

---

# An Exploratory Study of Best Practices for Audio Descriptions in Visual Media for Individuals with Visual Impairments

**Jonathan Anderson**

DePaul University  
243 S. Wabash  
Chicago, IL 60604  
jande162@mail.depaul.edu

**Chris Poynton**

DePaul University  
243 S. Wabash  
Chicago, IL 60604  
iamchrispoynton@gmail.com

**Diana Speicher**

DePaul University  
243 S. Wabash  
Chicago, IL 60604  
dspeicher27@gmail.com

**Aditya Yellamraju**

DePaul University  
243 S. Wabash  
Chicago, IL 60604  
adiyellamraju@gmail.com

**Abstract**

People who are blind do not experience movies or television in the same way that sighted people do. Instead, they often use Audio Description (AD); AD provides contextual details that are important to understanding the scenes that comprise a story. However, the quality of AD varies, in part due to the different means of creation. We aimed to understand user preferences when it comes to AD to derive a set of guidelines that can be used to create better and more consistent AD. Through four one-on-one interviews with four users of AD who were blind, we generated four categories of guidelines; (a) level of detail, (b) delivery style, (c) sound and audio, and (d) platform features.

**Author Keywords**

Accessibility Guidelines; Human-Computer Interaction; Visual Impairments; Audio Descriptions; VOD;

**CSS Concepts**

• **Human-centered computing** ~ **Empirical studies in accessibility**

**Introduction**

People who are blind cannot access visual-based content in videos, movies, and television shows in the same way sighted people can; this potentially excludes them from enjoying visual-based media. Audio descriptions (AD) that provide contextual details can make visual-based content accessible. However, there

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*CHI 2020 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.*

© 2020 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-6819-3/20/04.

DOI: <https://doi.org/10.1145/3334480>.

is a lack of universal formatting standards for content listed on Video on Demand platforms (e.g. Netflix), and the availability AD is limited [4].

AD provides additional spoken content and information for people who are blind or low vision [10]. The additional contextual information helps viewers comprehend scenes which can evoke emotional details otherwise not realizable, creating a more robust experience for people with visual impairments [7]. According to Snyder [8], “[AD] provides a verbal version of the visual - the visual is made verbal.”

The production time required to create AD is lengthy and costly [1]. Due to the high cost and significant production times many videos, movies, and television shows lack AD [3]. Technology is emerging that reduces production time and cost by automating the creation of AD making more visual-based content accessible.

Branje and Fels [1] explored the use of a Graphical User Interface (GUI) that assists amateur describers as they record, insert, and edit their descriptions. They found their AD was suitable to viewers who were blind which could potentially eliminate the need for professional describers and reduce production costs [1]. Pantula and Kuppuamy [6] assessed a tool that automatically inserted human narrated AD into a scene prior to it starting. They determined that when AD is presented prior to the start of a scene users with visual impairments demonstrated greater comprehension and ability to follow the narrative. These tools blend human narration with automated technology to create AD. Other tools remove the human aspect altogether, further reducing production time and cost.

Campos et al. [2] explored software both created and inserted AD automatically. The software used algorithms that parsed the show’s script by searching

for key phrases and word frequency. The words and phrases were then used to create a summary of the scene which was inserted as an AD. They found that while the software was able to produce and insert the AD, human intervention was needed to ensure accuracy and quality of the descriptions [2].

Lopez et al. [5] reported that despite the variety of methods used to create AD, the quality varies tremendously. This in part is due to the lack of unified guidelines leading some describers to adopt more creative description practices while others maintain more rigid approaches [9].

In this project, we explored AD through interviews with four people who were blind; we aimed to develop AD guidelines. Our findings contribute to guidelines and platform best-practices.

## Methods

In the following section, we describe our participants, data collection methods, and data analysis methods.

### *Participants*

We recruited four participants through personal networks. All participants were at least 18 years of age and identified as being blind. Three participants lived in Chicago, Illinois, and one participant lived in Portland, Oregon. All four participants had used AD when consuming visual media (VOD, DVDs, Blu-Rays, etc.) at least three times in the last month; see Table 1 for participant details.

Name	Age	Gender	Video Usage	Yrs. Blind
James	37	M	Daily	37
Martin	30	M	Weekly	30
Richard	30	M	Daily	30
Julia	48	F	2-3 times/wk	48

Table 1: Participant Demographics

### *Data Collection*

We conducted four interviews in-person. James and Martin were interviewed in separate student rooms in the DePaul CDM building. Richard was interviewed in a conference room at Friedman Place, a housing community for the blind on the Northside of Chicago. Julia was interviewed at the Tigard Public Library. All interviews lasted 60-70 minutes.

After receiving verbal consent, we asked warm up questions focused on participant's media consumption, and device preferences. We then asked about participant's use of AD. This included questions about their preferences surrounding level of detail, length, and pacing. We also asked questions about their preferred platforms, streaming services and desired features. We ended with a set of demographic questions. Participants were given \$20 gift card for their participation. All interviews were audio recorded and later transcribed.

