6

Towards a user specification for immersive audio description

Chris Hughes, Pilar Orero and Sonali Rai

Abstract

This chapter reports on an initial study conducted as part of the EU-funded Immersive Accessibility (ImAc) project, whose broad objective it is to define the requirements for accessibility services such as audio description and audio subtitles within immersive environments. This study follows a user-centric research design approach. A focus group comprising blind and partially sighted people was conducted to establish end-user expectations, recommendations and perceptions in order to elicit first principles for developing system requirements. The chapter concludes with recommendations for the initial user specification for the ImAc project.

6.1 Introduction

The aim of the Immersive Accessibility project is to establish how traditional access tools such as subtitles, audio description, sign language and audio subtitles may be used to transform the experience of consuming immersive content for people with hearing and/or sight loss. In contrast to the controlled environments designed for the delivery of mass media, ImAc will offer a fully customisable platform to present immersive content to users with varying degrees of sensory loss and learning disabilities. A range of open source tools will also be designed for the content production industry and showcased at the final project conference.

The ImAc Consortium brings together stakeholders from organisations across sectors including broadcast, charitable and volunteer organisations, academia, research and development with the aim of combining a range of different yet relevant skills to deliver on this project.

The project aims to create accessible and fully personalised services for all citizens using access features such as subtitles, audio description and sign language to transform the accessibility of immersive content. This will develop novel resources for the broadcasting industry, providing adapted content and ensuring accessibility in immersive environments. In order to ensure content is accessible across platforms, broadcasters need resources and tools which can be integrated with current production workflows and allow the repurposing of existing content whenever possible. ImAc will provide new, and enhance existing, tools to deliver access features for 360-degree content.

The ImAc project will also demonstrate the newly developed tools and platforms through open access pilot testing. The execution and evaluation of pilot studies is crucial to ensure that the developed services, tools and the platform are functional and can be operated in real-world scenarios. Therefore within the project there will be pilot studies in three countries: the UK, Spain and Germany. This will help to standardise accessibility data in immersive content environments. For the inclusion of accessibility services in immersive content, it is important that standardisation bodies acknowledge the need for these services and define requirements for their delivery from an early stage.

Finally, ImAc aims to maximise the impact on society by delivering useful solutions. The project development will be driven by real user needs and will seek to continuously involve users by means of user-centric design methodologies. Tools developed for this purpose should meet the requirements of experienced broadcasters from the start, and tools and services are to be thoroughly tested by means of pilot studies.

6.2 Literature review

Audio description (AD) research is usually conducted within the field of audiovisual translation. Generally, this is because the audio describer has a similar function to that of the language translator, who has to render the work written in a source language (multimodal text) into a target language (AD). The audio describer must provide an account of the visual image and disambiguate any meaningful sound which is not explicit (Szarkowska and Orero, 2014[BIB-029]). The image to be described may vary from a movie to a play, an opera or a picture hung in a museum. In the case of a museum picture the AD renders not only the picture, but also where it hangs, the frame, its size and so on. Matching user expectations to existing audio description production and guidelines is an interesting issue with much research (Maszerowska *et al*., 2014[BIB-017]). Classifying AD according to its genre has also been studied, mostly regarding TV (Van der Heijden, 2007[BIB-034]), cinema (Fryer and Freeman 2012a[BIB-008], 2014[BIB-011]; Fryer *et al*., 2013[BIB-012]; Perego, 2014[BIB-026]), comedy (Fels *et al*., 2006[BIB-007]), Bollywood (Rai, 2009[BIB-028]) and opera (Matamala and Orero, 2008[BIB-018]). Other authors have narrowed down the analysis to film components such as lighting (Maszerowska, 2012[BIB-016]), secondary details (Orero and Vilaró, 2012[BIB-022]), film credits (Matamala and Orero, 2011[BIB-019]), leitmotifs (Vilaró and Orero, 2013[BIB-035]) and sound (Szarkowska and Orero, 2014[BIB-029]; Orero *et al*., 2016[BIB-024]). Udo and Fels (2009[BIB-033]) refined their research to audio describing not only a play but, more specifically, audio describing Shakespeare.

