

## **Subtitling for the deaf and hard-of-hearing in immersive environments: results from a focus group**

**Belén Agulló, Universitat Autònoma de Barcelona**

**Anna Matamala, Universitat Autònoma de Barcelona**

### **ABSTRACT**

Immersive media such as virtual reality or 360° content is increasingly present in our society. However, immersive content is not always accessible to all, and research is needed on how to cater for the needs of different kinds of users. This article will review the current situation of immersive technologies and their applications in media. Also, research studies carried out so far concerning subtitling and SDH in immersive media will be discussed, as well as current implementation of subtitles in immersive content, such as VR video games or 360° videos. Finally, the results from a focus group carried out in Spain with deaf and hard-of-hearing users, as well as professional subtitlers, on how to subtitle 360° video content will be presented. The results from the focus group shed light on how to address the new implications brought by immersive media in regard to subtitling. Some of the main areas discussed in the results are: subtitle position, contrast considerations for a correct reading, and how to indicate the location of the speakers, among others. Also, results show that users are willing to accept the implementation of new features in SDH in immersive content, such as icons for non-speech information or improvements to current standards.

### **KEYWORDS**

Subtitles, subtitling for the deaf and hard-of-hearing, SDH, immersive content, virtual reality, 360° content, focus group.

## **1. Introduction**

Immersive media are increasingly present in our society and considerable research efforts are put in developing better immersive technologies (Jayaraj *et al.* 2017). According to a report by VR Intelligence (2017), nearly half of the surveyed VR companies (46%) reported a strong or very strong growth in sales. However, some reports agree that the main reasons preventing users to buy VR headsets are economic, because prices are too high, and the lack of available content (VR Intelligence 2017; Sketchfab 2017).

Even if data point to expected growth, immersive content is not always accessible to all, and research is needed on how to cater for the needs of diverse users. Audiovisual Translation (AVT) and more specifically Media Accessibility (MA) (Remael *et al.* (eds) 2014; Greco 2016), is the field in which research on access to audiovisual content has been carried out in the last years, generally focusing on access services such as audio description (AD), subtitling for the deaf and hard-of-hearing (SDH) or sign language (SL) interpreting, among other. Still, most research has dealt with traditional media such as TV or cinema (Perego *et al.* 2015; Romero-Fresco (ed.) 2015), museums (Jiménez Hurtado *et al.* 2012; Szarkowska *et al.* 2016; Neves 2018) or live events (Orero and Matamala 2007; Udo

and Fels 2011). In these environments, as in many others such as the localisation and game industry, accessibility has generally been considered an afterthought, despite many voices asking for the inclusion of accessibility in the creation process (Romero-Fresco 2013). To date, little research on accessibility in immersive media has been carried out and immersive technologies are on the rise, but they are still not fully implemented. This scenario was seen as an opportunity to start researching access services while immersive media were being developed, and the ImAc project was set up.

ImAc is a European project funded by the European Commission that aims to research how access services (subtitling, AD, audio subtitles, SL) can be integrated with immersive media. The project aims to move away from the constraints of existing technologies into an environment where consumers can fully customise their experience. The key action in ImAc is to ensure that immersive experiences address the needs of different kinds of users. ImAc follows a user-centred methodology (Matamala *et al.* 2018), so the first step in the project has been to ask users about their expectations. Two focus groups have been carried out in two countries (Germany and Spain) with different user profiles to gather feedback, define user scenarios and establish user requirements regarding SDH in 360° video content. This article aims to discuss the results of the focus group carried out at the Catalan Media Corporation (CCMA) in Spain.

The article begins with an overview of immersive technologies and immersive content in media, in order to contextualise our research. It then explains the limited research that has been carried out so far concerning subtitling in immersive media. Section 5 describes the methodology for the focus group and its implementation, and Section 6 reviews the results.

## **2. Immersive content: an overview**

Immersive content allows users to feel as if they were physically transported into a different location. There are different solutions that can provide such experience. Fulldomes are one of those solutions. These are based on panoramic 360° videos projected on a dome structure, such as those that can be seen in planetariums, museums or flight simulators, for example.

Stereoscopic 3D technology represents another type of immersive technology that has had a relative presence in cinemas and homes during the past decade. Specifically, stereoscopy or 3D imaging:

refers to a technique to create or enhance the illusion of depth in an image by presenting two offset images separately to the left and right eye of the viewer. The two 2D images are then combined in the brain to give the perception of 3D depth. The visual cortex of the brain fuses the two images into the perception of a three-dimensional scene or composition. (Agulló and Orero 2017: 92)

However, the quality of the immersive experience and the sense of depth depend on the display designs, which for 3D content are diverse and

lacking in standards (Holliman *et al.* 2011). Depending on the display design, the way of accessing stereoscopic 3D images differs (anaglyph glasses, head-mounted displays, active/passive glasses, etc.). This lack of standardisation, the intrusive nature of 3D and some uncomfortable side effects (headache or eyestrain) might prevent stereoscopy to become the main display for audiovisual (AV) products (Belton 2012).

The failure to adopt 3D imaging as mainstream display for AV products may have opened a door for VR and 360° content, as a new attempt to create engaging immersive experiences. Sherman and Craig (2003) define four key factors in VR (virtual world, immersion, sensory feedback and interactivity) which result in the following definition:

**virtual reality** [is] a medium composed of interactive computer simulations that sense the participant's position and actions and replace or augment the feedback to one or more senses, giving the feeling of being mentally immersed or present in the simulation (a virtual world). (Sherman and Craig 2003: 13)

VR has also been referred to as a new medium, like the telephone or television, by authors such as Steuer (1992), who considers VR as a set of technical components, for example computers, head-mounted displays, sensors, among others. Most recently, VR has also been referred to as a set of computer-generated images that reproduce a reality and allow users to interact with their surroundings with the appropriate equipment (BBC 2014). However, some authors (Biocca 1992; Steuer 1992) consider technical definitions of VR to be limited. They suggest defining VR in terms of human experience and introducing the concept of *presence*, which can be defined as “the experience of one’s physical environment; it refers not to one’s surroundings as they exist in the physical world, but to the perception of those surroundings as mediated by both automatic and controlled mental processes” (Steuer 1992: 75).

