

**SAMANTHA SAULD:**

Thanks for joining this webinar entitled "The Best Way to Create Audio Description." I'm Samantha Sauld from 3Play Media, and I'll be moderating today. I'm joined by James Herndon, who is a technical writer and accessibility consultant at Equal Entry. And with that, I'll hand it off to James, who has a wonderful presentation prepared for you.

**JAMES HERNDON:**

All right. Let me just share my deck here. So thank you, Samantha. Hello, everyone. My name is James Herndon. And as Samantha mentioned, I work for Equal Entry as a technical writer and accessibility consultant. My specialties are media alternatives. That includes text alternatives, captions, and, of course, audio descriptions. And I also write and produce a number of training videos about accessibility for various companies.

In the past few months, I have been working here in Atlanta with several people at the Georgia Libraries for Accessible Statewide Services. And we've been working together to determine the best ways to write and provide audio description for 360-degree video experiences in VR. And I will be giving a talk about that at the CSUN Assistive Technology Conference in March. And we hope to make that talk available online. So audio descriptions have been on my mind quite a bit lately, and I'm excited to dive into today's webinar and share my thoughts with you all about the best ways to create them.

So for the structure of this webinar, we're going to begin by talking about what audio description is, just so we can move forward with a clear definition in mind. After that, we're going to look at the types of pre-recorded audio descriptions that are typically used. There are other types of audio description, but we're going to be limiting our focus to the pre-recorded ones. Then we're going to move on to best practices and talk about some helpful abstract principles you should keep in mind when you're writing and narrating audio descriptions.

After that, we'll look at some examples of what I consider good audio

description. And we'll do that to help make the best practices feel more tangible. I'm the type of person who always wants examples, and I hope that these examples will help you all appreciate the principles of good audio description. Following that, I want to give you all a chance to briefly write some audio description of your own. I'll share a GIF image and ask you to write some audio description in the chat window, and then we'll talk about what people shared.

I want to tell you all upfront that there's no need to be shy. I'm not going to call out anyone's text as a bad example. We're going to keep it positive and fun and educational, so please join in. And then finally, if there's time for questions, I will do my very best to answer them. And with that, let's get started.

What is audio description? Simply put, audio description is a secondary narration track intended for blind and low-vision consumers of visual media. Now, this is typically divided into two types-- pre-recorded, which is what we're mostly going to be focusing on, or live. Live, you're talking about theater. Someone goes to see *Hamilton*, for example. They may have the option of wearing a pair of headphones through which a remote describer might be seated up near the sound person, who is providing real time descriptions about what's happening visually onstage.

And that can happen in any type of theater experience, or maybe even a sporting experience, anything where somebody needs real time description of what's happening, similar to ASL. And that is something that is increasingly common, but it is a whole other set of rules and set of circumstances and set of best practices that we won't have time to get into. So we're going to keep our focus on the prerecorded, which is typically what you are going to end up having to work with if you are working in the digital world and need to provide audio descriptions yourself.

And I'm guessing a lot of you have probably already experienced audio description in some form or another. If you've subscribed to Netflix, you

may have even noticed when you go to click on the Language Settings, you see English, Spanish, and so on. There's often an option for English with audio description track. It'll say English, dash, audio description.

If you've never seen that, next time you're streaming something, check on the language settings and see if you have that option. And if you do, go ahead and open it and listen to it. Honestly, the best way to understand something like audio description, like captions or any other aspect of accessibility, is just to take a look at it for yourself. And I think you'll find that once you experience it a little bit hands on, it's much simpler and easier to get a handle on than you might imagine.

Who else benefits from audio description? This is similar to other aspects of accessibility where we often think about what we're doing as a kind of remediation for, or benefit for, a very small group of people with a very specific set of needs. But that, like other aspects of accessibility, is really just not the case. It tends to benefit a much wider group of people than initially imagined. And the people who benefit are often people who didn't realize they would benefit until they encountered it.

In the case of audio description, we've had a lot of feedback from people on the autism spectrum. Many people on the spectrum find that audio description helps them better understand emotional and social cues that are only demonstrated through actions or facial expressions that happen visually on the screen. And getting the audio description reiteration helps to reinforce that.