On the other hand authors have studied the creation of AD, this time looking at the adequacy of the target text with the audience, considering general audience engagement (Afonso *et al*., 2010[BIB-001]; Chmiel and Mazur, 2012[BIB-004], 2016[BIB-005]; Fryer and Freeman, 2012b[BIB-009], 2014[BIB-011]; Wilken and Kruger, 2016[BIB-039]; Wissmath and Weibel, 2012[BIB-040]; Walczak, 2017[BIB-038]). Some authors have limited the audience to children (Schmeidler and Kirchner, 2001[BIB-030]; Palomo López, 2008[BIB-025]; Orero, 2012[BIB-021]; Krejtz *et al.,* 2012[BIB-015]; Starr, 2018[BIB-031]). Other studies have proposed changes to the target audio described text to aid comprehension (Cabeza-Cáceres, 2013[BIB-003]) and engagement (Walczak and Fryer, 2017[BIB-038]); and a recent study (Starr, 2018[BIB-031]) developed AD as a bespoke accessibility service for a cognitively challenged target audience of autistic individuals needing help with emotion identification.

Regarding the format of the audio description, Walczak (2018) and Knigge and Erkau (2014[BIB-014]) looked at the reception of AD on smart phones, but very little research has been performed regarding AD in other formats such as virtual reality or 360-degree scenarios. This can be considered the first publication dealing with such a complex environment, and points to the many lines of research opening in this field. Research to date has focused on many aspects of the technology, reception and production of AD where, until now, an image—or a scene—has been matched by one dedicated audio description oral text with its unidirectional associated soundtrack. The new format developed in the ImAc project will allow for multiple audio descriptions for an image – or scene – according to the end user relative position. The associated soundtrack will also change, offering additional information regarding environmental conditions such as directionality and distance. Finally, user interaction with the new audiovisual content and its audio description will also change, from the traditional passive audience to an interaction where the audience has a choice in what is consumed.

6.3 Research methodology

Following a user-centred design approach (Norman, 2013[BIB-020]) enables us to engage with the requirements of the UN Committee on the Rights of Persons with Disabilities (CRPD)[1] where the maxim 'nothing about us without us' was established. It provides the classification of a 'person with disability' by referencing a medical model, defined within the International Classification of Functioning, Disability and Health (ICF, 2001). This provides a classification within health and the health-related domains. As the functioning and disability of an individual occur within a societal context, ICF also includes a list of environmental factors to consider, and these are implemented through a standard questionnaire issued by the UN World Health Organisation's WHODAS 2.0 (WHO, 2001), which is primarily related to clinical issues. The UN approach profiles users through a single health-related marker. Within our research disability is not a health-related issue but a communication condition, i.e. an impairment (Ellis, 2016[BIB-006]) not always related to a disability. For example education will have a greater impact on the overall development of a person with sight impairments than a diagnosis of inherited *retinitis pigmentosa*, which will help define the type and level of adjustment that will be needed to achieve the educational goal.

User-centred design (UCD) is a philosophy that seeks to place the end user at the centre of the design process (Orero and Matamala, 2016[BIB-023]). From the field of engineering, Norman (2013[BIB-020]) provided a set of guidelines that designers could follow in order for their interfaces to achieve good usability outcomes. In addition, the

ISO standard 9241-210 (ISO, 2010[BIB-013]) deals with UCD to 'address impacts on a number of stakeholders, not just those typically considered as users', referring to the approach as human-centred design. This standard defines human-centred design as 'an approach to systems design and development that aims to make interactive systems more usable by focusing on the use of the system and applying human factors/ergonomics and usability knowledge and techniques'. It also describes the potential benefit of following a design approach that improves usability and human factors: 'Usable systems can provide a number of benefits, including improved productivity, enhanced user well-being, avoidance of stress, increased accessibility and reduced risk of harm.' Putting the user at the core of the design process is also the guiding principle of universal design (Story *et al.*, 1998[BIB-032]), where the aim is to create accessible products, environments and services for all users regardless of their physical or cognitive abilities, thereby satisfying the requirements of the United Nations' CRPD.

