson, M. (2019). VR for Your Ears: Dynamic 3D Audio Is Coming Soon. Retrieved from IEEE Spetrum on February 19, 2019: <https://spectrum.ieee.org/consumer-electronics/audiovideo/vr-for-your-ears-dynamic-3d-audio-is-coming-soon>
- Jones, S. (2017). Disrupting the Narrative: Immersive Journalism in Virtual Reality. *Journal of Media Practice*, 18(2–3), pp. 171–185. doi: 10.1080/14682753.2017.1374677
- López, M., Kearney, G., & Hofstädter, K. (2016). Enhancing Audio Description: Sound Design, Spatialisation and Accessibility in Film and Television. *Proceedings of the Institute of Acoustics*, 38(2). Retrieved from (December 16, 2018): <http://enhancingaudiodescription.com/wp-content/uploads/2016/11/RS2016-paper-Lopez-et-al.pdf>
- Maszerowska, A., Matamala, A., & Orero, P. (Eds.) (2014). *Audio Description: New Perspectives Illustrated*. Amsterdam: John Benjamins.
- Nielsen, L. T., Møller, M. B., Hartmeyer, S. D., Ljung, T., Nilsson, N. C., Nordahl, R., & Serafin, S. (2016). Missing the Point: An Exploration of how to Guide Users' Attention during Cinematic Virtual Reality. *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology* (pp. 229–232). New York: ACM. doi: 10.1145/2993369.2993405
- Portillo, C. (2018). Films Unseen: Approaching Sound Design Alternatives to Enhance Perception and Mental Imagery Models among Audio Description Audiences on Sci-fi Films. Paper presented at *Accessibility in Film, Television and Interactive Media Conference*. Retrieved from (December 17, 2018): <http://enhancingaudiodescription.com/conference/>

- Reviere, N., & Remael, A. (2015). Recreating Multimodal Cohesion in Audio Description: A Case Study of Audio Subtitling in Dutch Multilingual Films. *New Voices in Translation Studies*, 13(1), pp. 50–78. Retrieved from (February 20, 2019): https://www.iatis.org/images/stories/publications/new-voices/Issue13-2015/Articles/Reviere_New_voices_PUBL.pdf
- Rothe, S., Hußmann, H., & Allary, M. (2017). Diegetic Cues for Guiding the Viewer in the Cinematic Virtual Reality. *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*. New York: ACM. doi: 10.1145/3139131.3143421
- Rothe, S., Tran, K., & Hußmann, H. (2018). Dynamic Subtitles in Cinematic Virtual Reality. Paper presented at ACM TVX 2018. Seoul. doi: <https://dl.acm.org/citation.cfm?id=3213556>
- Sauro, J., & Lewis, J. (2016). *Quantifying the User Experience: Practical Statistics for User Research*. Amsterdam: Morgan Kaufmann.
- Snyder, J. (2008). Audio description: The Visual Made Verbal. In J. Díaz-Cintas (Eds.), *The Didactics of Audiovisual* (pp. 191–198). Amsterdam/Philadelphia: John Benjamins. doi: 10.1075/btl.77.18sny
- Streit, M., & Victoria, J. M. (2017). Guiding Viewer's Attention in 360 video and VR through Audio. Paper presented at the 4th International Conference on Spatial Audio.
- Suojanen, T., Koskinen, K., & Tuominen, T. (2015). *User-centered translation*. London: Routledge.
- Syrett, H., Calvi, L., & van Gisbergen, M. (2017). The Oculus Rift Film Experience: A Case Study on Understanding Films in a Head-mounted Display. In R. Poppe, J.J. Meyer, R. Veltkamp, & M. Dastani (Eds.). Springer. doi: 10.1007/978-3-319-49616-0_19
- Viswanathan, L. N., McDaniel, T., Krishna, S., & Panchanathan, S. (2010). Haptics in Audio Described Movies. *IEEE International Symposium on Haptic Audio Visual Environments and Games*, (pp. 1–2). Phoenix. doi: <https://doi.org/10.1109/have.2010.5623958>

APPENDICES:

Appendix 1: Pre-questionnaire

1. Sex:
 - a. Female

- b. Male
 - c. Other
 - d. I prefer not to reply
2. Age:
 3. Main language: (open question)
 4. Please, describe your current job: (open question)
 5. Have you ever audio described a 360° video? Yes / No
 6. For how long have you been working in the field of audio description? (open question)
 7. How many hours of audio description have you produced in your professional life?
 - a. Less than 50 hours
 - b. 51-150 hours
 - c. 151-300 hours
 - d. More than 300 hours
 8. In what language or languages do you normally audio describe?
 9. What software do you normally use?
 10. Please indicate your level of studies.
 - a. Primary education
 - b. Secondary education
 - c. Further education. Please specify (open field)
 - d. University. Please specify (open field)
 11. If you replied "Further education" or "University" in the previous question, please specify. (open question)
 12. If you have received specific training on audio description, please indicate it here.
 13. What devices do you use on a daily basis? Multiple replies are possible.
 - a. TV
 - b. PC
 - c. Laptop
 - d. Mobile phone
 - e. Tablet
 - f. HMD
 - g. Other: (open field)
 14. How often do you watch Virtual Reality content (for instance, 360° videos)?

	Never	Occasionally	At least once a month	At least once a week	Every day
On a smartphone					

On a tablet					
On a PC					
On a smartphone plugged to HMD					
In HMD					

15. If you have never used Virtual Reality content such as 360° videos or only occasionally, please indicate why. Multiple answers are possible.
- Because I am not interested.
 - Because it is not accessible.
 - Because I have not had the chance to use it.
 - Other reasons. Please explain: (open field)
16. Please state your level of agreement with the following statement: "I am interested in Virtual Reality content (such as 360° videos)."
- I strongly agree
 - I agree
 - Neither agree nor disagree
 - Disagree
 - Strongly disagree
17. Do you own any device to access Virtual Reality content?
- Yes (If yes, which one? _____)
 - No
 - I don't know or I don't want to reply
18. If you replied "yes" to the previous question, please specify which device(s). (open question)

Appendix 2: Post-questionnaire

System Usability Scale

1 – strongly disagree, 5 – strongly agree

- I think that I would like to use this system frequently.
- I found the system unnecessarily complex.
- I thought the system was easy to use.
- I think that I would need the support of a technical person to be able to use this system.
- I found that the various functions in this system were well integrated.
- I thought there was too much inconsistency in this system.

-
7. I would imagine that most people would learn to use this system very quickly.
 8. I found the system very cumbersome to use.
 9. I felt very confident using the system.
 10. I needed to learn a lot of things before I could get going with this system.

Preference questions

Now please reply to the following questions in your own words.

11. What did you like most about the AD editor?
12. What did you like less about the AD editor?
13. What do you think could be improved, and how?
14. Did you miss any functionality? If yes, can you tell us which?
15. Do you find the feature for setting the angle for the AD easy to use? Explain why.
16. Were the preview modes useful for you? Explain why.
17. Do you think it will take you longer to audio describe videos in 360°? Why?
18. Do you think 360° videos will impact on your work as an audio describer?
19. Other comments: (open field)