Auditory learners-- 20% to 30% of students say that they retain information best through sound. So if a person is watching something for educational reasons, listening to it or watching it with the audio description turned on helps them retain the information. Language students-- listening is a key step in learning a language and associating it with appropriate actions and behaviors. I have European friends who have told me that watching Netflix with the audio descriptions on has improved their English simply because they are getting additional

context, additional description, additional language education, of course.

And then more broadly, literally anyone desiring more flexibility in their media consumption can benefit from audio description, anyone who wants to enjoy videos in eyes-free environments. I live in Atlanta, Georgia, and we are not known for our pleasant traffic. And a number of times when I know I've got a long traffic day, I've downloaded a number of Netflix shows onto my phone, and I've put the Bluetooth setting on in my car, and with audio description on, I can listen to shows as if they were an old radio drama.

And it's great. It's like a very detailed podcast, and I really enjoy it. And I think it's yet another example of how accessibility benefits everyone. And even if it's something that you don't imagine will benefit you right now, it's something that is likely to benefit you later on in your life.

So now we're going to look at the types of pre-recorded audio description. There are two main types. The first, and the one that I believe is most common, and the one that I imagine a lot of you will be asked to produce the most, is the standard audio description. And that is when a narrator talks through a presentation, describing what is happening on the screen or stage without interrupting dialogue. And that's usually what we use for movies or other digital shorts.

And it's standard. And we call it standard because in most cases, it's awkward to pause a movie or a video and provide an extended description. Sure, we could do that, but it can kill the momentum and suspense of a movie or a video. And the goal in accessibility is to create an equivalent experience, right?

And so if we're trying to do that, it's a bit of an ask to provide this so-called equivalency to somebody and ask them to spend extra time that somebody else doesn't have to spend. Somebody with a visual disability, if they have to devote 30 minutes to a video that a sighted person only needs to devote 20 minutes to, that's not necessarily an equivalent experience. So when we can, it's generally best to provide

standard audio description unless it's truly, truly not possible to. But we'll get to extended in just a minute.

Now, in standard audio description, you will find that it is most common for description to be inserted in pauses in dialogue. And we're going to look at an example of that in just a second here. This is a show called *Chewing Gum*. And I want you all to listen to how the narrator slips in descriptions of what is happening visually but is not communicated in dialogue and how the narrator does that in between the pauses in the speech.

[VIDEO PLAYBACK]

- I'm nervous.

- Start with the right levels of eye contact,  
and the rest will be easy.

- Yeah.

- It's the difference between this--

- Candace stares ahead.

- And this.

- Candace purses her lips and rolls her  
neck.

- Wow.

- It's the come to bed face.

- What?

- Neutral face.

- Tracy stares ahead.

- Come to bed face.

- Tracy raises a brow and smiles.
- Uh, OK. A little bit less.
- Tracy shakes her face, smiles, and bats her eyelashes.
- I feel sick.
- OK. OK. And--

[END PLAYBACK]

**JAMES HERNDON:**

Another common place to put in a standard audio description is in a musical sequence or montage. This is sort of the same principle behind inserting it in pauses in dialogue, but you have much more time freedom, because the musical sequence or montage is going to be not just a few seconds-- you know, one to three seconds-- it could be 20 seconds, a minute, 2 minutes.

And you have a lot more time to get more detailed, and potentially, if you're a good describer, much more vivid. And so this is where you will find a lot of describers either playing catch-up or trying to truly immerse the listener as deeply as possible in the description. And this is a very common area to really, really dig deep into the description.

I'm going to show you a clip from the animated version of Disney's *The Lion King*. And pay attention to the difference in the detail of the description in this clip versus the previous clip where we were timed down to fit descriptions just between moments of dialogue, whereas here, we have a musical sequence.

[VIDEO PLAYBACK]

- (SINGING) It's the circle of life.
- Hundreds of animals gather at the bottom of the Pride Rock, a tall, flat ledge that towers over the rest of the savanna.

Zazu, a small blue bird with a large beak, flaps to the ledge. He bows to Mufasa, a powerful, dignified lion with a thick red mane.