### *Data Analysis*

We used open inductive coding using Atlas.ti. to analyze the transcripts. We transferred our codes to an online affinity diagramming tool (Stormboard) to collaboratively organize and group our codes into common and salient themes.

## **Findings**

We categorized our findings into four themes: (a) level of detail, (b) delivery style, and (c) sound and audio, and (d) platform features.

### *Level of Detail*

We classified the level of detail as any aspect of the characters or scene that is described in an AD. We further classified this into two themes: (a) aesthetic, and (b) context.

### AESTHETIC

All four participants said that any level of description was better than no description, and that they generally avoid media that does not contain AD. All four participants told us that clothing descriptions, facial expressions, and scenery were very important to help gain a better understanding of the environment and setup of a scene; however, they did not feel that there was any specific priority. Julia said, "I don't think there's one specific detail that's most important." Similarly, James told us, "The most important ones, I think, are the ones that a sighted person is more likely to focus on at any given moment."

### CONTEXT

All four participants also mentioned that ADs should be robust and have enough visual cues to provide context to the scene. The details mentioned by each participant were scene-specific and played directly into both the interactions and subtleties of the plot or scene. James said, "As long as you can tell what's going on in key moments, then I think that's a good set of descriptions." Richard elaborated further on the importance of potentially sensitive or politically incorrect details, saying that some ADs may leave out race, even when they are important to understanding the conflict of a scene. s

### *Delivery Style*

We defined delivery style as any aspect of the AD that related to how the AD was written and performed. We further classified this into three themes: (a) script and pacing, (b) the role of the describer and their narration style, and (c) word choice and tone.

### SCRIPT AND PACING

James, Richard, and Julia all mentioned the importance of a professionally written script that fits within both the film script and the pacing of the film. Richard was the most outspoken on this, saying that the script was

“very much overlooked,” and that it required “more care and consideration.”

James, Richard, and Julia elaborated on the importance of pacing within the script, and how that can affect their viewing experience. Julia said, “I think it’s important for the descriptions to fit within the timeframe of the pause/silence,” and that longer pauses for descriptions would “throw off the pacing of the movie.” Participants also noted that this was often challenging, especially in scenes that involved a lot of characters and/or were fast paced. James discussed the importance of keeping the pace of the movie by returning to previous details later in the scene, an idea he referred to as “pretty ingenious.”

#### ROLE OF THE DESCRIBERS AND THEIR NARRATION STYLE

All four participants conveyed the importance of having a professional describer. Martin said that professional describers have “a special aptitude for it” due to their training and experience. Both Martin and Richard discussed this in contrast to having their friends describe the scene, with Richard saying that his friends gave “incredibly horrible descriptions” that did not provide any insight into what was going on.

All four participants also described the importance of delivery style, sharing a preference for the use of emotion over a flat delivery style. Julia said her experience with *LaLa Land* was good, largely due to “how the describer used emotion in his own voice.” Richard made similar comments, saying “just hearing how the describer can be so involved with the story and the characters and things, that was awesome,” when describing his experience watching xxx. James mentioned that this was even useful in horror films, discussing how the describer used a “whispery voice and then the sudden shock that scares everybody.” However, Richard felt that this may be a generational

preference, and felt that “visually impaired people who are senior citizens like the flat narration.”

#### WORD CHOICE AND TONE

Richard and James discussed the importance of word choice and selecting words that fit with the tone of the film. Richard cited an experience when watching *The 40 Year Old Virgin* where the AD used a less offensive term than what was found in the movie, saying “that was a little strange to clean up your script when the movie is clearly not kosher.” He also recounted the use of slang, saying “if a lot of characters are using slang for an item, I would like to have the describer use the same slang term.” James noted that he did not generally have a preference but said in reference to overly sexual descriptions that if “something sticks out and really takes my headspace to a different space where it shouldn’t be, then yeah that’s a problem.”

#### *Sound and Audio*

We defined sound and audio category as the media and AD’s audio recording quality and audio output options while watching media.

All four participants noted the importance of sound and audio when adding AD to media. James and Julia discussed the AD audio track interfering with the standard film audio track. Julia recalled being “frustrated when the audio soundtrack volume gets lowered so that the descriptions can be heard.” James was more concerned with overlap, saying that when AD is “on a separate mono channel...you’re going to lose all that nice information on the soundtrack’. Martin discussed the idea that AD can interfere with more than just dialogue, saying “other sounds that’s not actually dialogue that can be helpful.”

James and Robert also mentioned the importance of having the AD audio quality match that of the film, specifically when using Dolby Atmos Surround Sound.

Robert noted that this allows people to “use the system to the fullest potential.”

#### *Platform Features*

Platform features included aspects of the AD that related to platform functionality and accessibility. We further classified this into two themes: (a) distinction that AD is available and (b) selecting a title to watch.

#### DISTINCTION THAT AD IS AVAILABLE

Julia, Richard, and James wanted ADs to be easier to find, and provided ideas to make this task easier. James liked that Disney uses an audio cue, saying “the chime is there to alert me that there is an audio descriptive track.” Richard said that Netflix provided a category for content with AD, making it “really easy to look for content.” However, he also noted that even sighted people had a problem finding content with AD because there was no symbol anywhere to indicate the availability of ADs.