This definition can be applied to different types of immersive content which, contrary to VR, does not need to involve an interactive response and an advanced equipment. Therefore, 360° or omnidirectional videos can also be defined as immersive content: they reproduce highly-realistic images recorded with special camera sets that represent a reality in which the users are observants and cannot interact (BBC 2014). Moreover, images can be combined with audio technologies such as immersive audio aimed to increase immersion. The spatial sounds are created with specially designed microphones that simulate the way in which human beings receive aural stimuli through the auditory system (BBC 2012).

Other less commercialised immersive technologies are mixed and augmented reality. Milgram and Kishino (1994: 1321) define those terms as follows:

Mixed Reality (MR) visual displays [...] involve the merging of real and virtual worlds somewhere along the ‘virtuality continuum’ which connects completely real environments to completely virtual ones. Probably the best known of these is Augmented Reality (AR), which refers to all cases in which the display of an otherwise real environment is augmented by means of virtual (computer graphic) objects.

In this definition, the authors explain their concept of “virtuality continuum” in which MR is the hypernym that includes the more specific term AR. Azuma (1997) further defines AR as a medium that combines the real world with virtual objects that appear superimposed in the real world. The properties of AR are that it “combines real and virtual objects in a real environment; runs interactively, and in real time; and registers (aligns) real and virtual objects with each other” (Azuma *et al.* 2001: 34).

VR and 360° content is accessed by different types of equipment. It can be directly viewed on a flat screen (for example, a computer, smartphone or a TV set) with a remote, touch pad or mouse to change the direction of the field of view (FoV), or it can be accessed with a head-mounted display (HMD), which can be either a tethered or a mobile device. Tethered HMD, such as PlayStation VR, Oculus Rift or Vive, incorporate high definition screens and are connected to high-performance computers or last generation consoles (Deloitte 2016). They are generally used for gaming purposes and the quality of the VR experience is higher. On the other hand, mobile HMD, such as Samsung Gear VR or Google Cardboard, are dependent on smartphone technology (such as accelerometer or gyroscope).

### **3. Immersive content in media**

Broadcasters and video content developers are starting to experiment with immersive content. According to a report on VR by the European Broadcasting Union (EBU 2017a), 49% of its members are starting to explore or further develop immersive content. Most EBU members think that the potential of immersive content is clear, because it offers new opportunities to tell stories from a different perspective. However, factors such as technical limitations, lack of knowledge and uncertainty about the return on investment are holding some of them back (EBU 2017a). The current preferred format is 360° video over VR or AR/MR and the trend in terms of duration is 5-10 minutes. Most stories told in 360°/VR are “history or news and current affairs products, as 360°/VR allows the user to gain better understanding of the story being told” (EBU 2017a: 9). Sports and music events are clear candidates for 360°/VR. In the report, attention is also directed at the challenges posed in terms of storytelling, since the plots are non-linear and the level of user interaction is not determined.

Immersive journalism is featured as a key concept in the EBU report. This relatively new concept entails “the production of news in a form in which people can gain first-person experiences of events or situation described in news stories” (De la Peña *et al.* 2010: 291). According to De la Peña *et al.*, VR is a perfect media that could help journalists in eliciting deeper emotions in the audience. Major broadcasters and newspapers such as the BBC, *The New York Times*, *The Washington Post* and *ABC News* have already started working on this. To name just two examples, in 2015 *The New York Times* decided to reward their subscribers with a pair of Google

Cardboard glasses, as a strategy to promote their own immersive content (Wohlsen 2015). The company even launched its own smartphone application (NYT VR). More recently, as reported by Jones (2017), *ABC News* recorded a 360° short film about North Korea with a relatively positive outcome. Immersive journalistic experiences have been proven to elicit emotions and increase audience engagement (Jones 2017), showing greater potential for journalism and broadcasting.

As a genre, fiction appears to be a rather underexplored area in relation to immersive content. However, there are some indicators which show that VR has had a moderate penetration in the entertainment industry. For example, the appearance of immersive content in major film and TV industry events is one of them. In 2016, the VR film *Henry* was awarded an Emmy for Outstanding Original Interactive Program (Oculus Team 2016). In 2017, the immersive short film *Pearl* was nominated for the Oscars, under the category of Best Animated Short Film (Hall 2017). Pixar has also developed a VR and interactive experience to market their animated film *Coco* (Lee 2017). However, most fictional immersive experiences are computer-generated products and not images recorded with 360° cameras. According to the BBC, “truly interactive VR video is in its infancy and can be expensive to create, but total or partial animation or CGI can be used very effectively and efficiently, while other production techniques may yet emerge or become more accessible over time” (Conroy 2017). Therefore, it could be inferred that immersive technologies are still not sufficiently developed to be implemented in creating successful fictional films or movies. Some of the reasons could be the hindrances posed by technology, which is delivering a quality that is still not considered suitable for the audience (EBU 2017a); also, the lack of knowledge in delivering well-written immersive stories. It is in this context of development that the integration of access services in the production line should be researched, adopting a user-centred approach. The users’ voice needs to be heard before the technology is fully implemented.