[END PLAYBACK]

**JAMES HERNDON:**

All right. Now we move on to extended audio description. This is a secondary version of the media in which the video is periodically paused automatically, automatically meaning programmatically. There's not somebody manually doing the pausing. This is set in the code of the video. And during those pauses, a narrator provides a detailed audio description of what is happening onscreen, and then the video programmatically resumes playback.

I'll show you some examples of those. Here is an example of some extended audio description.

[VIDEO PLAYBACK]

- Jane presses the button on the pedestrian crossing. Above the button she presses, there is a bicycle and a man icon. They are lit up and red. When she presses the button, these icons turn from red to green.

- When you are ready to cross the road, press the button and wait for the light to turn from red to green. Don't cross until the light has turned green.

[END PLAYBACK]

**JAMES HERNDON:**

So here we have a video, the purpose of which is to demonstrate how the crosswalk works. And I think it's fair to argue that the light is an integral part of that, right? So maybe you could get away with using standard, but I don't know that you're providing an equivalent

experience when the point is to provide a clear sense of how this crosswalk works. And I think it's safe to say that the light is an integral part of that. And so this is a case in which extended is necessary because we need time to describe what is going on with that light.

Here is one more example.

[VIDEO PLAYBACK]

- Work order interface modifications--  
Facility Management Software System.

[MUSIC PLAYING]

A screenshot of the Work Order List Screen in FMSS with work orders populating the list.

- The next time you open Work Order Tracking, you may notice some changes designed to simplify work order entry and modification.

- The same work order list screenshot with two tabs circled-- Simple Work Order and Full Work Order.

- The biggest change that you will notice is the option for a Simple Work Order screen.

[END PLAYBACK]

**JAMES HERNDON:**

Now, this is a complicated user interface, right? And the person who chose to make this video, they are adding pauses so that they have more room to provide audio descriptions that describe that complicated user interface. Personally, I think this is the type of video that could probably be reshot in such a way that standard audio description could

work just fine.

I also think it could be rescripted in such a way that no audio description would be needed, because if this shot were to be zoomed in to the upper left and the narrator said, I select Simple Work Order, the second of eight tabs, that's something a person who uses a screen reader can follow. Right? And that's an example of a situation in which audio description would not be necessary because the equivalent experience is already provided in the dialogue. And we'll talk about that more in just a minute.

So which should you use, standard audio description or extended audio description? As I say here in the bullet, the decision really depends on the ratio of pauses in the audio to the complexity of the visual material. If you have a lot of gaps in the audio, it should be fairly straightforward to use standard because you have plenty of pauses in which to insert your narration. If it's a business presentation and somebody is showing PowerPoint slides, similar to what I'm doing now, and there's a constant visual action happening, that does pose more of a challenge.

And of course, the more visual stylizing you have going on, that's going to make it even more difficult. It's not always the case, and I'm going to hope to prove that to you as I get into some of the examples later. But it really does depend. 3Play Media has an excellent quiz about standard versus extended audio description if you all look on their site for that. I highly recommend-- if you're unsure about which one for your specific piece of content you should do, take that quiz and see which one it points you to. And that might help you get a better sense of direction for your particular project.

All right. So let's move on to some best practices and get a sense of the principles that we're working with. Now, as I was just saying with the Work Order Modifications video, really, the best thing we can do with audio descriptions and writing them well-- or I should say keeping them in mind when we're working-- is try to script in such a way that the audio descriptions aren't necessary at all. If we plan ahead and we write

dialogue and describe things in such a way that what is happening onscreen is indicated by what is being spoken, we don't need audio description. The equivalence is already there. Audio description is what we need if what is happening onscreen is not accounted for.

And I show a picture of Mr. Rogers here because some of you may be familiar with this story. In an early episode of the show, Mr. Rogers was talking about his fish and mentioned that he was feeding them. And then as episodes moved on, he might be telling a story about something else, and casually shake some fish food into the feeder, and move on to do something else. And he got a letter from a blind girl who said she was very worried about the fish that he had mentioned a few episodes earlier, and were they getting fed?

