ENHANCING AUDIO DESCRIPTION: SOUND DESIGN, SPATIALISATION AND ACCESSIBILITY IN FILM AND TELEVISION

M Lopez Department of Theatre, Film and Television, University of York, York, UK
G Kearney Department of Electronics, University of York, York, UK
K Hofstädter CoDE Research Institute, Cambridge, UK

1 INTRODUCTION

Questions on accessibility are crucial as sight loss affects approximately 2 million people in the UK and estimations indicate that this number will increase to 2,250,000 by 2020 and 4 million by 2050.\(^1\) Significantly, 87% of visually impaired people access audio-visual entertainment such as film and television on a regular basis.\(^2\) Currently, the only existing accessibility system for visually impaired people wanting to watch film and television is Audio Description (AD), which is a pre-recorded audio commentary that provides information that clarifies the narrative, such as descriptions of actions, gestures and places.

This paper discusses the potential of using digital audio technologies and sound design techniques to provide a wider range of accessibility options that cater for different needs and preferences, while also creating an inclusive version of the soundtrack that is incorporated to the production and postproduction workflows and, as a result, represents more accurately the filmmakers’ vision.

2 AUDIO DESCRIPTION FOR VISUALLY IMPAIRED FILM AND TV AUDIENCES

AD was first proposed as an accessibility measure in the 1960s by Chet Avery who was inspired by the existence of closed captioning for people with hearing impairments, but it was not until the 1970s that Gregory Frazier developed the concepts behind AD.\(^3\) The UK has been at the forefront of developments in AD and
it is one of the countries in Europe in which it is more widely available. In 1992 ITC (Independent Television Commission, now Ofcom) led the European consortium Audetel (Audio Described Television), whose aim was to study the needs of visually impaired users in relation to television as well as explore the technical requirements for broadcasting. In 2003 the Communications Act dictated that broadcasters had to ensure 10% of their programming had AD, with the BBC, Channel 4, ITV and Sky committed to including AD in at least 20% of their content. In UK cinemas, a big breakthrough came in 2002 with the release of *Harry Potter and the Philosopher’s Stone*, the first film screened using a DTS system that ensured the synchronisation of the AD track to the film. In 2003 the provision of AD in the UK was increased with the funding provided by the UK Film Council to 78 cinemas in England, and nowadays 40% of the cinemas in the country include AD services. For home entertainment, around 500 DVD/Blu-ray titles include an AD track. However, recent years have seen the increase in popularity of online streaming services, in which accessibility is limited; although, BBC iPlayer and 4oD include an AD track on part of their programming and in April 2015 Netflix started adding AD to selected titles, AD provision still remains absent from other streaming services such as Amazon Instant and NowTV. The Royal National Institute of Blind People (RNIB) in partnership with MovieReading have developed and trialled a mobile application designed to overcome these limitations. Through this app users can download their required AD track and play it while watching the media content in question. The app then synchronises the AD track to the original soundtrack, allowing users to access the audio described version of the film.

The main strength of AD is its potential to allow visually impaired audiences to construct a story that is alike the one experienced by sighted people. AD users find TV programmes more interesting, informative and enjoyable, while also experiencing an increase in confidence and self-esteem as they can discuss TV programmes without fear of having misinterpreted the narrative or without the need of a sighted friend to describe the content.
Despite the advantages of AD, the system does have some inherent problems. The main shortcoming is the fact that it is an accessibility measure outside the creative process involved in a film or TV production. As a result, although the describer is meant to provide information in an objective manner, what s/he provides is her/his own interpretation of a particular piece. AD is not overseen by the director of the production, and as a result does not necessarily express the artistic vision of the piece. A further disadvantage of traditional AD practices is the focus on providing an experience that is as equally informative as the one offered to sighted audiences, but not necessarily equally entertaining. Furthermore, users of AD often complain about lack of intelligibility related to loudness, the complexity of the soundtrack and the audio mix. Last but not least, AD follows a ‘one fits all’ model that disregards current research on its success being affected by three main factors: expectations, needs and experience. Expectations can be related to the age group, with young users aged 15-20 wanting a service that ensures independence but does not make them feel different from their sighted peers. Different needs can be linked to different types of visual impairment and whether users have any residual vision. Experience can be associated to whether users have watched films or TV shows in the past before their sight loss and how much they know about the media.

