

Fiction Podcasts Model Description by Design

GEORGINA KLEEGER

Let's say you want to produce a movie or video with access features such as captioning and audio description (AD). If this is the case, I applaud you. Clearly, you recognize that, as in architecture and product design, access always works better and is more in line with the original ideas of the producer when it is in the plan from the outset rather than added on later in postproduction. Also, there's some self-interest involved. You know that captioning has long been popular with people who are not deaf or hearing impaired. More recently, as more and more AD is available on streaming platforms such as Netflix and Amazon Prime, there is evidence that many people who are not blind or visually impaired also enjoy it for the way it allows them to consume visual media while doing something else with their eyes. AD can be helpful to people learning another language for the way it provides new vocabulary. And neurodiverse audiences appreciate how it can enhance comprehension by directing attention to specific aspects of a scene. So designing for access by taking charge of AD from the outset will mean that your project will reach a larger audience and allow consumers to choose how they will consume it.

Because I am blind and consume AD, and have written somewhat critically about current standards, you may come to me seeking my advice. I sense you there, with your fingers poised above your keyboard, waiting for a list of dos and don'ts, or perhaps a recommendation of a show or movie where the AD was especially good. I'm sorry to say that I will disappoint you. When I recommend the AD on something I've been watching as adequate or helpful, sighted friends often come back with criticism that the describer left out crucial visual details. Or else they may shrug and think, "I guess that's what the blind people want." As a blind person, I don't know what I'm missing and so tend to prefer AD that is not distracting. Other AD users have different preferences. Often I find AD simply perplexing, supplying odd bits and pieces of information I can't quite imagine needing to know. I am left to shrug and think, "So that's what sighted people find noteworthy." Keep in mind that I am of the generation that grew up listening to TV and movies without AD; I even studied film in college, and so I have a basic understanding of the conventions of visual storytelling. For the most part, without AD I can always follow the plot with only minimal additional information from sighted companions, whether in real time when screening at

home or after the fact when screening in the movie theater. Usually I can derive meaning the same way I do in real life, from characters' words, ambient sounds, and the context of the plot.

There are now many sources of guidelines for AD I could direct you to with the caveat that you should read them critically. The Audio Description Project of the American Council of the Blind is a comprehensive clearing house for information about blind access to all genres of visual media, films, television, museum collections, live theater, and sporting events. It also provides information for people seeking training as a professional describer, technical specifications for the different equipment used, and numerous examples and lists of guidelines and best practices. There are now also books, such as Louise Fryer's *An Introduction to Audio Description*, which essentially offers a self-guided training course with both theory and practice in the form of exercises would-be describers can try out on their own. These instructions tend to assume that the art or media in question is already complete, and that description will be added on in postproduction and without the input of the creator.

But you want to do something different. Since it's part of your plan to have AD, you can write your script to allow space for description. You can also attend to the nonverbal aspects of your soundtrack that deliver meaning. Producers like you who wish to design for accessibility from the outset would do well to remember that since the advent of the talkies, film became both an audio and visual medium. The work of the soundtrack supports the visual content but also sets the mood and makes meaning. Actors' vocal performance enhances the meaning of the words spoken, while sound effects and music convey context and anticipate action. Audio describers are already enjoined to refrain from describing something that can be derived from an audible cue. For instance, if the phone rings, and the character can be heard picking it up and saying, "Hello," the describer does not need to mention that the character is answering the phone. Dialogue can also deliver descriptive information that the describer does not need to repeat. For example, if a character says, "What are you smiling at?" there's no need to describe the other character's expression.

As inspiration, you can find examples of media where some form of description is already built in—that is to say, description that is not necessarily directed at an exclusively blind audience. Take, for instance, baseball on the radio. Of course, radio has always been accessible to blind people. While news broadcasts deliver information primarily through speech, producers have always included ambient sound when recording in the field, and reporters and commentators may include visual details to help orient the audience. Although it is a complex visual spectacle, baseball is particularly well suited to audio broadcast, because it is comparatively slower than the back-and-forth sports like basketball and soccer. Long intervals of relative inaction are interrupted by sudden, dramatic events. During those long intervals, announcers fill the air with historical and

biographical accounts of past games and players, recital of statistics, and bemused observations about the endearing antics of fans in the stands. And while the majority of play-by-play may simply be the counting of balls and strikes, the best broadcasters know how to paint a picture with both their words and voices. In fact, it is frequently the vocal performance that conveys the meaning. Imagine hearing this: “It’s a line drive deep to right field.” Even an unpracticed listener instantly recognizes from the way the announcer delivers those words whether the ball is likely to be caught or to clear the fence for a home run.