#### **4. Subtitling and SDH in immersive media**

AVT and MA research on immersive content is at an early stage. Nevertheless, we can find studies that have addressed the challenges of creating and consuming subtitles in stereoscopic 3D content (Vilaró 2011; Lambooij *et al.* 2013). Some of the main issues when integrating subtitles in 3D imaging is that superimposing 2D subtitles on a 3D image can generate effects such as ghosting, which hinders the readability of the subtitles and can cause headaches and eyestrain (Agulló and Orero 2017). However, the implementation of some techniques such as the positioning of the subtitle close to the screen plane or the use of lighting, shades and colours to reduce contrast between the screen and the subtitle (González-Zúñiga *et al.* 2013) could contribute to minimising the impact of such issues.

Reception studies on access services in VR and 360° content are almost non-existent, with research on subtitling by the BBC being an exception. Although not focused specifically on SDH, audiences with hearing loss might be potential users of intralinguistic subtitles. The BBC research designed four subtitle scenarios for 360° content (Brown *et al.* 2018: 3-6): (a) subtitles equally spaced by 120° in a fixed position below the eye line; (b) subtitles following head immediately always in front of the user; (c) subtitles following head with lag in front of the user; and (d) subtitles appearing in front of users, and then fixed until they disappear. 24 participants, frequent users of TV subtitles, took part in the study. They randomly viewed the four variables on an Oculus Rift HMD, and replied to a questionnaire (Brown *et al.* 2018: 7-9). Results show that the preferred solution was (b), in line with a similar behaviour to subtitles in 2D. Their conclusion was that sometimes the simplest solution is the best (Brown *et al.* 2018: 29-33), but it remains to be seen whether results would be the same with longer content.

#### **4.1. Revisiting subtitling parameters in immersive content**

User reception studies in subtitling and in SDH have allowed the definition of a set of preferred parameters for users (Jensema *et al.* 1996; Romero-Fresco 2009; Matamala and Orero (eds) 2010) in a myriad of aspects such as: the number of characters and lines per subtitle; subtitle editing; font type and size; boxes, borders and shadows; justification and spacing; paralinguistic information, and subtitle speed (Neves 2005; Arnáiz-Uzquiza 2012; Romero-Fresco (ed.) 2015). Immersive environments, however, pose specific challenges that need to be considered.

Subtitle positioning, which is a widely researched (Bartoll and Martínez-Tejerina 2010) and standardised parameter for 2D products, is one of the main issues when designing subtitles for immersive content, since user behaviour in the immersive environment is unpredictable (Arrés *forthcoming*). While safe areas for subtitling are already defined and guidelines are provided for content consumed in TV or flat screens (EBU 2017b), recommendations for safe areas in immersive devices such as HMD are still lacking. The FoV for users in VR environments is wider than the FoV in 2D products on a flat screen. But to the best of our knowledge, eye-tracking studies, showing where the users direct their attention when reading subtitles in VR or 360° content, are lacking. Therefore, the safe area for subtitles in immersive environments needs to be defined very carefully and tested later on.

The interactivity and the freedom of movement that are inherent to immersive products also impact other subtitling parameters such as character identification, because there is the possibility that a character speaks but is located outside the FoV of the user. In 2D SDH, character identification is usually solved by using different colours, name tags or speaker-dependent placement of subtitles or a combination of these (Neves 2005; Arnáiz-Uzquiza 2012). However, immersive content

introduces a new dimension: direction, which is particularly relevant in immersive environments.

Another SDH parameter that may change in subtitles for immersive content is the display of non-speech information, such as music or sounds. Immersive technologies offer new opportunities for implementing new features in SDH. In previous studies, the implementation of graphic elements such as icons to display non-speech information has been suggested as a way to optimise reception (Civera and Orero 2010). Other authors have also tested the reception of emotions via emoticons (Arnáiz-Uzquiza 2015) and other creative approaches (Sala Robert 2016). Although these are not extended practices for SDH, the technical advances provided by immersive technologies could open the possibility of introducing new elements that might counterbalance other VR limitations. However, it remains to be seen whether alternative approaches such as using icons may help reduce that kind of discomfort.

#### **4.2. Some examples from current practices**

Some randomly selected current experiences in subtitling immersive environments can give us food for thought as to the opportunities and challenges subtitles in immersive AV products may pose. The Spanish television series *El Ministerio del Tiempo* (The Ministry of Time) launched an immersive experience in the form of an interactive episode. The episode “El tiempo en tus manos” (The time is in your hands) is one of the first fictional TV episodes launched in an immersive format. The interlingual subtitles (Spanish into English) in this short episode are positioned slightly below the centre of the screen, following the movement of the head and floating through the screen as the user’s head moves. The transition of the subtitles, when the users move their heads, presents a slight delay in reaction time. Therefore, when the movement is abrupt, subtitles are not positioned in the centre of the screen, but float in the direction of the movement of the head. They only settle into a fixed position in the centre of the image when the user’s head is still. The font type is a white sans serif font without a background box. The justification is centred. The segmentation rules are not followed, and some linguistic issues are found, such as missing information.

Another example is the clip *The Displaced* created by *The New York Times*, in which children who have been driven away from their homes explain their current situation as refugees. In this video, subtitles are burnt in in three different fixed positions in the 360° video, so when the users move their head to explore the scenario, they will always find the subtitles somewhere in their FoV (Brown *et al.* 2017). The font is white sans serif, smaller than the previous example, which hinders readability. Moreover, these subtitles do not include a background box and the contrast is very low, meaning that sometimes the text is very difficult to read.

Video games provide other examples of subtitles in immersive environments, although subtitling practices in video games do not always

follow the same rules as in other AV content (Mangiron 2013). In the game *Eve Valkyrie*, for PlayStation VR, intra- and interlingual subtitles are located in a fixed position in the centre of the screen. Therefore, if the users turn their head towards a different part of the scene, they will not be able to read the subtitle. This strategy might result into less freedom of movement, but it might also help avoiding distractions from the main action when the narrative requires the user's attention. The font of the subtitles is sans serif and yellow. Subtitles contain more than 2 lines in many cases, and do not follow segmentation rules.