And so from then on out, every time he walked by the tank, he would stop and pick up the fish food and say, I'm feeding the fish. And you noticed he did it several other ways on the show where if he would pick something up, he would say, I'm holding whatever it was he was holding, or I'm about to do this. He would narrate what it was that he was doing so that blind children watching the show would be able to understand what was going on. And that's an example of somebody in the scripting phase planning ahead so that audio description wouldn't be necessary. And that really is a kind of best case scenario.

Now-- and that's not to say too that you can't also do a mix of the two. An example of that is *The Great British Baking Show*. This is a show where much of the show is not audio-described because the pauses between dialogue are so brief. If any of you have seen the show, it's just constant patter. However, the contestants on the show often describe what they're doing while they're doing it. And because of that, audio description is not necessary, because what they're saying is what you're seeing onscreen.

I'm going to show you a clip here. It's two parts. I will show you a brief clip where the principle I've just described is illustrated. What the characters-- excuse me, contestants are saying is indicated by what's

happening onscreen. And therefore, audio description is not necessary. And then I cut to a clip of the judging segment of the show. And you have facial expressions of contestants which are important to understanding the show, and those facial expressions are not described. And so an audio describer steps in to describe the facial expressions. So let's watch those two clips.

[VIDEO PLAYBACK]

- Right, so I've got the genoise. Tip the eggs and sugar into a bowl. Whisk the mixture until it's reached 43 degrees.

- (SINGING) I've never met a genoise.  
Stop it! [LAUGHS]

- If the bakers don't whisk for long enough--

- There is where I get a good, strong arm.

- They won't get enough air into the genoise sponge to create the volume that Pru is expecting.

- What do you think?

- I think it's wonderful.

- Oh, thank you.

- The balance of spices that you put in is perfect. It's nice and moist inside there, and the design as well. It's faultless, actually.

- Faultless?

- Faultless.

- Thank you. Thanks very much.

- Michael and Henry beam as Michelle screams silently with excitement.

- Oh, bless you.

[END PLAYBACK]

**JAMES HERNDON:**

Now, as I mentioned, you have the contestants enacting the dialogue, right? So for the genoise, tip the eggs and sugar into the bowl. Whisk the mixture until it reaches 43 degrees. We see the whisking. We see the thermometer saying 43 degrees.

That part doesn't need audio description because what's happening onscreen matches the dialogue. Audio description is saved for the silent facial reactions. "Michael and Henry beam as Michelle screams silently with excitement." And in that moment when audio description is necessary, what's happening is described in the present tense as it's happening, and the active voice is used. And the narrator sounds confident, interested, warm, and authoritative.

Now, if audio description is necessary for the full video, that's where certain principles need to be kept in mind. The DCMP guidelines-- that stands for the Described and Captioned Media Program. The description must be accurate, prioritized, consistent, appropriate, and equal.

When we say accurate, we're talking about how there must be no errors in word selection, pronunciation, diction, or enunciation. That seems fairly straightforward, but it's really asking you to not freestyle your audio description. Don't play your video, press record on your microphone, and wing it. Plan it and script it the way you would anything else professionally.

Prioritized-- content essential to the intended learning and enjoyment outcomes is of primary importance. Now, that is essentially asking you

to triage the content. And while in many programs, there is a lot happening onscreen and a lot to describe, it is asking you to make some judgments about what has to get across in order for the experience not to break down, what absolutely must come through.

Consistent-- both the description content and the voicing should match the style, tone, and pace of the program. That one can get tricky, and I've certainly seen things where that was not the case. And it's interesting to think about when you see it, because I'm sure-- if all of you have watched audio description or if you begin watching it, you will definitely notice as soon as it happens. Because one of the hallmarks of good audio description is that consistency and is that sense of the audio description drawing attention to the program and the content and not to itself.

An example might be, say I have a client who is the owner of a meditation retreat in Hawaii, and they want me to write audio description for a bunch of footage of people in Lotus poses and slow-motion waterfalls and so on. And I write very accurate descriptions of all of that, but the voicing I use is my best Freddie Mercury impression.

They're probably not going to be happy with me. I say probably. Never say never. But I'm guessing they're not going to be happy with me because that doesn't match the style, tone, and pace of the program. They're going to be looking for a voice that conveys a sense of calm and serenity because that matches the tone of what they want.