Radio broadcasts of baseball assume that the audience cannot see what the announcers do. There have been occasions when the announcers weren’t seeing the action on the field either. When Ronald Reagan was a radio sportscaster in the 1930s, he narrated games from accounts he received via telegraph, and then added the play-by-play description with the embellishments of his imagination. More recently, during the pandemic seasons for 2020–2021, most baseball announcers did not travel with the teams and so delivered their commentary from watching the game on TV. Many complained that the camera angles and framing did not allow them to see certain plays, and so they had to resort to some speculation in order to describe them.

Baseball on the radio provides a century-old model of AD built into the medium. A more contemporary example is the narrative podcast. Long-form investigative journalism, as well as fictional dramas, tell their stories with both words and sounds. For example, in the many popular true-crime podcasts, the text may quote from missing persons reports or eyewitness accounts of crime scenes. In fiction podcasts, description may be subtly included in characters’ speeches.

Fiction podcasts, like radio plays of the past, use ambient sounds to set the scene, sound effects to represent characters’ movements and actions. And while the Foley artists of the past used physical objects and sometimes their own voices to duplicate footsteps, thunderstorms, and gunshots, the sound designers of today’s fictional podcasts have a vast array of digital tools and techniques for sonic world-building.

The Truth podcast, produced by Jonathan Mitchell, is a collection of short radio dramas, described on its website as “movies for your ears,” and features skillful sound design to convey setting and situation, typically without any kind of narrative voiceover. A private home, a lively dance club, and a lunar spacecraft are all rendered with recognizable sonic detail without characters needing to mention where they are. In “Visible,” which first aired in February 2016, written by Louis Kornfeld with Diana McCorry, a newly blind man experiments with a new seeing-eye app. The app has a female voice with the slightly stilted delivery of smart speakers and Siri, though its speech rate is slower than the pace more experienced blind people tend to prefer. It provides enhanced navigation information, telling the blind man exactly how far he is from the elevator

and announcing obstacles in the way. The man is able to venture forth from the medical facility where he is recovering from the accident that caused his sight loss, and cross a busy street without injury. These different environments are rendered with recognizable sonic detail. In the park across the street, the app counts and describes the song sparrows the man hears chirping. And then, as if to show the progress he's making at aural interpretation, he is the one who discerns their departure, from the collective fluttering of all their wings.

The man then asks the app to take him to an art gallery. There he asks the app to describe an abstract painting. The app does so, following the conventions of verbal description of art, mapping the composition and naming all the colors: "This picture is an abstract arrangement in a mosaic pattern composed of intersecting vertices of colored squares, alternating in hue from dark brown at the periphery to light orange in the center." The man is dissatisfied even while the app insists that it has provided an accurate description. A gallery employee offers to help. Her initial description also follows the conventions, with some art historical references thrown in: "It's an abstract mixed media in the style of Bauhaus, which captures some of the spirit of late Klee." The man interrupts her to ask how the painting makes her feel. She then provides a much more subjective and emotive description: "As I said, it's an abstract. It's warm and small and personal. I think it does a nice job of evoking a sense of memory. There's something bittersweet to it. When I look at this it makes me think of the feeling you have when you're just drifting off to sleep. Do you know the way one thought will become very vivid in your mind and then it vanishes and you can't remember what it was? There's that afterglow. That's what this picture feels like." The man thanks her, adding that her response is exactly what he was going for when he painted it, thus revealing that he is the artist.

The story somewhat self-consciously draws attention to a basic aspect of the podcast medium—that the audience derives all meaning from the combination of text, vocal performance, and sound design. Significant visual details must be communicated in words. The blind man's adaptation is presented as a reorientation to aurality and an assessment of the limits of assistive technology. When it comes to visual art, human interpretation is still preferable to artificial intelligence.