*Summer Lesson* and *London Heist*, both games for PlayStation VR, use a similar strategy to implement subtitles. In this case, intra- and interlingual subtitles are always displayed in front of the user, at the bottom of the FoV and centred, which is less intrusive for the scene. They both use sans serif fonts in white. The subtitles in *London Heist* include a black background box to facilitate reading. This strategy could be appropriate for immersive environments, because the user has freedom of movement and it would be very complex for the professional subtitler to foresee where the video background could interfere with the reading and, therefore, change the position of the subtitle, as is the case in current subtitling practices for 2D content. Finally, another strategy for implementing subtitles in VR games appears in *Battle Zone*, for PlayStation VR. In this game, subtitles are integrated in the scene, as they would appear in a head-up display. In this example, the subtitles are not obtrusive in the scene because they appear as if they were part of the environment, integrated in a futuristic spaceship.

The previous examples have shown how some critical issues such as subtitling positioning have been implemented in a selection of subtitled immersive content. Others, such as character identification or non-speech information display, have not been addressed at all. However, it is paramount to gather user feedback in order to generate subtitles that can be easily implemented and accepted by end users. Focus groups such as the one presented in the next section can contribute to such an end.

## **5. Gathering user feedback: focus group methodology**

Focus groups were considered appropriate for identifying user needs at the beginning of the ImAc project, before access services in immersive media were actually implemented. A shared methodology was developed for the five focus groups (three on AD and two on SDH) which took place in four different countries, and ethical clearance was obtained. The preparation stage for the focus groups involved two main steps.

To the best of our knowledge, access services such as subtitles or audio description in immersive media were not fully implemented at the time of conducting the focus groups. Therefore, in order to identify the most relevant questions to be posed to participants, it was necessary to define user types and scenarios. Two main user profiles were defined: those creating the services, i.e. professional users (for instance, IT, graphic

designers, subtitlers, audio describers, and SL interpreters), and those consuming the services, i.e. home users (for instance, deaf, hard-of-hearing, blind, low vision users, the elderly). At this stage it was decided that the focus would be mainly on those consuming the services, gathering additional data from a few professionals and opening the door to future research with other professional profiles such as content creators. It was also decided that home users would be advanced, meaning they would have some knowledge or special interest on the technologies being developed.

The focus group presented in this paper was organised by the Catalan Media Corporation (CCMA) in collaboration with Universitat Autònoma de Barcelona and was held on 28 November 2017 at the CCMA premises. The main aim of the focus group was to obtain feedback regarding expectations, recommendations and desires from professional and home users when consuming and editing SDH in 360° videos, as well as SL access services. The focus of this article is on subtitling, so only the results related to SDH will be reported. The results of this focus group reflect the needs of the Spanish SDH audience. Another focus group regarding SDH was carried out in Berlin (Germany) by another project partner for which the results have not been published. The contents of the focus groups were different and so was the target audience (different language and subtitling habits). Therefore, the results were not fully comparable, and it was decided not to include them in the present study.

### **5.1. Participants**

There were 14 participants (6 males, 8 females): 10 advanced home users (6 signers, 4 oralists) and 4 professional users (2 subtitlers, 1 technical expert, 1 representative from a user association). Age range was 21-40 (3), 41-60 (7), and +60 (4). Three participants had secondary education studies, four participants had further education studies, six had university studies and one person did not reply to this question. Three of them reported having a device to access VR content (VCR, glasses and PC, respectively). Mobile phones were the technology most frequently used by the participants on a daily basis (14), followed by TV (14), laptop (12), PC (10) and tablet (8). The advanced home users were deaf (8) and hearing-impaired people (2), most having the disability from birth (4) or when they were between 0-4 years (5) or 41-60 years (1). The preferred devices for watching online video content was PC (7) and laptop (7), followed by smartphone (5), tablet (3) and TV (3).

Even though the recommended group size is up to 10 (Bryman 2004: 507), the profiles of deaf and hard-of-hearing users are diverse (Báez Montero and Fernández Soneira 2010: 26) and it was therefore considered that including the maximum number of home users would provide a more accurate demographic sample. The diversity in the results of the focus group confirmed that this approach was appropriate. Moreover, following standard recommendations on focus groups, it was deliberately decided to

over-recruit in order to allow for no-shows (Wilkinson 1999: 188) and a higher number of short suggestions (Morgan 1998: 75).

## 5.2. Procedure

The focus group included five stages. First, participants were welcomed by the facilitator who briefly explained the aim of the ImAc project. The focus group took place in a meeting room equipped with a table and chairs for the participants and a computer and a large TV screen to show the examples to be discussed. A SL interpreter was present, as well as two researchers who took notes and summarised the conclusions in real time. Secondly, the aim of the focus group was explained to the participants, and they were asked to sign informed consent sheets. The third step was filling in a short questionnaire on demographic information. Finally, the group discussion began. To trigger the discussion, the facilitator gave a short introduction to VR and 360° content and explained how 360° content can be accessed, showing VR glasses to the participants. He explained that 360° content can also be accessed on a flat TV screen using a mouse to move around the 360° scene. As a specific example, an excerpt of the TV show *Polònia* was shown to participants on a flat TV screen. Different types of subtitles were presented to give users some ideas about how SDH could be implemented in immersive content and to stimulate their imagination: subtitles located in a fixed position, subtitles located close to the speaking character, and subtitles located each 120° in the 360° view. The facilitator also posed questions about how users would like to interact with this type of access services and what features a future platform giving access to these services should have. Together with these stimuli, the facilitator also used a list of guiding questions grouped under major topics to generate participants' reactions, taking special care to allow participants to raise aspects that they considered relevant even if not included in the list. A balance between an open-ended and a structured approach was sought, and the result was a lively discussion in which interesting suggestions were made.