UCD has four defined activity phases: (i) identify the user and specify the context of use; (ii) specify the user requirements; (iii) produce design solutions; and (iv) evaluate design solutions against requirements. This chapter makes reference to the first phase.

6.4 Focus group

This study focuses on the accessibility of 360-degree video content from an AD perspective, as tested by a focus group of blind and partially sighted people. The aim of the focus group was to gather feedback from regular users of AD on viewing and interacting with 360-degree content within an immersive environment.

A review within the ImAc project of immersive environments highlighted several key areas on which feedback was sought during the focus group. Firstly, we needed to consider the differences between the traditional curated linear approach where the viewer consumes the content that the director wishes to show on the screen, and a non-linear approach where the viewer has control over the content and can choose their own path by interacting with the environment. We also needed to understand how using audio could both enhance the immersive experience and address issues surrounding the creation of a fully accessible interface for 360-degree content.

In the absence of any commercially available audio described 360-degree video content, live description was delivered in a clip available on YouTube in a 360-degree format. We wished to explore how a small group of participants would respond to the material. During the test the participants were first shown each video, followed by a period of discussion. A professional audio describer and a technology expert were both present during the entire focus group session and contributed to steering the discussions.

There were 7 participants within the group split between female (3) and male (4) and ranging in age from 26 to 53 years (average age, 38 years). All of the participants are native English speakers and had achieved a university level of education. Five of the participants described themselves as blind, whereas two participants defined themselves as low-sighted. All of the participants stated that they had experienced the disability from birth.

6.4.1 Technology preferences

A survey of users' technology preferences was undertaken before the trial commenced. All participants responded that they use both a television and a mobile phone on a daily basis. However, when further questioned about computer usage, six of the participants claimed to use a laptop and a tablet each day, whereas only one participant stated that they used a desktop computer.

Before participating in our study, none of the participants had ever had access to VR technology or experienced VR content. In terms of consuming online video content, six of the participants showed a strong preference towards using smartphones and tablets, whereas one participant claimed to prefer using a desktop computer for this purpose.

All participants were asked about their preferences for using assistive technologies such as screen readers (JAWS, VoiceOver, TalkBack) and magnification tools (for instance, Zoomtext). Of the seven participants, five used only screen readers, one used only magnification tools, and one used a combination of both. All participants told us that when interacting with online media they would prefer assistive technology to be able to help them identify content, browse content from a library and switch accessibility services, such as AD and subtitles, on and off. They also indicated that it is essential for the assistive technology to interact with functions such as play, stop, pause, forward and rewind.

6.4.2 Sight level

For this study, it was essential to understand the current level of sight that each of the participants could experience. The study group contained a complete range of abilities. Two of the participants could not see anything at all; five participants could identify the windows in the room based on the light. Three participants could identify the shapes of furniture in a room. One participant was able to identify a friend over a metre away, whereas another participant could only recognise a friend if they were in close proximity to their face.

Further, the participants were asked to describe the barriers they experienced whilst watching television. All of the participants expressed difficulty in seeing the buttons on a remote control. None of the participants was able to see fine detail and text on the television screen and everyone stated that they found it difficult to understand what was happening onscreen. More specifically, six of the participants had difficulty seeing a picture on the screen at all; in fact, five participants could only identify the television screen by the light it emitted; and two participants could see nothing at all.

The sight level had a significant impact on the participants' current ability to watch and follow programmes and films on television. One participant told us that they are able to follow the programme using residual sight, whereas another said that their solution was to sit closer to the screen. Six of the group generally attempt to pick up as much as they can about the programme from the sound alone, but they stressed that this was not a preferred option. They also told us that they regularly rely upon friends or family to explain what is happening, or use AD to work out what is happening, on the screen.

6.5 Results

The participants were given four tasks to complete, designed to help with our understanding of the way AD fits within the context of immersive content. Each participant viewed the 360-degree videos on a laptop connected to a projector. The 360-degree rendering was provided by the ImAc player and the audio recorded in order to be transcribed into the results at a later date. The focus group lasted for about one hour.