Depictions of blindness in fictional podcasts may be rare, but this kind of self-reflexive move to image description occurs elsewhere. *Within the Wires*, written by Jeffrey Cranor and Janina Matthewson, "is an immersive fiction podcast using found audio from an alternate universe," the website announces. The "found audio" includes relaxation tapes, black-box recordings from an airplane, and voicemail messages. Each season features a single character narrating the story in brief segments. Sound design is limited to a rendering of the sounds of sometimes obsolete audio recording and playback equipment—the click and whir of tape recorders and slight distortions of the speaker's voice. The ten episodes

of the second season are formatted as audio guides to museum exhibits of fictional artist Claudia Atieno, narrated by her friend and fellow artist Roimata Mangakahia. From the biographical details and the narrator's increasingly emotional delivery, the listener pieces together a story about Atieno's life and mysterious disappearance and death, but the script includes lengthy descriptions of paintings that follow the standards for image description for blind and visually impaired museum visitors.

Here are two humanoid forms, one holding the other. On first glance, the erect figure in the background appears to be cradling the limp figure in the fore.

Look at their faces. Or at least the indentations that replicate human faces on each figure. The one being held has almost no countenance, perhaps a shadow for eyes, and the grayish blob to the left of that could be a distorted mouth, caught midway into a cry or a song. (Season 2, cassette 10)

While the imperative—"Look"—implies that the audio tour is intended for a general rather than blind audience, the speculation about what "perhaps . . . could be" represented would make the text equally informative for blind people. And of course, since the actual audience of the podcast, whether blind or sighted, cannot see the fictional paintings being described, as in the example from *The Truth*, the script self-consciously draws attention to the nonvisual nature of the medium.

In the *Everything Is Alive* podcast, produced by Jennifer Mills and Ian Chillag, each episode features a different putatively inanimate object—a can of soda, an elevator, a bar of soap, among many others—interviewed by the host, Chillag. The sound design duplicates that of a standard interview show. Occasionally the host calls on an outside expert whose remarks are recorded as if on the phone. On some occasions, the host asks the object to describe its visual appearance: "Do you just for . . . I guess for our listeners that can't see you . . . Do you just want to describe what you look like, for us?" But here description intertwines aspects of the object's lived experience and personality. For example, the episode titled "Connor" begins this way: "My name is Connor and I'm a painting, a portrait painting of President William Howard Taft. You see him from the waist up and he's wearing a double-breasted suit, sort of leaning back looking at the left foot, your left foot but my right. I'm a good painting, not a great painting, you know, I'd admit to that." When Chillag asks him how he distinguishes himself from President Taft, the subject of the painting, he explains, "No, I'm an image, it's an illusion, it's an optical illusion, it's an easy mistake made, I'm not him but I look like him, or I look like, part of me looks like him. . . . You know, a little bit of me looks like the wall behind him." The episode then goes on to explore the painting's inner life: its mild resentment that other portraits in the

gallery are more popular, its sexual fantasies about the *Mona Lisa*, its trauma at once being touched by a toddler. Much is conveyed by the actor Connor Ratliff's voice, which is of medium and slightly nasal tenor. His tone is of someone who accepts his lot in life while still harboring longings for greater recognition.

Taken together, these three randomly selected examples could serve as models for image description of art, suggesting that description that attempts to be authoritative or to deploy art historical terms may be less effective than something more emotive. Why not ask the painting to describe itself? But these examples (particularly the first two) also provide models for would-be screenwriters of naturalistic situations where description occurs in the dialogue and so need not be added on separately as part of an AD track.

Although it is unlikely that mainstream podcast producers think about it this way, blind people are an ideal audience for the medium. Indeed, there are a number of informational podcasts produced for and by blind people, such as *Reid My Mind*, hosted by Thomas Reid, and *Be My Eyes* and *Say My Meme*, hosted by Will Butler, which treat topics of interest to blind people, including AD. In claiming that blind people are an ideal audience for podcasts in general, I am not subscribing to the myth that blind people are compensated for lost sight with augmentation of their other senses, especially hearing. Rather, people who have been blind for any length of time learn to attend to the sounds around them for general information, navigation, and so forth. Anyone, blind or sighted, interested in honing their listening skills should consider *The World according to Sound* by Chris Hoff and Sam Harnett (with support from the Lighthouse for the Blind and Visually Impaired in San Francisco).