As the focus group took place, one researcher was drafting a list of conclusions. Reading these conclusions and agreeing on them was the last step of the focus group, which lasted 90 minutes. At the end of the session, participants were thanked for their participation and they were told about the next steps in the project.

## 6. Focus group results

Data analysis followed a qualitative approach, due to the number of participants and the methodological tool chosen. As explained above, two researchers took notes on a shared document and summarised the conclusions in real time. After the focus group, the notes were thoroughly revised and tagged. This procedure allowed to identify three main areas in which users voiced their views: (1) feedback from advanced home users concerning the services; (2) feedback from advanced home users

concerning the interaction with a future platform giving access to the services; and (3) feedback from professional users concerning content creation. The analysis also allowed to define aspects in which there was consensus among users and aspects in which opinions diverged, as described next.

### **6.1. Advanced home users: services**

In general, home users considered that subtitles in immersive media should be based on approved subtitling rules (AENOR 2003) and, if necessary, improvements might be implemented to adapt existing rules to the new needs posed by immersive environments.

Regarding the position of the subtitles, users suggested that subtitles should always appear in a fixed position in relation to the users' FoV. They also agreed that subtitles should always appear at the bottom — except in specific cases, such as football matches. There was a brief discussion about the possibility of customising the position of the subtitles. It was even suggested that the placement of subtitles in real time should be changed while watching the 360° video. However, users finally disregarded this option, since they all agreed that subtitles at the bottom of the FoV was the most comfortable solution. It remains to be seen whether research will actually confirm this is the best solution.

Most participants were concerned about the fact that sometimes the subtitles could not be read because of the background image. They stated that it is important, therefore, to have the possibility to choose subtitles with a black background box to facilitate the reading. Also, some participants expressed their worry about the fact that subtitles in immersive media could be disruptive if they appear in a close-up or some other scenarios where the subtitle is obstructing the image. They said that subtitle editors should pay special attention to avoid disrupting the immersive experience.

For character identification, users stated that it is necessary to maintain colour coding to identify characters, as this is already done in SDH for 2D content.

Concerning the display of non-speech information (sounds, music, paralinguistic information, etc.), different options were proposed. In general, users requested that basic subtitling elements that have been previously approved in the regulations (for example, how to indicate music) should be retained. However, they accepted that new technologies may bring new possibilities. Some users preferred to receive non-speech information in the form of text in brackets, as is now the case in most subtitled TV programmes. Other users, considering the new technology in use, preferred to receive this information as icons. In that sense, users suggested the possibility of using a closed list of icons. For example, a lightning icon to indicate the sound of a storm. Regarding the position of non-speech information, users did not reach a consensus. Some stated

that they preferred them to be located at the top, others at the bottom close to the subtitle (dialogue), and others would like to move them to a different location. In general, participants did not like the idea of having non-speech information at the top, as it is stated in the current Spanish UNE rule for SDH, because they do not have time to read both the subtitle (at the bottom) and the non-speech information (at the top). They suggested to change this in immersive environments and place the non-speech information in form of icons or text between brackets close to the subtitle area, stating that this would be easier to process. Also, some hard-of-hearing participants stated that they do not need non-speech information in the subtitles, and they would prefer to deactivate them if possible. In this sense, users would like to be able to customise the position of non-speech information.

Users also considered the challenges that the new dimension brought by immersive media (i.e. space and the need to indicate directions) would entail when it comes to SDH in immersive content. In that sense, users stated that it was difficult to know where to look to see the character speaking. They considered that the subtitle should indicate how you need to move your head (four directions), with icons (arrows), indicators between brackets (to the left, to the right) or some sort of mechanism. It was suggested that a compass or radar could be used to that end and that it should be always visible on the screen. Participants also agreed that the radar or compass should be close to the subtitle, otherwise it could be distracting.

As for the subtitle content, users insisted that it should include all the information, both on screen and off screen; in other words, dialogues taking place both within and outside the user's FoV. They suggested that this could be indicated with ON and OFF tags. They also stated that there are different needs among users and, consequently, subtitles must be adapted to different profiles. For example, there could be different levels of speed (faster/slower). However, users considered that summarised or simplified subtitles do not generally help deaf people, because this type of subtitle make it more difficult to follow the AV content. Nevertheless, they conceded that simplified subtitles may be useful for users with other types of needs and could be considered an alternative. It was clear that user profiles are diverse, and customisation should be a priority.

## **6.2. Advanced home users: platform interaction**

Users were asked about the options and features that they would like to have in an immersive platform which would give access to virtual content with access services implemented. At the time of the focus group, no prototype was available. Therefore, thought-provoking questions were presented to participants based on hypothetical user scenarios.

Regarding interaction with an immersive interface, users positively valued the possibility of personalisation, i.e. having different layers that could be activated or not. For example, some participants preferred subtitles only

for dialogues, others needed non-speech information indications and others wanted to have as much indications as possible, including directions. However, some elements were not considered in need of customisation, such as the position of the subtitles, which it was reported should always be at the bottom of the field of view, because they considered it would be easier to read. Moreover, both professional and home users considered that the user should customise this future platform the first time but then those parameters should be recorded by the interface for future use. Users also suggested that this customisation should be transferrable from one device to the other (importing profile, that is, the user profile could be imported) and they requested the possibility of creating more than one profile. They also considered the possibility of transferring a profile from the user's device to another external device (for example, at a friend's home).

Regarding interaction with access services, users positively valued the possibility of alternative interactions (for example, voice commands), although they did not find it necessary for their specific needs and indicated that implementation costs should be taken into account. However, they added that if this platform were to be developed for other profile types (for example, blind users), it could be an additional resource.

Regarding companion screens, participants liked the possibility of using the smartphone to interact with the platform as a touch-screen (like a "mouse") and to customise their preferences. One user even suggested the possibility of including a finger-sensor that would allow users to see their own fingers on the virtual image. There were different opinions regarding the need of reproducing the same content on the smartphone screen, since the smartphone is often used as an element to access additional content. When accessing AV content together with other people, users did not want to consume subtitles on a different screen because this made them feel excluded.