Appropriate-- consider the intended audience, the objective, and seek simplicity and succinctness. That's asking for, of course, clear communication, but it's also asking you to consider who your audience is. And if you find yourself teetering between one possible description path and another, ask yourself, who am I speaking to? And then equal-- equal access requires that the meaning and intention of the program be conveyed.

Now, the company I work with, we have three guidelines that we come

back to a lot for audio description. We say three arbitrarily because it's simple and direct, and the more I work on audio description, the more these three things seem to be what come up again and again and again. And they are, number one, ground your descriptions in the most familiar terms possible.

I will re-emphasize that as we go through some of the examples in just a minute. But you're going to see that given how visually stylized and surreal some videos can be, your descriptions can be-- it's very easy to go off the rails quickly and become very florid in your descriptions. And while that may be fun and offer you poetic license, and the writing may be technically good, it's not accomplishing the task. The task is to convey what is happening onscreen immediately for somebody who is listening and wants to experience what's happening in real time.

Two-- only describe the scene, not the unseen. And what I mean by that is, when you're seeing facial expressions and people are lifting their eyebrows or frowning or doing other things, you can trust the person listening to associate emotion or meaning to those expressions.

You can describe the-- you can describe the facial expression, but you don't need to tell the person listening that the character, for example, is furious at so-and-so. You can say that they are frowning, or that they are wrinkling their forehead, and trust that the person listening can say, oh, that indicates they must be upset. If you are getting too deep into meanings or intentions, and so on, you're bringing a little too much subjectivity into it. And the goal is to be objective, as we just talked about.

Three-- always ask yourself, if my video were aired as a podcast, would it make sense? That's honestly one of the best things you can do if you're trying to decide-- particularly if you're trying to decide, do I need audio descriptions at all, if you're trying to see, could I-- can I get away with not having them for this? Did I sufficiently describe what's happening?

If you're wondering that, share your video with somebody who's never seen it. Ask them to bring it up on their laptop with the laptop turned around in the other direction so they can't see the screen, and say-- play it, and then ask them about it afterwards, and see, did they get everything? Was it all coming through? And if it did, great. But if a lot was lost on them, then maybe it didn't come through.

OK, so now the fun part. Let's look at some examples of what I consider to be good audio description. Here's the first example. This is a scene from the movie *Okja*. We're going to play the clip, and then I will talk about why I think it's good.

[VIDEO PLAYBACK]

- (SINGING) This fun world.

- A view looks out over a foggy mountain range beneath a sky filled with stratus clouds. A title appears-- *Okja*. In a lush woods, a young teenage girl peels apart a pod of white tufted seeds and blows, sending them airborne. Text reads, "10 years later, far from New York."

More seeds take flight as she blows again. Short, dark brown hair frames the carefree smile on her face. Behind her, a full-grown superpig approaches from the brush. It's gray and the size of a small elephant with a pig-like midsection, long, floppy ears, and a snout like a hippo. A scar runs from one of its nostrils toward its mouth. It nuzzles the girl, who wears a splotchy pink, lavender, and blue jacket.

[END PLAYBACK]

**JAMES HERNDON:**

So the first thing I want to call attention to is the fact that all of the onscreen text in this clip was announced. We had the title, *Okja*. That was announced. And we also had the time and place indicator-- "10 years later, far from New York." Another thing you may have noticed

was that all of the onscreen text was announced as onscreen text. And that helps avoid any confusion because it distinguishes onscreen text from the narration track itself.

The second thing is that all descriptions are grounded in the most familiar terms possible. We start with a foggy mountain range, which is simple enough. Then we have a girl peeling apart a pod of seeds and blowing them away. Even if you've never done this or never even seen someone do this, it's still, I would argue, a tangible image.

And then we get to what could be a tricky part, which is providing an audio description for Okja. The narrator could have accurately described Okja in very alien terms, but that probably would not lend itself well to quick absorption and comprehension by the listener. So what the narrator says is "a full-grown superpig approaches from the brush."