Each ninety-second episode focuses on sounds in a given environment or situation. These may be the sounds of nature (mud pots or giraffes after dark) or they may be human-made sounds—the tinkle and hum of beer bottles on a conveyor belt in a brewery, the crunch and rustle of an artist cutting and crumpling heavy paper. The audience can listen to a horse walking, trotting, cantering, and galloping on a treadmill, or the sound of liquid flowing into a bottle through a funnel. The sounds are offered with little to no commentary, and the aftereffect is to heighten the listener's aural awareness and perhaps allow them to imagine scenarios where this sound would be prominent.

Fictional podcasts, movies for the ears, offer models for soundtracks that require little or no additional description. If AD turns Netflix videos into audiobooks, narrative podcasts offer models for soundtracks that describe themselves.

As delighted as I am, as a blind consumer of media, with the accessibility of these podcasts, I have to note that the medium is not accessible to people who are deaf and hearing impaired. Some podcasts provide transcripts, though it is not always easy to find them, and they are typically working scripts for the voice artists, without any rendering of nonverbal sound, music, or actors' tonal

qualities. So these transcripts fall short as access tools for many of the same reasons that captions for film and TV do. The captions often don't mention nonverbal sounds or music, or else are inconsistent in supplying this information. And even when sounds are mentioned, they are merely named rather than described. So, as often happens in the world of media, something that is accessible to one group may be anything but accessible for another.

Alt Text as Poetry is a collaborative project from disabled artists Bojana Coklyat and Finnegan Shannon designed to raise awareness about accessibility in the online world. The main focus is on alt text, or the kind of short-form AD that can be added to a website or social media post to provide visual information to people who use screen readers. But the two founders of the site also commissioned disabled composer JJJJerome Ellis to create music for the site, and in the interest of promoting access, they provide a verbal description of the music, in both one- and five-minute versions: "A melodic whirring begins slowly, with sounds akin to playful electronic bells dot [*sic*] the soundscape. Another beat, which feels like a very quick rushing back and forth is added to the mix. As the music warms up, a voice quickly joins in with a, 'Yo.' Fingers snap twice and add more texture to the layered composition." While a five-minute verbal description of music would be beyond the scope of real-time captioning, it suggests the possibility of a secondary source of information the consumer could consult asynchronously. Filmmakers who focus on sound design as a facet of nonvisual access would automatically be in a better position to describe significant sounds and include this as part of the captions.

Alt Text as Poetry goes beyond web accessibility compliance to model the ways that access can be an occasion for artistic innovation. As people both with and without sensory disabilities consume media in multiple modalities, producers like you who are interested in broadening their audience and making their creations more equitable would do well to explore new ways to present material. These new forms of media will probably not look or sound like older forms and will be produced with a heightened attention to both visual and aural aspects. What starts as an accommodation for people with sensory disabilities can become a new art form, with new possibilities of enjoyment for a much larger audience.

BIBLIOGRAPHY

- Butler, Will. 2022. *Be My Eyes*. <https://www.bemyeyes.com/>.
- Butler, Will. 2022. *Say My Meme*. Podcast. <https://www.bemyeyes.com/podcasts-show/say-my-meme>.
- Coklyat, Bojana, and Shannon Finnegan. 2022. *Alt Text as Poetry*. <https://alt-text-as-poetry.net>.
- Cranor, Jeffrey, and Janina Matthewson. 2017. *Within the Wires*. Podcast, second season. Presented by *Night Vale Presents*. <http://www.nightvalepresents.com/withinthewires>.
- Fryer, Louise. 2016. *An Introduction to Audio Description*. Routledge Press.

- Hoff, Chris, and Sam Harnett. 2022. *The World According to Sound*. Podcast. <https://www.theworldaccordingtosound.org/>.
- Kornfeld, Louis, and Diana McCorry. 2016. "Visible," In *The Truth*. Podcast, produced by Jonathan Mitchell. <http://www.thetruthpodcast.com/about>.
- Mills, Jennifer, and Ian Chillag. 2019. *Everything Is Alive*. Podcast. <https://www.everythingisalive.com/>.
- Reid, Thomas. 2022. *Reid My Mind*. Podcast. 2022. <http://reidmymind.com/>.