6.5.1 AD within linear storytelling

The first task was designed to inform our understanding of AD's role within linear storytelling and the manner in which the interactive aspects of 360-degree videos impact an AD track. In conventional television we assume that the consumers are looking directly at a TV screen and therefore that the AD describes what they are seeing on the screen in front of them. However, within a 360-degree video the viewer can be looking in any direction. Therefore, this task was designed to determine the importance of AD in an environment where the viewer is able to choose the line of vision. In this task the participants were guided along a set path. This meant that the choice of view was predetermined for the participant in a similar way to watching traditional directed television.

The video, available on YouTube (BBC Earth, 2016[BIB-002]), is a documentary titled *Attenborough and the Giant Dinosaur*, and was created for BBC One's BBC Earth Unplugged series. The participants were required to watch the entire video (run time, four minutes). The AD was delivered 'live' during our pilot study by a professional audio describer who worked without a script, although prior knowledge was employed, as were notes relating to the video content.

6.5.1.1 Feedback

Participants praised the live audio description, and all participants believed that the track gave them enough information to understand the complete picture, even the components which lacked alternative audio cues. During the discussion each participant provided feedback on their experience of the AD, such as:

Participant 1: The important thing is the description complemented the narration, it wasn't repetitive and it wasn't overly descriptive.

Participant 3: I think it managed to describe everything that was on the screen even though there wasn't much time.

It was interesting to observe that even though participant 3 was sight-impaired and therefore could not confirm that everything on the screen had been described, they believed that the AD contained enough information to get an understanding of everything that would be on the screen. Although the description was appreciated in terms of making the story clear, some felt it lacked the elements to build the atmosphere needed for an immersive experience:

> *Participant 4*: What I didn't get from the description was the ambience – you hear the rustling of the foliage, large clomping feet, but what's the weather like, what birds can you see in the sky? This is where your imagination comes in.

> *Participant 5*: I still can't get a proper visual picture of the dinosaur in my head, so the colour, has it got a spiky back?

In relation to these comments the describer, who as an expert was present throughout the focus group, pointed out that the lack of sufficient gaps in narration meant there was a need to prioritise what could be described. It was identified that none of the participants wanted the AD to go over the voiceover of the documentary.

6.5.2 AD within the interactive aspects of 360-degree video

The second task was focused on understanding the impact of AD when the viewer is able to choose their own path through the immersive video. The personalisation facilitated by 360-degree video means that the viewer may choose not to look at the key action and therefore the AD was adapted in the moment ('live') to describe what the viewer was actually looking at. The video from the first task was used again, and the AD was delivered 'live' by the same professional audio describer, who had previously made notes on the video.

In order to simulate a 360-degree environment, a cursor provided within the video was used to allow the participant to choose direction by using the mouse. For example instead of watching the presenter gaze at the dinosaur during the first break in voiceover, the view panned left to the mountains in the distance and consequently the audio description changed as it referred to the on-screen elements now in view. This resulted in participants mostly feeling disconnected with the storyline, finding the description to be repetitive and missing important details.

Participant 2: I thought that sounded disjointed and for me if it is a description of 360 then it would take a lot more than just words. This didn't give me anything really.

While the 360-degree movement seemed to enhance the overall immersive experience visually, it was almost impossible to simulate that experience in audio given the lack of sufficient gaps in voiceover. Some participants commented that it was difficult to comprehend a 360-degree view.

Participant 4: I don't understand, was it vertical now, like a vertical axis, looking down? Was that the view? Was it from the perspective of the dinosaur?

Participant 5: For some people who can't see and never have, it is already a challenge to understand 360.

Participant 6: I don't believe this! Where is the TV screen? I can't get my head around it.

Participants agreed that a lot more information would be needed in addition to AD to make the environment more immersive for people with significant sight loss. In order to be able to navigate the 360-degree video they would need to know more information about the environment and therefore have the cues to tell them where to direct their gaze. Getting the right balance without this information would be difficult.