### **6.3. Professional users: SDH creation**

Professional users expressed their proposals regarding the production of SDH in immersive environments. They agreed that vertical positioning of the subtitles could be an interesting option for separating dialogue subtitles from non-speech information, although they considered that home users must be able to decide or set up where they prefer to locate the subtitle.

Regarding the production of subtitles, they stated that they preferred an on-screen display (player) showing one dynamic angle of the 360° view, so that they could choose which angle to see using cursors or mouse movements. Professional users considered that they should be able to test the results with both HMD and flat screen (for instance, a PC screen).

Regarding the subtitling tool, users indicated that they would need a subtitling editor similar to the existing ones for SDH, but it should add the

360° displaying and the possibility of adding emoticons and text messages to show sound actions that take place parallel to the dialogue subtitles. They also highlighted the need for the editor tool to offer original 360° immersive audio because it is important to identify where the sound/dialogues come from, as this information is requested by end users.

## **7. Conclusions**

This article has described the role of immersive media in our society and has put forward the need to make them accessible to all users. The emphasis has been put on how subtitles can be integrated in immersive environments, with reference to the limited existing research and practice. Adopting a user-centred approach, the results of a focus group on SDH developed as part of the ImAc project have been presented. According to participants' feedback, SDH for 360° videos should: (1) be located in a fixed position and always visible in relation to the FoV and preferably at the bottom; (2) have a background box to avoid contrast issues with an unpredictable background; and (3) include a system to indicate directions when the speaker is outside the FoV, such as arrows, a compass or text between brackets. Also, results show that home users are willing to accept the implementation of new features in SDH in immersive content, such as icons for non-speech information, because of the new possibilities and dimensions brought by this medium. Moreover, customisation options appear to be a desirable feature among participants. Users also show their agreement and interest in continuing established practices and regulations for SDH, such as the Spanish subtitling standard UNE 153010 (AENOR 2003). However, they agree in introducing some changes to improve the current standards. For example, they would like to have the non-speech information or direction information closed to the subtitle area and not at the top as it currently is, in order to avoid distractions.

One of the limitations of the present study is that it only applies to the Spanish audience and should be replicated in other countries to confirm the validity and generalisation of the conclusions. Another focus group was carried out in Germany for the ImAc project,<sup>1</sup> with similar results (German participants would also like the subtitles always visible in the FoV and also suggested using an arrow to indicate the location of the speaker). However, the contents and examples were not the same, because the audiences spoke different languages and, therefore, the results are not comparable. It was not the intention of the project to compare the two focus groups, but rather gather a general feedback from end users that would set the basis to start developing a prototype for access services in 360° videos. The prototype will be later tested with a larger number of participants in the next stages of the project.

Therefore, the next step will be to transfer user feedback into user requirements and implement the features in immersive content in order to verify whether it is technically possible or whether there are limitations.

Once implemented, user testing will be necessary to verify or reject the validity of the proposed SDH models.

In conclusion, it is clear that research is needed in the field of MA for immersive media. For this purpose, the ImAc project will be a perfect laboratory environment for the development of a successful SDH model for immersive content. This is a significant step in the field of AVT and MA, since the design of accessibility will be taken into account before the technology is fully implemented in society.

## Acknowledgements

ImAc has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 761974. The authors are members of TransMedia Catalonia, a research group funded by Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya, under the SGR funding scheme (ref. code 2017SGR113). This article reflects only the authors' views and the funding institutions hold no responsibility for any use that may be made of the information it contains). This article is part of Belén Agulló'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 Universitat Autònoma de Barcelona.

We would like to thank CCMA, and especially Francesc Mas and Jordi Mata, who were involved in the focus group management at the CCMA premises, and to all participants who took part in the focus group.

## References

- **Agulló, Belén and Pilar Orero** (2017). "3D Movie Subtitling: Searching for the best viewing experience." *CoMe - Studi di Comunicazione e Mediazione linguistica e culturale* 2, 91-101.
- **Arnáiz-Uzquiza, Verónica** (2012). *Subtitling for the deaf and the hard-of-hearing: Some parameters and their evaluation*. PhD thesis. Universitat Autònoma de Barcelona. <https://www.tdx.cat/handle/10803/117528> (consulted on 08.01.2019).
- **Arnáiz-Uzquiza, Verónica** (2015). "Eye tracking in Spain." Pablo Romero-Fresco (ed.) (2015). *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe*. Bern: Peter Lang, 263-287.
- **Arrés, Eugenia** (forthcoming). "Traducir en entornos de realidad virtual: cómo garantizar la inmersión del usuario." To appear in *New Technologies Applied to Translation Studies: strategies, tools and resources*.
- **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*. Madrid: Asociación Española de Normalización y Certificación.
- **Azuma, Ronald** (1997). "A Survey of Augmented Reality." *Presence: Teleoperators and Virtual Environments* 6(4), 355-385.