3 SOUND DESIGN AND ACCESSIBILITY

AD guidelines have failed to acknowledge how advancements in digital audio could be game-changers in the process of conveying information as well as providing an entertaining experience to visually impaired audiences. The film and TV industries have also been slow in embracing the potential of sound design to foster inclusivity as well as provide new creative challenges. This section discusses creative practice in the fields of audio films, audio games, immersive theatre and documentary filmmaking that pushes the boundaries of their respective fields to embrace inclusivity.
3.1 Audio Films

Previous research explored the design of an alternative to AD, called Audio Film, which is a format of sonic art that eliminates the need for visual elements and a describer, by providing information through sound, sound processing and spatialisation.\textsuperscript{22} Sound effects are used both to represent actions and as soundmarks to help the listeners identify the different spaces in the narrative.\textsuperscript{23} Artificial reverberation is employed to provide each space with a characteristic sound while spatial audio (6.1 surround sound) is used to suggest the layout of the spaces as well as indicate the movement of the characters. A pilot study with visually impaired volunteers has demonstrated the viability of this format as well as the need for further research.\textsuperscript{24}

3.2 Audio Games

The field of electronic Audio Games, that is, games in which audio is the main way of communication and entertainment, is at the forefront of developments in the use of sound design for accessibility, while also incorporating notions of inclusivity from the start of the design process. \textit{Papa Sangre} (Somethin’ Else, 2010),\textsuperscript{25} \textit{The Nightjar} (Somethin’ Else, 2011)\textsuperscript{26} and \textit{A Blind Legend} (DOWNiNO, 2015)\textsuperscript{27} represent three of the best regarded Audio Games released in the last few years. \textit{Papa Sangre}, whose design involved RNIB and visually impaired volunteers, won the Most Innovative Game Mobile Gaming Awards in 2011, whereas \textit{The Nightjar}, was nominated for two BAFTA Game Awards in 2012.\textsuperscript{28} These audio games fuse together narration (either through voice over, dialogues, or a combination of both), real and abstract-sound effects as well as diegetic and non-diegetic sounds. Atmos tracks are used to indicate geographical locations and music is used to set the tone of each game. Furthermore, binaural audio, which is used extensively in Audio Games,\textsuperscript{29} helps the player locate objects relying on her/his hearing and, accompanied by volume automation, gives a sense of the distance of the playable character to an object or person. The Audio Games in question also use real time
reverb to indicate a change in the environment as well as auditory/sonic displays, which are recognisable sounds that give information to the user.\textsuperscript{30} For example, in \textit{A Blind Legend} a non-diegetic sound of a bell is used to indicate the importance of an upcoming event.

\subsection*{3.3 Drama Productions}

In the field of drama, professional companies have started incorporating ideas on the \textquote{aesthetics of access}, that is, how accessibility can provide creative challenges and possibilities.\textsuperscript{31} In this field, Extant (a professional company of visually impaired artists) has been investigating ways in which multi-sensory stimuli could be used to provide a similar theatrical experience to both sighted and visually impaired participants.\textsuperscript{32} \textit{Flatland} (Extant, 2015) is a participatory theatre piece that can be experienced regardless of one\textquotesingle s sight condition.\textsuperscript{33} Sighted people are blindfolded, and all participants are given a haptic device (called \textit{animotus}) to navigate the space and bone conducting headphones through which they can hear the pre-recorded story of \textit{Flatland} in the voice of the actor/guide. Moreover, the theatrical space is populated with electronic textiles that trigger sound effects linked to the story, and a combination of live action and pre-recorded material reproduced through loudspeakers, provide further immersion as well as channels for communication.

\subsection*{3.4 Documentary Filmmaking}

\textit{Notes on Blindness} (Middleton and Spinney, 2016) is a documentary on the academic John Hull, who after losing his sight started keeping an audio diary to record his experiences. All dialogue in the film is either from those audio diaries or from interviews conducted with John and Marilyn (his wife) for the film.\textsuperscript{34} \textit{Notes on Blindness} has a unique approach to accessibility as it provides several alternatives for visually impaired people, including the original film soundtrack, two different audio described versions and an enhanced option that uses additional interviews
as well as supplementary sound design and music. This film, as a result, acknowledges the need for alternatives in the field of accessibility, which is a main concern in the Enhancing Audio Description project.

4 ENHANCING AUDIO DESCRIPTION

Enhancing Audio Description (EAD) is a research project funded by the Arts and Humanities Research Council (AHRC) in the UK, which explores the use of digital technologies to transform the design and implementation of AD for film and television and as a result, change the ways in which visually impaired audiences experience audio-visual presentations. The project investigates ways in which AD can be updated through digital technologies to provide both an informative and entertaining experience. Moreover, by incorporating AD into film and television workflows we aim to provide an audio track that is closer to the artist’s vision.

At the centre of EAD is the belief that disabilities should not limit the options on how to experience audio-visual media and that the diversity of preferences by visually impaired people cannot be reduced to one accessibility method, but on the contrary requires a user-centred personalised method that allows audiences to make choices on access strategies.