6.5.3 Spatial audio within 360-degree environment

The third task used spatial audio (binaural sound) to create the illusion of a 360-degree environment by providing sound from all around the participant. There was no video present so the focus was on whether the participants could be fully immersed within a scene using only the sound cues.

The demonstration example used was the *Virtual Barbershop* sound file provided by QSoundLabs (2007[BIB-027]), which had been designed as a high-quality demonstration of binaural audio technology. The audio places the participant in the seat at a barbershop, where the barber is making conversation whilst walking around the participant cutting their hair. It is an example which worked well in our focus group, as the participants were all familiar with sitting in a barber's chair, where customers are encouraged not to move their heads. Many of the sound cues would also have been familiar to the focus group participants, such as the snipping of hair.

Two stereo speakers were employed to allow the focus group to participate as a group. Participants responded enthusiastically to the clip:

> *Participant 3*: For me this was immersive!

> *Participant 4*: You're actually in the room!

> *Participant 6*: It was very impressive that last one, I found it very interesting. I know who's around this table because I'm in the room. Felt as if it was happening around me!

> *Participant 7*: Honestly, that's what I need to be completely immersed.

It was noted that the narration in the clip integrated elements of audio description within the soundscape, for example "now I'm moving to the right" and "look at my pair of scissors". These cues, combined with the perception of depth delivered in spatial audio, enhanced the experience for the focus group.

6.5.4 Establishing user preference for tools required to access 360-degree content

Whilst discussing the significance of the headset, which would be needed to track head movements and subsequently trigger specific descriptions, most participants felt that a head-mounted display would be unnecessary as audio would be the key access feature for them.

The issue of integrating immersive AD with assistive technology tools such as speech readers and magnification was briefly discussed. It should be noted here that all participants were regular users of video on-demand services such as Netflix, Amazon Videos and BBC iPlayer, which are configured to work with assistive technology. Individuals in our study reiterated that this was essential in order to enable independent access. However, since the characteristics of immersive environments are such that they may not allow control via traditional tools such as keyboard and mouse, voice control was regarded as most appropriate. Once again, members of the focus group had previously used voice control on smartphones, i.e., VoiceOver on IOS, and Talkback on Android devices.

6.6 Discussion

Overall, the focus group agreed that an audio-led immersive environment was easier to comprehend than an environment described only using AD. It was felt that the visual display of a 360-degree environment was somewhat irrelevant without elements of it figuring in the AD track. Head-mounted displays were regarded as unnecessary by the group with the exception of one participant, who is not a regular AD user.

The participants in our study strongly felt that the description track needed to complement the main narrative, and that any deviation from the primary storyline would lead to unnecessary disorientation. The five key elements of AD – *who*, *what*, *why*, *where* and *when* – were prioritised over the description of 360-degree elements by those in our pilot study. These were considered more acceptable when offered in the immersive audio environment of the *Virtual Barber Shop*, where sound alone was used to create a fully immersive environment but the script followed a single narrative with integral clues to the setting.

Our sight-impaired audience suggested that since 360-degree content was designed to be consumed independently, the AD script could be written in the second person, for example 'you're only a few paces away from the stage where the band is playing', etc. The professional describer added that this may help pull listeners into the scene and enhance the immersive experience in general.

The group discussed various factors that could contribute positively to the immersive experience, including the number of voices that would be considered appropriate in an immersive environment, and whether directionality and placement of the audio description would impact the viewer experience for people with sight loss. For example one of the questions that was debated was whether the audio description should be the voice in the viewer's ear, or alternatively, come from somewhere behind the audience. However, no definite conclusions could be reached in the absence of further examples.

On the subject of accessing content, there was a consensus on using a combination of voice control and integration with assistive technology tools – in particular, magnification and speech readers. It was discussed that traditional interfaces such as a keyboard and mouse might not be appropriate for accessible immersive environments, and that voice control may therefore be the only way to access content. It is important to state here that some participants who had previous experience of using voice-controlled environments such as Amazon Fire TV, Amazon Alexa and Google Home indicated that they would be happy to use voice as a control interface.

