

An Audio Art Form: Accessing the Visual of TV with Audio Description

KERR CASTLE¹

¹ Theatre, Film and Television Studies, University of Glasgow, Glasgow, UK
e-mail: k.castle.1@research.gla.ac.uk

Abstract

Audio description (AD) emerges as an exemplary intervention which seeks to translate the visual of television into an easily digestible audio form for visually impaired audiences. Relaying critical visual information as spoken word – including text, body language, movements and expressions – alongside each programme’s existing soundscape, AD transforms both the visual and aural characteristics of TV. The resulting text intends to engage directly with the needs of visually impaired viewers, a more inclusive media form that enables those with sight loss to better participate in contemporary culture. This paper seeks to examine this process of ‘audiovisual translation’ (AVT) and consider how this “enabling service” (Fryer & Freeman, 2012) changes the very architecture of television. Working with a current AVT practitioner, it examines the complexity of the audio describer’s task and explores examples of audio described TV from a practitioner perspective. Critically engaging with debates around the describer’s role, what gets translated, and alternative description styles, this paper sheds light on the mechanics of AD and the sorts of decisions and practical issues describers face daily. Ultimately considering the “viewing” experience offered by audio described programmes, and how this new way of hearing, as opposed to seeing, strives to remain artistic and expressive, this paper examines what the audio described text means, both for AD users and for Television Studies academics, raising significant questions around what television is and what we might learn from its translation to audio. Finally, this paper highlights important issues concerning UK television’s current levels of accessibility and the service’s capacity to support, include and cater for those with sight loss. Championing audio description’s ability to reveal a world of sight to those without it, this paper considers the provision of TV for visually impaired viewers not only as a caring intervention, but as a legitimate art form.

1 Introduction

The Royal National Institute of Blind People (RNIB) simply defines audio description (AD) as “commentary that describes body language, expressions and movements, making [a film or television] programme clear through sound” (2016a). In her seminal text, *An Introduction to Audio Description: A Practical Guide*, AD expert and practitioner Louise Fryer understands audio description as a form of “audiovisual translation” (2016: 2-3); that is, the mediation of visual information as aural commentary for visually impaired viewers.¹ Perhaps most significantly, within the context of this conference, Elena Di Giovanni notes that this translation

“allows for the restoration of agency in the enjoyment of films and audiovisual texts by [...] blind and partially sighted [viewers]” (2014: 137); an inclusive means by which they can continue to follow, enjoy and participate in their favourite films and television shows. What audio description aspires to offer, then, is a *comparable experience to viewing*; an “enabling service” that helps visually impaired audiences consume audiovisual media in a different way (Fryer & Freeman, 2012: 422).²

Primarily concerned with how TV is translated for audio description, this paper looks at audio description from a practitioner’s perspective, exploring the various

¹ An experienced audio describer herself, Fryer’s 2016 publication is really the first text of its kind to address the practice in such detail, with the resulting text now serving as a manual of sorts for practicing audio describers, both old and new.

² Its intervention does not simply afford access to visual information contained within a text, but changes the very shape of the text itself, introducing a virtual companion of sorts, the audio describer, who assumes the role of tour guide; taking the AD user beyond the original soundscape, enabling them to uncover detail and interpret information that would not otherwise be available to them.

challenges of AD and how describers seek to produce emotive, immersive and entertaining audio described television texts. In this instance, the describer, Jonathan Penny, is an AD scripter for Ericsson Broadcast and Media Services, though the term ‘describer’ can equally apply to those who voice AD texts. The following paper is the result of two separate interviews with Penny in August 2016 concerning his work in audio description. Beginning with some of the basic principles of AD and how the service should operate, this paper will progress to examine some of its challenges, exploring debates around the describer’s role, what gets translated, and different styles of description. Ultimately considering the “viewing” experience offered by audio described TV programmes and how this new way of hearing, as opposed to seeing, strives to remain artistic and expressive, the paper will conclude with thoughts on the future of the service and its place within academia.