Two key methods are explored as part of the project. The first one studies the potential of spatial audio to convey information in the visual channel, and the second is the use of first person audio description. Both methods require the incorporation of AD to the creative process, transforming it from just an accessibility measure into part of the creative workflows. To boost the chance for the incorporation of findings to the film and TV industries the research team is engaging with key stakeholders from the film, television, interactive media industries and accessibility services to consider the practicalities of such implementations.
4.1  Spatial Audio

The Audio Film format demonstrated that spatial audio is key to conveying clear narratives.\textsuperscript{35} However, it is also necessary to explore how it can be used as a tool to communicate cinematic visual elements such as the positions of objects and characters in space, types of shot, camera movements, camera angles and depth of frame.

Although the description of strictly cinematic elements, such as camera angles, is discouraged in the Ofcom guidelines on AD,\textsuperscript{36} previous research by Fryer and Freeman on Cinematic AD has shown that users favoured the inclusion of filmic terms to the descriptions.\textsuperscript{37} Such preference is probably due to the fact that people with acquired blindness have access to visual memories and those with congenital blindness can find spatial and auditory equivalents that make such descriptions meaningful. Moreover, Cinematic AD allows for a greater level of independence than traditional AD as it allows users to deduce the effect of the images presented to them, the same way than a sighted person would.\textsuperscript{39}

EAD takes the notion of Cinematic AD a step further by encouraging the communication of such information through audio spatialisation, instead of verbal descriptions. Success in this field is dictated by the accuracy of the sound spatialisation. Prior work on Audio Films has focused on sweet-spot listening over a 6.1 loudspeaker array. However, recent trends in the cinematic industry have led to the development of surround systems which significantly increase the loudspeaker numbers beyond this in an effort to improve sound localisation accuracy for distributed (non sweet-spot) listeners and heighten the sense of immersion. Examples include Auro 3-D (10.2 surround) and Dolby Atmos (64 channels).\textsuperscript{40} Such increased localisation accuracy could significantly benefit AD for visually impaired people.
Outside of the cinema, sound immersion can be difficult to achieve for audio-visual presentations. For example, in the home, as it is impractical to have dozens of loudspeakers placed around the living room.\textsuperscript{41} Methods for soundtrack processing over loudspeaker arrays located at the visual display are being developed in an attempt to deliver fully immersive surround sound in the living room. Similarly, soundtracks presented over headphones are heard ‘in-head’ as opposed to being externalised at accurate positions in 3-D space. Novel signal processing algorithms have been developed to deliver the correct auditory cues to the ears via headphones such that the cognitive processes for sound localisation are tricked into thinking that the sound is localised outside of the head and at an accurate position in 3-D space, known as binaural rendering.\textsuperscript{42}

4.2 The I-voice

Despite the huge potential of spatial audio to convey information to visually impaired audiences, verbal descriptions might still be required for colours, facial expressions, gestures and physical appearance. In these cases we are exploring how third person description could be replaced by the use of the \textit{I-voice}.\textsuperscript{43} The term \textit{I-voice} refers to the use of voice over by one of the characters to express her/his inner thoughts and help the identification of audiences with the character in question; in the context of AD, this would be the equivalent of a first person description. Furthermore, such \textit{I-voice} would be scripted during the production process by the scriptwriter with the help of an audio describer. A similar notion was successfully trialed by Fels \textit{et al.},\textsuperscript{44} who found that the creative team embraced the opportunity of having control over the AD track, a process that was easily incorporated to the workflows. Visually impaired volunteers found the first person narrative enjoyable and believed they would be able to share the experience with sighted friends. \textsuperscript{45}
4.3 User-Centred Design

The EAD project incorporates the user to the centre of the design process, considering their needs and opinions. With this aim in mind, we conducted eight focus groups (four in Cambridge and four in York, to facilitate geographical spread) with a total of 42 participants. Group size ranged from 2 to 10 participants and meetings lasted two hours each. The age of participants varied from 21 to 93 years old, with 60% female volunteers and 40% male.

Volunteers were invited to express their opinions on a specific short film chosen as a case study. The 20-minute short film, entitled *Pearl* was created by final year students of the BSc in Film and Television Production at the University of York. The film was shot in low-lighting conditions, making it an ideal example of a sound-driven film. Moreover, its unconventional storyline (the film tells the story of a girl that coughs up pearls), presents a challenge for visually impaired audiences.

Participants were shown the film without any accessibility features, and the focus group leader encouraged them to discuss its content. After this initial discussion, the participants were provided with information on the plot and invited to watch isolated extracts from the film and discuss how these could be made accessible. Furthermore, the focus group leader introduced notions to be investigated during the project and sought opinions on their potential interest. It is the findings of these focus groups that will be used to initiate the creative process of accessibility applied to *Pearl*.