Further research, including specially produced content and a wider focus group comprising people with different sight levels, could clarify the importance of sound in an immersive environment.

6.7 Conclusions

It was clear from this study that participants, regardless of their sight level, wanted the opportunity to have the same experiences as consumers with normal sight. However, they were pragmatic in understanding that although they wanted to consume the same immersive media, if they had no vision it was pointless to have a full head-mounted display when only a tracked set of headphones would be needed. However, those with partial sight would still benefit from a head-mounted display as it would provide the same cues they have in the real world.

The most successful experiences were encountered during the first task where participants were taken on a directed path through the immersive content, and during the third task, i.e. the audio-only binaural demonstration. In the first instance, this is likely to be because this is how they are familiar with content, and without being able to fully see and experience the environment, it was very hard to choose to follow the action. Secondly, the *Virtual Barbershop* experience provided the most excitement and immersion as it was developed and designed to be a very specific audio-only demo. Nevertheless, it would be technically difficult to adapt this technique to work with general video content.

We therefore conclude that it is important to make immersive content fully accessible and available to all audiences regardless of vision level; however, it is now clear that providing AD for immersive content is not going to be straightforward. It may be necessary to provide a directed mode, where viewers are taken on a designated route through the visual material and guided through the key action in the media. If it is not possible for sight-impaired audiences to navigate 360-degree content independently, perhaps this raises the question as to their ability to contribute to a user-centred design approach. However, sight loss and visual acuity are clearly defined on a scale with a tipping point prior to which 360-degree content can still be enjoyed. As always, audio description must fit around the existing dialogue, never running simultaneously with the characters speaking. Binaural sound is also key to immersion, and further studies must be conducted to identify whether the virtual location of the audio describer voice is important.

Funding statement or declaration of conflicting interests (mandatory)

This publication is part of the EU-funded project ImAc, grant agreement no. 761974.

References

BIB-001 Afonso, A., Blum, A., Katz, B. F., Tarroux, P., Borst, G., and Denis, M. (2010) 'Structural properties of spatial representations in blind people: scanning images constructed from haptic exploration or from locomotion in a 3-D audio virtual environment', *Memory & Cognition*, 38(5), pp. 591–604.

BIB-002 BBC Earth (2016) *360° Attenborough and the Giant Dinosaur*, Available at: https://www.youtube.com/watch?v=IDkbkvOFxng (accessed 3 September 2019).

BIB-003 Cabeza-Cáceres, C. (2013) 'Audiodescripció i recepció. efecte de la velocitat de narració, l'entonació i l'explicitació en la comprensió fílmica' [Audio description and reception. The effect of speed of narration, intonation and explicitation in film comprehension]. PhD thesis, Universitat Autònoma de Barcelona.

BIB-004 Chmiel, A., and Mazur, I. (2012) 'AD reception research: some methodological considerations', in E. Perego (ed.), *Emerging Topics in Translation: Audio Description*, Trieste: Edizioni Università di Trieste, pp. 57–80.

BIB-005 Chmiel, A., and Mazur, I. (2016) 'Researching preferences of audio description users – limitations and solutions', *Across Languages and Cultures*, 17(2), pp. 271–288.

BIB-006 Ellis, G. (2016) 'Impairment and disability: challenging Conceptscof "normality"', in A. Matamala and P. Orero (eds.), *Researching Audio Description: New Approaches*, London: Palgrave Macmillan, pp. 35–45.

BIB-007 Fels, D., Udo, J. P., Diamond, J. E., and Diamond, J. I. (2006) ‘A comparison of alternative narrative approaches to video description for animated comedy’, *Journal of Visual Impairment & Blindness*, 100(5), pp. 295–305.

BIB-008 Fryer, L., and Freeman, J. (2012a) ‘Cinematic language and the description of film: keeping AD users in the frame’, *Perspectives: Studies in Translatology*, 21(3), pp. 412–426.