2 Describing Television

For those unfamiliar with audio description and what it sounds like, Channel 4’s audio described *We’re the Superhumans* advertisement for the Rio Paralympics in 2016 perfectly illustrates the importance of audio described television shows and the creative potential of AD. A lavish celebration of the “superhumans” and their diversity, the advert highlights the sort of sporting achievements one might expect to observe in the Games along with various examples of individuals overcoming the specific constraints of their disabilities in daily life:

Audio Described: We’re the Superhumans –
<https://www.youtube.com/watch?v=EKGOWavFvHc>

With Adam Hills from *The Last Leg* (Channel 4, 2012-) replacing the traditional describer voice, the advert crucially begins by inviting all viewers – not just those who are blind or partially sighted – to “strap in” and experience television as the visually impaired audience would. Enabling the sighted viewer to compare and contrast their way of seeing with the colourful, energetic and concise descriptions in the AD commentary, the advert simultaneously raises awareness that the description service is available and highlights how essential it is for existing users.³

³ Channel 4 launched an accessibility survey earlier in 2016, promoted with a TV ad combining all three access services (subtitling, audio description and signing), again highlighting the existence of these services and seeking viewer feedback to ascertain how they could be improved upon (RNIB, 2016b). The advert is available to view here: <https://youtu.be/4iUm64DCebs>

It is also worth noting that Ofcom legislation only requires that 10% of television content broadcast by BBC, ITV/STV/UTV or Channel 4 be audio described, compared to the far more substantial subtitling

Distinct from the other access services – subtitling and signing – audio description requires the describer to produce their AD scripts from scratch (Fryer, 2016: 3), reading the screen’s contents, then interpreting and prioritising them; typically including features such as location, characters present, current speaker, onscreen text or graphics, but also actor performance, facial cues, actions and reactions. What is perhaps most notable for the sighted audience in the Rio Paralympics example is the number of quick-fire descriptions Hills must race through. While his intentionally long list draws attention to the number and diversity of participants featured in a comedic way, it also usefully demonstrates the overall complexity of description; constantly battling against time and fully aware not everything can be described. Consequently, some descriptions are out of sync, or played in early to set up the following scene (e.g. the blind footballer shushing), or even overlooked entirely (as with the two fencers). Yet, as Fryer notes, AD is “written for a listener, not a reader” (2016: 2). While it might seem ideal to list every athlete and every action, it would also be extremely complicated for the listener to follow, overwhelming them with unnecessary detail and bringing unwanted attention to the description process itself. As Penny explains, AD is intended to be “a layer of the programme”:

“At the end of the day, we’re trying to use as few words as possible to convey elements of the plot that aren’t obvious from the dialogue. Rather than describing all of the visuals, [we are] trying to aid comprehension.” (2016a)

The describer is tasked, then, with immersing viewers in such a way that audio description becomes an accepted part of the programme, a valuable aid as opposed to a distracting intervention that further hinders the viewing experience (Sweeney, 2012). As Fryer and Freeman note, “AD is at its best when the user is unaware of it” (2013: 16). To achieve this effect, describers must rework the television text in subtle ways, respecting the original tone and intentions of programme creators, but appreciating that blind and partially sighted audiences process information differently and require an adaptation better suited to their needs. What they are presented with instead, then, is a *version* of the featured programme, as opposed to a definitive translation, with

requirement of 100% for BBC output and 90% for ITV/STV/UTV and Channel 4 (2016: 1.12). Originally influenced by the “paucity of affordable technology able to receive the service” in the 2003 Audetel Communications Act (Fryer, 2016: 19), the 10% quota remains unchanged despite the increasing affordability of AD-supported digital platforms. While a voluntary AD target of 20% is consistently surpassed by all three broadcasters (Ofcom, 2016: 1.15), blind and partially sighted audiences are still greatly underrepresented.

the addition of audio description ultimately changing the very architecture of the show.