- **Azuma, Ronald et al.** (2001). "Recent Advances in Augmented Reality." *IEEE Computer Graphics and Applications* 21(6), 34-47.
- **Báez Montero, Inmaculada Concepción and Ana María Fernández Soneira** (2010). "Spanish deaf people as recipients of closed captioning." Anna Matamala and Pilar Orero (eds) (2010). *Listening to Subtitles: Subtitles for the Deaf and Hard of Hearing*. Frankfurt: Peter Lang, 25-44.
- **Bartoll, Eduard and Anjana Martínez-Tejerina** (2010). "The positioning of subtitles for the deaf and hard of hearing." Anna Matamala and Pilar Orero (eds) (2010). *Listening to Subtitles: Subtitles for the Deaf and Hard of Hearing*. Frankfurt: Peter Lang, 69-86.
- **British Broadcasting Corporation (BBC)** (2012). "Binaural Sound: Immersive spatial audio for headphones." <http://www.bbc.co.uk/rd/projects/binaural-broadcasting> (consulted on 19.12.2017).
- **British Broadcasting Corporation (BBC)** (2014). "360 Video and Virtual Reality: Investigating and developing 360-degree video and VR for broadcast-related applications." <http://www.bbc.co.uk/rd/projects/360-video-virtual-reality> (consulted on 19.12.2017).
- **Belton, John** (2012) "Digital 3d cinema: Digital cinema's missing novelty phase." *Film History: An International Journal* 24(2), 187-195.
- **Biocca, Frank** (1992). "Virtual Reality Technology: A Tutorial." *Journal of Communication* 42(4), 23-72.
- **Brown, Andy et al.** (2017). "Subtitles in 360-degree Video." Judith Redi et al. (eds) (2017). *TVX '17 ACM International Conference on Interactive Experiences for TV and Online Video* (Hilversum, The Netherlands, 14-16 June 2017). New York: ACM, 3-8.
- **Brown, Andy et al.** (2018). "Exploring Subtitle Behaviour for 360° Video. White Paper WHP 330." <https://www.bbc.co.uk/rd/publications/whitepaper330> (consulted on 13.02.2019).
- **Bryman, Alan** (2004). *Social Research Methods*. New York: Oxford University Press.
- **Civera, Clara and Pilar Orero** (2010). "Introducing icons in subtitles for deaf and hard of hearing: Optimising reception?" Anna Matamala and Pilar Orero (eds) (2010). *Listening to Subtitles: Subtitles for the Deaf and Hard of Hearing*. Frankfurt: Peter Lang, 149-162.
- **Conroy, Andy** (2017). "The BBC and Virtual Reality." *BBC Research & Development*. <http://www.bbc.co.uk/rd/blog/2016-06-the-bbc-and-virtual-reality> (consulted on 21.12.2017).
- **De la Peña, Nonny et al.** (2010). "Immersive Journalism: Immersive Virtual Reality for the First-Person Experience of News." *Presence: Teleoperators and Virtual Environments* 19(4), 291-301.
- **Deloitte** (2016). *Technology, Media & Communications Predictions*. <https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Technology-Media-Telecommunications/gx-tmt-prediction-2016-full-report.pdf> (consulted on 20.12.2017).
- **European Broadcasting Union (EBU)** (2017a). *Virtual Reality: How are public broadcasters using it?* <https://www.ebu.ch/publications/virtual-reality-how-are-public-broadcasters-using-it> (consulted on 18.12.2017).
- **European Broadcasting Union (EBU)** (2017b). *Safe Areas for 16:9 Television Production*. <https://tech.ebu.ch/docs/r/r095.pdf> (consulted on 31.01.2018).
- **González-Zúñiga, Diekus, Carrabina, Jordi and Pilar Orero** (2013). "Evaluation of Depth Cues in 3D Subtitling." *Online Journal of Art and Design* 1(3), 16-29.

- **Greco, Gian Maria** (2016). "On Accessibility as a Human Right, with an Application to Media Accessibility." Anna Matamala and Pilar Orero (eds) (2016). *Researching Audio Description. New Approaches*. London: Palgrave Macmillan, 11-33.
- **Hall, Charli** (2017). *Watch VR's first Oscar-nominated short film in Polygon*. <https://www.polygon.com/2017/1/24/14370892/virtual-reality-first-oscar-nominated-short-film-pearl> (consulted 24.10.2017).
- **Holliman, Nicolas et al.** (2011). "Three-dimensional displays: a review and applications analysis." *IEEE Trans. Broadcast* 57(2), 362–371.
- **Jayaraj, Lionel, Wood, James and Marcia Gibson** (2017). "Improving the Immersion in Virtual Reality with Real-Time Avatar and Haptic Feedback in a Cricket Simulation." Wolfgang Broll et al. (eds) (2017). *Adjunct Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality*. Los Alamitos, California: Conference Publishing Services, 310-314.
- **Jensema, Carl, McCann, Ralph and Scott Ramsey** (1996). "Closed-captioned television presentation speed and vocabulary." *American Annals of the Deaf* 141(4), 284-292.
- **Jiménez Hurtado, Catalina, Seibel, Claudia and Silvia Soler Gallego** (2012). "Museos para todos. La traducción e interpretación para entornos multimodales como herramientas de accesibilidad universal." *MonTI. Monografies de Traducció i Interpretació* 4, 349-383.
- **Jones, Sarah** (2017). "Disrupting the narrative: immersive journalism in virtual reality." *Journal of Media Practice* 18(2–3), 171–185.
- **Lambooi, Marc, Murdoch, Michael, Ijsselsteijn, Wijnand and Ingrid Heynderickx** (2013). "The impact of video characteristics and subtitles on visual comfort of 3D TV." *Displays* 34(1), 8-16.
- **Lee, Nicole** (2017). "Pixar's VR debut takes you inside the entrancing world of 'Coco'". *Engadget*. <https://www.engadget.com/2017/11/15/pixar-coco-vr/> (consulted on 21.12.2017).
- **Mangiron, Carme** (2013). "Subtitling in game localisation: a descriptive study." *Perspectives: Studies in Translation Theory and Practice* 21(1), 42-56.
- **Matamala, Anna and Pilar Orero** (eds) (2010). *Listening to Subtitles: Subtitles for the Deaf and Hard of Hearing*. Bern: Peter Lang.
- **Matamala, Anna et al.** (2018). "User-centric approaches in Access Services Evaluation: Profiling the End User." *Proceedings of LREC Workshop "Improving Social Inclusion using NLP: Tools, Methods and Resources", ISINLP2 2018*. <https://ddd.uab.cat/record/189707> (consulted on 13.02.2019).
- **Milgram, Paul and Fumio Kishino** (1994). "A taxonomy of mixed reality visual displays." *IEICE Transactions on Information Systems* 77, 1321–1329.
- **Morgan, David L.** (1998). *Planning Focus Groups*. Thousand Oaks: Sage.
- **Neves, Joselia** (2005). *Audiovisual translation: Subtitling for the deaf and hard of hearing*. PhD Thesis. Roehampton University. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.129.1405&rep=rep1&type=pdf> (consulted on 08.01.2019).
- **Neves, Joselia** (2018). "Cultures of Accessibility: Translation making cultural heritage in museums accessible to people of all abilities." Sue-Ann Harding and Ovidi Carbonell Cortes (eds). *The Routledge Handbook of Translation and Culture*. London/New York: Routledge, 415-430.