BIB-009 Fryer, L., and Freeman, J. (2012b) ‘Presence in those with and without sight: implications for virtual reality and audio description’, *Journal of Cybertherapy and Rehabilitation*, 5(1), pp. 15–23.

BIB-010 Fryer, L., and Freeman, J. (2013) ‘Visual impairment and presence: measuring the effect of audio description’, *Proceedings of the 2013 Inputs-Outputs Conference: An Interdisciplinary Conference on Engagement in HCI and Performance*, Brighton, UK, 26 June.

BIB-011 Fryer, L., and Freeman, J. (2014) ‘Can you feel what I’m saying? The impact of verbal information on emotion elicitation and presence in people with a visual impairment’, in A. Felnhofer and O. D. Kothgassner (eds.), *Challenging Presence: Proceedings of the 15th International Conference on Presence*, Vienna, 17–19 March. pp. 99–107.

BIB-012 Fryer, L., Pring, L., and Freeman, J. (2013) ‘Audio drama and the imagination: the influence of sound effects on presence in people with and without sight’, *Journal of Media Psychology*, 25, pp. 65–71.

BIB-013 ISO (2010) ‘9241-210 Ergonomics of human-system interaction, Part 210: Human-centred design for interactive systems.’ Available at: www.iso.org/standard/52075.html (accessed 3 September 2019).

BIB-014 Knigge, M., and Erkau, J. (2014) ‘Cinema and theatre: accessible entertainment, with benefits for all’, in H. A. Caltenco, P.O. Hedvall, A. Larsson, K. Rassmus-Gröhn and B. Rydeman (eds.), *Proceedings of the International Conference on Universal Design, UD 2014*, Lund, Sweden, 16–18 June.

BIB-015 Krejtz, I., Szarkowska, A., Krejtz, K., Walczak, A., and Duchowski, A. (2012) ‘Audio description as an aural guide of children’s visual attention: evidence from an eye-tracking study’, *Proceedings of the ACM Symposium on Eye Tracking Research and Applications Conference,* New York, ACM, 28–30 March, pp. 99–106.

BIB-016 Maszerowska, A. (2012) ‘Casting the light on cinema – how luminance and contrast patterns create meaning’, in R. Agost, E. Di Giovanni and P. Orero (eds.), *Multidisciplinarity in Audiovisual Translation,* Special Issue of *MonTI*, 4, pp. 65–85.

BIB-017 Maszerowska, A., Matamala, A., and Orero, P. (eds.) (2014) *Audio Description: New Perspectives Illustrated*. Amsterdam: John Benjamins.

BIB-018 Matamala, A., and Orero, P. (2008) 'Opera translation', *The Translator*, 14 (2), pp. 429–453.

BIB-019 Matamala, A., and Orero, P. (2011) 'Opening credit sequences: audio describing films within films', *International Journal of Translation*, 23(2), pp. 35–58.

BIB-020 Norman, D. (2013) *The Design of Everyday Things*. New York: Basic Books.

BIB-021 Orero, P. (2012) 'Film reading for writing audio descriptions: a word is worth a thousand images?', in E. Perego (ed.), *Emerging Topics in Translation: Audio Description*, Trieste: Edizioni Università di Trieste, pp.13–28.

BIB-022 Orero, P., and Vilaró, A. (2012) 'Eye tracking analysis of minor details in films for audio description', *MonTI*, 4, pp. 295–319.

BIB-023 Orero, P., and Matamala, A. (2016) 'User-centric audio description: a topsy-turvy research approach', in A. Manco and A. Mancini (eds.), *Scrittura brevi: segni, testi e contesti [Short writing: signs, texts and contexts]*. Naples: Università degli Studi di Napoli, pp. 376–387.

BIB-024 Orero, P., Maszerowska, A., and Cassacuberta, D. (2016) 'Audio describing silence: lost for words', in A. Jankowska and A. Szarkowska (eds.), *New Points of View in Audiovisual Translation*, Oxford: Peter Lang, pp. 219–236.