4.3.1 Initial Findings

*Spatial Audio*

The majority of the volunteers commented that the use of spatial audio to replace or complement the verbal information given in AD would be advisable as it would add to the atmosphere of the film and as a result enhance the cinematic experience. Furthermore, one of the volunteers mentioned that the convention that
determines that voices are panned to the centre channel is unsuitable for accessibility purposes, as it does not allow visually impaired audiences to determine the movement of the characters. However, one participant mentioned that too much spatial information could become irritating and distract audiences from the film. Those who gave a more neutral feedback felt that sounds should only be spatialised when relevant to the storyline.

EAD seeks to explore spatialisation both through the use of loudspeaker arrays and binaural audio through headphones. However, feedback on the use of headphones was mixed. Some volunteers commented on their preference of using headphones as this was a less annoying solution for the rest of the family and could help visually impaired people who also have a hearing impairment. On the other hand, other participants mentioned that their families were happy to listen to the AD and they would prefer not to be cut off by having to wear headphones.

The I-Voice

Most of our participants were in favour of experimenting with first person narration and felt it could create an accessible version of films that felt more organic: "...it would be like a poetic device...Like a radio play, inner monologues, I think that would be interesting...it might be more organic...a part of it, rather than a stranger's voice cutting in." However, other participants thought AD was best done in the conventional way, with an objective third person narration and were concerned that first person narration would add an extra level of complexity and perhaps even detract from the enjoyment of the film.

Personalisation of Accessibility

Focus group participants indicated the need to personalise AD and for the sector to acknowledge that the variety of personal preferences and sight conditions cannot necessarily be reduced to one single form of accessibility. Opinions ranged from wanting and needing minimal description to full detailed accounts of visual
elements on screen. As expressed by a volunteer "... maybe a smaller version to a more informative version, that might be something for people that are just at the beginning of their sight loss stages... and people that do have severe sight loss might want to benefit more from a more descriptive narration... And we all have different levels of sight loss... Our needs are different."

Personalisation was also indicated in relation to first person narration, for example, by indicating that visually impaired audiences might want to access the story from different characters' perspectives. The possibility of accessing enhanced and spatialised sound effects that convey parts of the story more clearly was also mentioned as something desirable.

Incorporating AD to the Creative Filmmaking Process

Different groups mentioned the importance of integrating AD to the creative workflows and expressed disappointment when they realised AD was outsourced instead of being part of the production of a film. One of the volunteers said: "The thing you have to try to get out of literature or out of a film is the creative novelty the really new thing which the author is trying to say and that mustn't be diluted by other people's opinion of what they think it is because other people often give you a very partial view. You must go to the complexity if you possibly can of what the author and the different people involved in the film have been trying to project and then you get a much richer experience." Another volunteer added "...you might as well just not even pay to go to the cinema and, you know, talk to a friend afterwards about what it was about...because you are not getting the real experience." On the other hand, other volunteers expressed their confidence in AD and their sense of getting an objective account of the film by an impartial describer.

5 CONCLUSIONS
The Enhancing Audio Description project explores how sound design techniques and spatial audio can be used to transform audio-visual experiences for visually impaired audiences, whilst also turning the AD soundtrack from a compliance exercise into an intrinsic part of the creative process. The first stage of this research placed the user at the centre by holding focus groups with visually impaired volunteers, before starting any creative work. The focus groups conducted so far have shown an interest in the use of spatial audio techniques as well as excitement regarding the potential of first person narration to create a more organic experience. Furthermore, volunteers emphasised the need for personalised accessibility systems that allow users to change the amount of description accessed, and whether they would prefer a more traditional or experimental form of AD. The next stage of this project includes the analysis of a survey on the present and future of AD completed by over 100 volunteers. After such analysis has been completed the research team will use the data gathered to start work on turning the short film *Pearl* into an accessible film.

## 6 ACKNOWLEDGEMENTS

This project is funded by the Arts and Humanities Research Council (AH/N003713/1) and is a collaboration between the University of York and Anglia Ruskin University. Special thanks to the York Blind and Partially Sighted Society, Cam Sight and RNIB for their help in disseminating the survey and organising and hosting the focus groups. EAD is establishing links with the television and film industries as well as the accessibility sectors thanks to an advisory panel including Howard Bargroff (Sonorous), Jerry Gilbert (Cam Sight), Lisa Holdsworth (Screenwriter), Elfed Howells (Dolby), Sonali Rai (RNIB), Jez Watts (Sensor Media), John Whiston (ITV) and Warren Wilson (Cam Sight). Last but not least, special thanks to all the volunteers that participated in the survey and the focus groups, this project would not be possible without their input.
7 REFERENCES


5. S. Rai, RNIB International AD Exchange Study: Observations from a focus group study. Media and Culture Department, Royal Institute of Blind People, (May 2011).


7. Idem


16. Idem


34. http://www.notesonblindness.co.uk/ [Last accessed on 8th April 2016].


38. Idem

39. Idem


45. Idem