- **Oculus Team** (2016). "Oculus Film Short "Henry" Wins an Emmy!" *Oculus Blog*. <https://www.oculus.com/blog/oculus-film-short-henry-wins-an-emmy/> (consulted on 24.10.2017).
- **Orero, Pilar and Anna Matamala** (2007). "Accessible opera: Overcoming linguistic and sensorial barriers." *Perspectives: Studies in Translation Theory and Practice* 15(4), 262–277.
- **Perego, Elisa, Del Missier, Fabio and Sara Bottioli** (2015). "Dubbing versus subtitling in young and older adults: cognitive and evaluative aspects." *Perspectives: Studies in Translation Theory and Practice* 23(1), 1-21.
- **Remael, Aline, Orero, Pilar and Mary Carroll** (eds) (2014). *Audiovisual Translation and Media Accessibility at the Crossroads*. Amsterdam/New York: Rodopi.
- **Romero-Fresco, Pablo** (2009). "More haste less seed: Edited vs. verbatim respoken subtitles." *VIAL (Vigo International Journal of Applied Linguistics)* 6, 109-133.
- **Romero-Fresco, Pablo** (2013). "Accessible filmmaking: Joining the dots between audiovisual translation, accessibility and filmmaking." *The Journal of the Specialised Translation* 20, 201-223.
- **Romero-Fresco, Pablo** (ed.) (2015). *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe*. Bern: Peter Lang.
- **Sala Robert, Elia** (2016). *Creative subtitles. Subtitling for all*. PhD Thesis. Universitat Pompeu Fabra. <https://repositori.upf.edu/handle/10230/27737> (consulted on 08.01.2019).
- **Sherman, William and Alan Craig** (2003). *Understanding Virtual Reality: Interface, Application, and Design*. San Francisco: Morgan Kaufmann Publishers (An imprint of Elsevier Science).
- **Sketchfab** (2017). *VR Industry Trends: A comprehensive overview on the state of the VR Industry in Q2-2017*. <https://sketchfab.com/trends/q2-2017> (consulted on 18.12.2017).
- **Steuer, Jonathan** (1992). "Defining Virtual Reality: Dimensions Determining Telepresence." *Journal of Communication* 42(4), 73-93.
- **Szarkowska, Agnieszka et al.** (2016). "Open Art: Designing Accessible Content in a Multimedia Guide App for Visitors with and without Sensory Impairments." Anna Matamala and Pilar Orero (eds). *Researching Audio Description. New Approaches*. London: Palgrave Macmillan, 301-320.
- **Udo, John Patrick and Deborah Fels** (2011). "From the describer's mouth: reflections on creating unconventional audio description for live theatre." Adriana Serban, Anna Matamala and Jean-Marc Lavaur (eds). *Audiovisual Translation in Close-Up: Practical and Theoretical Approaches*. Bern: Peter Lang, 257-278.
- **Vilaró, Anna** (2011) "Does 3D stereopsis change the way we watch movies? An eye-tracking study". Paper presented at *4th International Conference Media for All* (London, 28 June-1 July 2011).
- **VR Intelligence** (2017). *Virtual Reality Industry Survey*. <http://vr-intelligence.com/vrx/docs/VRX-2017-Survey.pdf> (consulted 19.12.2017).
- **Wilkinson, Sue** (1999). "Focus Groups Methodology: A Review." *International Journal of Social Research Methodology* 1(3), 181-203.
- **Wohlsen, Marcus** (2015). "Google Cardboard's New York Times Experiment Just Hooked a Generation on VR." *Wired Magazine*. <https://www.wired.com/2015/11/google-cardboards-new-york-times-experiment-just-hooked-a-generation-on-vr/> (consulted on 21.12.2017).

## Biographies



**Belén Agulló** is a predoctoral researcher in the Department of Translation, Interpreting and Eastern Asian Studies at Universitat Autònoma de Barcelona, where she is working on the Horizon 2020-funded project Immersive Accessibility (ImAc). Her PhD focus is subtitling for the deaf and hard-of-hearing in immersive media. Previously, she worked for more than 5 years in the game localisation industry. She teaches game localisation in different masters in Spain and France. Her research interests include audiovisual translation and media accessibility.

E-mail: [belen.agullo@uab.cat](mailto:belen.agullo@uab.cat)



**Anna Matamala**, BA in Translation (UAB) and PhD in Applied Linguistics (UPF), is an associate professor at the Universitat Autònoma de Barcelona. Currently leading the TransMedia Catalonia group, she has participated and led projects in audiovisual translation and media accessibility. She has taken an active role in the organisation of scientific events (M4ALL, ARSAD), and has published in journals such as *Meta*, *Translator*, *Perspectives*, *Babel*, *Translation Studies*. She is currently involved in standardisation work.

E-mail: [anna.matamala@uab.cat](mailto:anna.matamala@uab.cat)

---

## Notes

<sup>1</sup> Results for the German focus group can be found in the public report: <http://www.imac-project.eu/documentation/deliverables/>.