BIB-025 Palomo López, A. (2008) 'Audio description as language development and language learning for blind and low vision children', in R. Hyde Parker and K. Guadarrama García (eds.), *Thinking Translation: Perspectives from Within and Without*, Boca Ratón, FL: Brown Walker Press, pp. 113–133.

BIB-026 Perego, E. (2014) 'Film language and tools', in A. Maszerowska, A. Matamala and P. Orero (eds.), *Audio Description: New Perspectives Illustrated*, Amsterdam: John Benjamins, pp. 81–102.

BIB-027 QSoundLabs (2007) *Virtual Barber Shop*. Available at: www.qsound.com/demos/binaural-audio.htm (accessed 3 September 2019).

BIB-028 Rai, S. (2009) 'Bollywood for all'. Available at: www.rnib.org.uk/sites/default/files/2009_09_Bollywood_AD_report.pdf (accessed 3 September 2019).

BIB-029 Szarkowska, A., and Orero, P. (2014) 'The Importance of Sound in Audio Description', in A. Maszerowska, A. Matamala and P. Orero (eds.), *Audio Description: New Perspectives Illustrated*, Amsterdam: John Benjamins, pp. 121–139.

BIB-030 Schmeidler, E., and Kirchner, C. (2001) 'Adding audio description: does it make a difference?', *Journal of Visual Impairment & Blindness*, 95(4), pp. 198–212.

BIB-031 Starr, K. (2018) 'Audio description and cognitive diversity: a bespoke approach to facilitating access to the emotional content in multimodal narrative texts for autistic audiences.' PhD thesis, University of Surrey. Available at: http://epubs.surrey.ac.uk/848660 (accessed 3 September 2019).

BIB-032 Story, M. F., Mueller, J. L., and Mace, R. L. (1998) *The Universal Design File: Designing for People of All Ages and Abilities*. Raleigh, NC: Centre for Universal Design. Available at: www.ncsu.edu/ncsu/design/cud/pubs_p/pudfiletoc.htm (accessed 13 December 2019).

BIB-033 Udo, J. P., and Fels, D. I. (2009) 'Suit the action to the word, the word to the action: an unconventional approach to sescribing Shakespeare's Hamlet', *Journal of Visual Impairment & Blindness*, 103(3), pp. 178–183.

BIB-034 Van der Heijden, M. (2007) 'Making film and television accessible to the blind and visually impaired.' MA thesis, Utrecht School of the Arts. Available at: https://pdfs.semanticscholar.org/6b7c/507262678e9d01083b8fe09897c2f3081ed7.pdf (accessed 3 September 2019).

BIB-035 Vilaró, A., and Orero, P. (2013) 'The audio description of leitmotifs', *International Journal of Humanities and Social Science*, 3(5), pp. 56–64.

BIB-036 Walczak, A. (2017a) 'Creative description: audio describing artistic films for individuals with visual impairments', *Journal of Visual Impairment & Blindness*, 111(4), pp. 387–391.

BIB-037 Walczak, A. (2018) 'Audio description on smartphones: making cinema accessible for visually impaired audiences', *Universal Access in the Information Society*, 17(4), pp. 833–840.

BIB-038 Walczak, A., and Fryer, L. (2017) 'Creative description: the impact of audio description style on presence in visually impaired audiences', *British Journal of Visual Impairment*, 35(1), pp. 6–17.

BIB-039 Wilken, N., and Kruger, J. L. (2016) 'Putting the audience in the picture: mise-en-shot and psychological immersion in audio described film', *Across Languages and Cultures*, 17(2), pp. 251–270.

BIB-040 Wissmath, B., and Weibel, D. (2012) 'Translating movies and the sensation of being there', in E. Perego (ed.), *Eye Tracking in Audiovisual Translation*, Rome: Aracne, pp. 281–298.

BIB-041 WHO (World Health Organisation) (2001) *World Health Organisation Disability Assessment Schedule 2.0*. Available at: www.who.int/classifications/icf/whodasii/en (accessed 3 September 2019).[2]

[1] www.ohchr.org/EN/HRBodies/CRPD/Pages/CRPDIndex.aspx.