3 The Challenges of AD

Penny cites that one of the mantras of audio description, applicable to any genre, is “What You Say Is What You See” (2016a; also see Fryer, 2016: 56). However, the describer’s task is far more complex and restrictive than this. First and foremost, AD is traditionally tasked with sitting in between any gaps in the dialogue, which immediately restricts your capacity to paint a complete picture for users. As Penny appropriately comments, “It’s like I’ve got five things that I can take with me and five that are going to have to stay; and usually it’s a lot more skewed in favour of the stuff [I’ve] got to leave out!” (2016b). This inevitably impacts not only what you are able to describe, but the way in which you can describe it. The significance of how you ‘Say What You See’ is also influenced by a multitude of other factors, including programme genre, language used onscreen, pace and tone of a scene, the presence of music, and the audience’s familiarity with characters, settings or programme conventions. For example, with a long-running soap like *EastEnders* (BBC, 1985-present), AD users familiar with the show would likely be able to distinguish between the sounds of people in the Queen Vic public house compared to the Bridge Street “caff” before they’ve even been told where a scene takes place (Penny, 2016a). All of this usefully highlights how the describer’s role goes beyond relaying visual information but also requires them to accommodate the existing soundscape, to simultaneously speak for audiences old and new about any given show, its characters or storylines, and, where appropriate, even reflect a show’s vernacular in their descriptions, staying true to the linguistic register of a programme.

Indeed, while objectivity is prioritised in the majority of description work, so as “to avoid any manipulation, spoon-feeding or a patronising attitude towards the target audience” (Mazur & Chmiel, 2012: 173), it would appear seemingly impossible for the describer not to flavour descriptions with their own style and approach to reading the screen. As Fryer confirms, “sight [itself] is inherently subjective,” and “in order to increase engagement and immersion in AV products, describers need to accept and move towards subjective description, authorised where possible by the work’s creator” (2016: 172). Yet, during interviews, Penny noted that practical AD work requires the describer position her/himself somewhere between these two viewpoints, stating that, “You’ve just got to put yourself in the mind-set of the people who made the programme and try and describe it in a way that reflects what you think their intention is”

(2016a). Consequently, describers mediate what programme makers have attempted to communicate, both visually and aurally, in the original television text, utilising supporting materials, like post-production scripts and previous episode synopses, to further aid comprehension and ensure greater accuracy (Penny, 2016a).⁴ The resulting programme is an amalgam of the fully realised TV product and the describer’s subsequent decoding and reinterpretation of it.

This positions the describer as the ultimate decision maker for AD audiences, which in some instances means sacrificing content for clarity, such as omitting a character’s shocked reaction at the end of one scene in a soap to set up the next, or focusing on “*affordances*, i.e. what an object allows [the character] to do with it,” as opposed to the object’s “*qualities*, such as colour, texture...” (Gibson cited in Fryer, 2016: 56). Yet, as Penny frequently emphasised during interviews, while the process can be restrictive, the describer aims as much as possible to produce engaging, creative, well-timed descriptions that promote interaction with, and immersion in, the television text. One such example of this is prematurely relaying details of a physical comedy performance so that the end of the description coincides with the laughter track (Penny, 2016b). This keeps the AD audience in the programme and ensures they get the same opportunity to get the joke. Similarly, describing the performances of actors in terms of their “symptoms” (Penny, 2016b) or physical cues – compared to bluntly stating that someone is happy, sad or angry, for instance – allows the AD user to read and interpret the bodily signals given off by the individual, just as the sighted audience would, and formulate their own assessment of what those signals might mean.