#### SELECTING A TITLE TO WATCH

James and Richard discussed how they selected a title to watch. James said that he previewed trailers and read reviews. However, Richard said that “trailers are pretty much useless...because they’re so visual to begin with,” and that he would prefer hearing a synopsis. However, to his knowledge there were no platforms that provided additional accessibility features to read plot synopses, if when provided.

### **Discussion**

Three of our categories of findings ((a) level of detail, (b) delivery style and (c) sound and audio) were aligned to those in the literature. However, platform features were not discussed in previous research; all four participants mentioned this as an important aspect to the experience of using AD.

Our four categories provided the framework for creating a set of AD guidelines.

#### *Guideline Recommendations*

These guidelines can be implemented through manual creation of AD (human narration) or through automated creation (synthetic narration).

#### LEVEL OF DETAIL

- Use robust descriptions that focus on details that sighted people would be drawn to.
- Provide details that give context to the scene (especially with gags or instances that rely on visual cues).
- Identify characters when necessary.
- Prioritize details that are important to the context of the scene first and add additional details later.

#### DELIVERY STYLE

- Use professionally written AD scripts that fit the tone and style of the actors’ script.
- Use words that match the feeling of the film or show (e.g. raunchy words, slang, etc.).
- Provide options for multiple narration styles (i.e. flat vs enhanced).
- Use a vocal tone that matches the feeling of the film or show.
- Record descriptions with professionally trained voice actors and describers.
- AD should not interfere with dialogue.

#### SOUND AND AUDIO

- Keep audio tracks consistent with film audio (e.g. using surround sound).
- Create a separate channel so users can have individual control of the AD track or be able to listen on third party devices (e.g. headphones).

#### PLATFORM GUIDELINES

- Use the standard AD symbol (with audio) to clearly indicate when AD is available.

- Create a category for browsing that includes all content with AD.
- Record a plot synopsis for people with visual impairments to listen to so they can more efficiently select a title to watch.

#### *Limitations of the Study*

Our study was limited by a small sample size which was not large enough to generalize a larger audience. We were also limited by homogenous demographics. A majority of our participants fell in the age range of 30-40 years old, which may not be reflective of the habits and preferences of younger and older audiences. All participants also used AD frequently. Findings may differ with participants who are not as experienced with AD.

All studies were conducted in different settings, some of which may have been foreign to the participant. This may have affected the participant's level of comfort, and the answers they provided.

#### *Future Work*

In future studies, we would like to increase our sample size and recruit from more diverse age groups. We would also like to use our initial findings to expand our interview questions to be more targeted at the specific aspects of AD we found in this study.

#### **References**

- [1] Carmen J. Branje and Deborah I. Fels. 2012. LiveDescribe: Can Amateur Describers Create High-Quality Audio Description? *Journal of Visual Impairment & Blindness* 106, 3: 154-165.
- [2] Virginia P. Campos, Tiago M.U. de Araujo, Guido L. de Souza Filho, and Luiz M.G. Goncalves. 2018. CineAD: A System of Automated Audio Description Script Generation for the Visually Impaired. *Universal Access in the Information Society* 17, 3: 510-514.
- [3] Katie Ellis. 2015. Netflix Closed Captions Offer an Accessible Model for the Streaming Video

Industry, but What about Audio Description? *Communication, Politics & Culture* 47, 3: 3-20.

- [4] Katie Ellis, Mike Kent, Kathryn Locke, and Melissa Merchant. 2016. Accessing Subscription Video on Demand: A Study of Disability and Streaming Television in Australia. Australian Consumer Action Network, Curtin University, Sydney.
- [5] Mariana Lopez, Gavin Kearney, and Krisztian Hofstadter. 2018. Audio Descriptions in the UK: What Works, What Doesn't and Understanding the Need for Personalising Access. *British Journal of Visual Impairment* 36, 3: 274-291.
- [6] Muralidhar Pantula and K.S. Kuppamy. 2019. AuDIVA: A Tool for Embedding Audio Descriptions to Enhance Video Accessibility for Persons with Visual Impairments. *Multimedia Tools and Applications* 78, 14: 20005-20018.
- [7] Marina Ramos. 2015. The Emotional Experience of Films: Does Audio Description Make a Difference? *The Translator* 21, 1: 68-94.
- [8] Joel Snyder. 2005. Audio Description: The Visual Made Verbal. *International Congress Series* 1282, 1: 935-939.
- [9] Agnieszka Walczak and Louise Fryer. 2017. Creative Description: The Impact of Audio Description Style on Presence in Visually Impaired Audiences. *British Journal of Visual Impairment* 35, 1: 6-17.
- [10] Agnieszka Walczak. 2017. Audio Description on Smartphones: Making Cinema Accessible for Visually Impaired Audiences. *Universal Access in the Information Society* 17, 4: 833-840.