4 The Language of Description

One example of this can be found in a short sequence from *EastEnders* (ep.5330), as a very troubled Grant Mitchell (Ross Kemp) struggles to come to terms with the events that brought him back to the Square. The sequence contains a range of symptoms, including Ian’s (Adam Woodyatt) immediate regret after his unexpected outburst to Grant, Grant’s sense of anger and frustration at his situation, Belinda’s (Carli Norris) dented pride as Grant rejects her, and the continued unravelling of Phil [Steve McFadden] and Grant’s relationship:

EastEnders [24:40-26:15] –
<http://www.bbc.co.uk/iplayer/episode/b07nwwny/ad/eas-tenders-09082016>

⁴ “For almost all of the big dramas that we do, we would have the post-production script, which is really helpful because it’s not always obvious what the intention of a scene is” (Penny, 2016a).

What makes this sequence interesting is the range of emotions on display, each communicated subtly in the AD commentary. Each description works to compliment the main dialogue and, where possible, allow space for silence; particularly between the feuding brothers (at this point in the British soap, Grant blames Phil for their mother's death). For example, towards the end of the sequence, the describer voice relays, "Grant stops and the brothers' eyes meet for a moment across the Square. [Pause] But Grant's jaw is set, and he turns away and continues onto Bridge Street". While this kind of audio described sequence might seem flat or functional to the sighted audience, I would argue such descriptions are in fact highly expressive in their own way. Although they articulate in relatively simple terms what is taking place onscreen, the language used in and of itself pertains to a very different way of seeing, in the same way that a screenplay is charged with activating the different visual elements of a story for its reader. A far more conscious cognitive process, as opposed to that of the sighted viewer, who likely reads the screen's ever-changing elements at a more unconscious level – description must leave space for interpretation and elaboration by the AD user; a sort of jumping off point, as it were.

Accordingly, language must be used in a creative way, summarising the constituent elements of a scene, but not limiting the AD user to one closed understanding of it. The way in which the silent exchange between Grant and Phil is described, for example, including the slightly graver tone of the voice as Grant turns away from his brother, communicates the significance of the scene in such a way as to activate the image for the AD user. Indeed, it is as if the exchange occurs in an imaginary space, one that the AD user is now able to access and observe, and is afforded the opportunity to make their own assessment of what this standoff means for both parties. Just as importantly, description here enables the AD user to better engage with further emotional aspects of the text; to empathise with Grant's sense of conflict or Phil's guilt and regret. As Penny notes, "You've got to use language in a creative way to be able to do that" (2016b), producing engaging descriptions that fundamentally transform the entire televisual experience for AD users.

5 Style

Perhaps the most significant ongoing debate in audio description is that of style; namely, how should describers speak to viewers and relay information on their behalf? Audio description's relatively formal mode of address has been a mainstay of the service since its inception, but as television has evolved, along with the

literacy of TV audiences, AD's style of delivery has been drawn into question. The primary contender to "traditional" AD is that of "cinematic" description; not only describing what is taking place, where and with whom, but how a scene has been realised technically, including details like camera movements, framing and shot transitions. Jonathan Freeman and Louise Fryer's investigation into the use of cinematic language in the description of film suggests "cinematic AD makes the audience viewpoint clear[er]" (2012: 416), enabling audiences to better understand their physical relation to the moving image and the added value of a film's visuals. However, it would seem that the practitioner's experience is somewhat different when it comes to the use of cinematic description for TV, with Penny arguing that describing a close-up shot can be "alienating for someone who might not know what that is" (2016a).⁵ He continues:

"If it was in a drama and you suddenly start talking about a 'tracking shot', it's like it's too [...] analytical, and for someone who was maybe just trying to be in the story, it would be jarring." (2016a)

Furthermore, while a film is more likely to allow space to linger on shots and let the visuals breathe, the same is not always true for television, which typically requires economical and concise description, quickly shifting from scene to scene.

While cinematic AD appears less at home on television screens, a further alternative highlighted by Penny is a more conversational approach; "basically imagining the audio describer is not bodiless, but is actually someone sitting next to you" (2016a), a sort of *companion*, watching with you and relaying information as a friend or family member might. One example of this conversational style in action is with a show like *Deal or No Deal?* (Channel 4, 2005-2016). In this game of chance, contestants would frequently state things like "I hope it's a blue" (the lower value boxes) to the player of the game before opening their box to reveal its true monetary value. Once the box had been opened, the

⁵ The "language" of description was determined during the original development of the Audetel service: "Between July and November 1994 an ambitious trial of a full Audetel service operated on peak-time ITV and BBC television, delivering over 6 hours of described programming a week to 140 receivers throughout the UK, mostly situated in the homes of visually impaired people. The service was carefully monitored to record practical engineering, logistical and editorial experiences as well as to generate a wealth of feedback from the users on receiver ergonomics and on the quality of the descriptions themselves. The trial was an overwhelming success, demonstrating not only the practicability of regular broadcasts, but also enormously increased comprehension and enjoyment among blind and partially sighted viewers" (ITC, 2000: 5).

describer could simply have stated the colour and box value, e.g. “Red, £20,” in a very formal, detached way. However, instead, the AD commentary almost became a part of the game, excitedly announcing things like, “But it’s a red!”, followed by the box’s monetary value. This performative aspect again reinforces the creative role of the describer (both scripters and voicers), embracing the light-hearted, informal tone of the show in the described commentary in order to keep the AD user in the show, a part of the moment onscreen.

6 Conclusion

As I hope to have evidenced throughout this short paper, describers consistently seek to activate TV texts for AD users in immersive, emotive, engaging and entertaining ways, not simply offering a stripped back, nuts and bolts version of TV, but instead an alternative realisation of it. As Giovanni explains, viewers are in fact observers; “they perceive [...] visual stimuli and cognitively elaborate [a text’s] reception” (2014: 137). In the same way sighted viewers will often unconsciously read body language, movements and expressions, AD users simply have a different way of seeing and require different “elements” to then elaborate their own meanings from a TV text. Audio description enables them to do this, and as such should be applauded and discussed more widely as a legitimate remediation of television that promotes greater audience equality.

Joel Snyder perfectly encapsulates the importance and richness of audio description, explaining:

“AD uses words that are succinct, vivid, and imaginative to convey the visual image that is not fully accessible to a segment of the population and not fully realised by the rest of us – sighted folks who see but who may not observe.” (2007: 100)

Comparing AD to “a kind of literary art form [...] a type of poetry” (2007: 100), Snyder highlights how the audio described text can in fact enable *all* audiences to see and understand more clearly. This notion of the poetic perfectly sums up the subtleties of word craft, pacing and inference required in audio description, producing rhythm, momentum and affording space to engage with a text more completely.

For the purposes of film and television academia, audio described television texts tell us a great deal about how we all instinctively read the television screen, how we make sense of what we see, and could prove particularly useful for undergraduate students as they learn more

about the interrelation of sound and image.⁶ Equally, Snyder’s comments reinforce that AD should not be overlooked as an art form in its own right, and is worthy of academic attention beyond its means of production.⁷ The audio described television text can be as artistically valid and worthy of analysis as its source material.

As for the service itself, increased awareness of its existence and increased levels of choice when it comes to styles of description are crucial moving forwards. AD needs to adapt to the needs and preferences of users, and move beyond the guidelines produced and feedback recorded during its inception. As Penny appropriately noted in our final interview, anecdotally, he hears a lot about people “who have turned [audio description] on by accident and don’t realise what it is,” leaving them bewildered and sometimes frustrated (2016b). If AD and, in fact, the whole of access services is about equality and improving usability, then public knowledge of the service clearly needs to change.

Bibliography

- Evans, Elizabeth Jane; Pearson, Roberta (2009) *Boxed Out: Visually Impaired Audiences, Audio Description and the Cultural Value of the Television Image*. *Participations*, 6:2.
- Fryer, Louise; Freeman, Jonathan (2012) Cinematic language and the description of film: keeping AD users in the frame. *Perspectives*, 21:3, pp.412-426.
- Fryer, Louise; Freeman, Jonathan (2013) Visual Impairment and Presence: Measuring the Effect of Audio Description. *Inputs-Outputs Conference 2013: An Interdisciplinary Conference on Engagement in HCI and Performance*. New York, USA: ACM, article no.4.
- Fryer, Louise (2016) *An Introduction to Audio Description: A Practical Guide*. London: Routledge.
- Garman, Judith (2011) Autistic Spectrum, Captions and Audio Description. *Mindful Research*. [Online] Available from:

⁶ As Elizabeth Jane Evans and Roberta Pearson note, “Analysing a piece of audio described television presents a particular difficulty for the sighted researcher. AD foregrounds information that a sighted viewer may ignore” (2009). Accordingly, learning to read and interpret the elements of television as a visually impaired viewer would with audio description could prove particularly useful and enlightening.

⁷ Furthermore, research by Judith Garman has identified the value of AD for individuals on the autistic spectrum, helping them to identify emotions and reinforcing information they might otherwise struggle with, such as identity (2011).

- <http://mindfulresearch.co.uk/2011/08/29/autistic-spectrum-captions-and-audio-description/> [Accessed 28 July 2016]
- Giovanni, Elena Di (2014) Visual and Narrative Priorities of the Blind and Non-blind: Eye Tracking and Audio Description. *Perspectives*, 22:1, pp.136-153.
 - ITC (2000) ITC Guidance on Standards for Audio Description. [Online] Available from: http://audiodescription.co.uk/uploads/general/itcguide_sds_audio_desc_word3.pdf [Accessed: 19 August 2016]
 - Mazur, Iwona; Chmiel, Agnieszka (2012) Audio Description Made to Measure: Reflections on Interpretation in AD based on the Pear Tree Project Data. In A. Remael, P. Orero, & M. Carroll (Eds.), *Audiovisual Translation and Media Accessibility at Crossroads. Media for all 3*. Amsterdam: Rodopi.
 - Ofcom (2016) Television Access Services Report 2015. [Online] Available from: <http://stakeholders.ofcom.org.uk/binaries/research/tv-research/access-service-reports/access-report-2015.pdf> [Accessed: 19 August 2016]
 - Penny, Jonathan (2016a) *AVT Practitioner Interview One*. Interviewed by Kerr Castle [in person]. Ericsson Broadcast and Media Services, 100 Brand Street, Glasgow, 8 August 2016.
 - Penny, Jonathan (2016b) *AVT Practitioner Interview Two*. Interviewed by Kerr Castle [in person]. Ericsson Broadcast and Media Services, 100 Brand Street, Glasgow, 12 August 2016.
 - RNIB (2016a) Audio Description (AD). [Online] Available from: <http://www.rnib.org.uk/information-everyday-living-home-and-leisure-television-radio-and-film/audio-description> [Accessed: 28 July 2016]
 - RNIB (2016b) Channel 4 launches an accessibility survey. [Online] Available from: <http://www.rnib.org.uk/channel-4-launches-accessibility-survey> [Accessed: 19 August 2016]
 - RNIB (2016c) How many people in the UK have sight loss? [Online] Available from: <https://help.rnib.org.uk/help/newly-diagnosed-registration/registering-sight-loss/statistics> [Accessed: 28 July 2016]
 - Snyder, Joel (2007) Audio Description: The Visual Made Verbal. *The International Journal of The Arts in Society*, 2:2, pp.99-104.
 - Sweeney, Mary (2012) 3 Challenges of Audio Description. *Ericsson*. [Online] Available from: <http://www.ericsson.com/broadbandmedia/what-we-think/blog/3-challenges-of-audio-description/> [Accessed: 28 July 2016]

Filmography

- *EastEnders*, 2016. Episode 5330. [Television Programme] BBC, BBC One, 9 August 2016 19.30.
- *We're the Superhumans* (Channel 4, 2016) [Audio Described Television Advertisement] Available from: <https://www.youtube.com/watch?v=EKGOWavFvHc> [Accessed: 8 September 2017]