And now, the use of the word "superpig" touches on my third bullet, which is matching the vocabulary to the content. Superpigs are the official name of the species in the movie. But to the person who can't see the screen, superpig doesn't mean anything. So what the narrator does is after grounding us in a familiar image of a young girl playing in the woods, he essentially defines the superpig visually in the most familiar terms possible.

He says it's gray and the size of a small elephant with a pig-like midsection, long, floppy ears, and a snout like a hippo. These are all pretty familiar ingredients that the listener can combine in their mind into a strange new thing in the imagination, and that thing is kind of likely to resemble what a sighted person can see on the screen. And that means that an equivalent experience has been provided, which is what we want.

And I would argue that because the content is described so simply, clearly, and concisely, the narration is easy to comprehend. And it seemed that at first glance, it might make a new audio describer say, oh, I'm going to need at least seven minutes to describe all of that. With

the right word choice and the right sense of how to transition from the familiar to the unfamiliar, standard audio description works great and provides a true equivalent experience.

All right. Let's move on to the next example. This is a clip from the movie *Black Panther*.

[VIDEO PLAYBACK]

- Beneath the giant stone panther carved into the mountainside, a tunnel leads to the cavernous vibranium mine, a hive of activity. An aircraft resembling a dragonfly rises over maglev trains and perches on a mezzanine walkway. Goods elevators rise to the opening in the mountaintop where W'Kabi stands with some of his men. Crates are loaded into the elevators as W'Kabi walks with Killmonger.

- Everything is on schedule.

- Have the spies been alerted?

- Yes. Some resistance to our new nation, but the war dogs in London, New York, and Hong Kong are standing by.

[END PLAYBACK]

**JAMES HERNDON:**

Just in case anyone's wondering, that nod that Killmonger makes in that last moment is described a second later. So what I really like about this one is in a lot of these movies where the camera takes you on a CGI roller coaster ride, it's very tempting to-- when you go to describe it, to think about it as you being on the roller coaster.

And we're flying under-- we're flying under a panther in the mountain, we're going through a tunnel, and now we're flying up into the mine, and

we're flying up toward the surface of the mountain. And I think if I were a new describer, that's how I might be tempted to describe it. But that is problematic because when you think about the primary purpose of audio description, it's to provide an equivalency for people with a visual disability.

And if you put yourself in that position, it's not very difficult to imagine that if you're a person with a visual disability, getting an equivalency that constantly reinforces the message we see, we see, we see, that's probably a pretty irritating thing to hear over and over. Right? And it's beside the point as well. The point of audio description isn't to describe what the audience is experiencing.

The point is to describe what's happening on the screen as accurately as possible. And that's why I love the beginning of this. "Beneath the giant stone panther carved into the mountainside, a tunnel leads them to a cavernous vibranium mine, a hive of activity." It's present tense, just taking you through almost not even a point of view. And it's very effective of drawing attention to the presentation, like I discussed earlier, not to itself.

And I think the narrator isn't reminding you that you're watching or listening to a movie, and that makes it much easier to immerse yourself in the video. I think the language here is concise and effective. The narrator doesn't go into a lot of unnecessary detail about what W'kabi's men are wearing or how many crates are in the elevator at one time. We just get crates are loaded into the elevators as W'kabi walks with Killmonger.

There are surreal and unfamiliar elements happening all throughout that mine. It's all comic book fiction, fantasy. And we just get a hive of activity. That gives an emotional cue of what's happening almost more than the visual, but I think it conveys an accurate sense of what's happening more than a pointillistic, dot-by-dot visual description would. We have this impossible comic book aircraft that gets described as an aircraft resembling a dragonfly. Boom-- we've got something familiar, a

dragonfly. It perches. We've got a familiar gesture.

We might not have ever been on a maglev train. I certainly haven't. But it's not very difficult to understand that they're talking about from maglev, they're saying, oh, magnetic levitation. And that makes this comic book concept just familiar enough to be imagined and not get so confused that we lose track of what's happening.

All right. Let's move on to the next example. This is a clip from the movie *Coco*.

[VIDEO PLAYBACK]

- The skinny boy looks ahead and gapes in wonder. He and his late family members approach a massive bridge made out of glowing orange flower petals. Skeletons come and go. As they step onto the bridge, their glow disappears.

- Whoa!

- Come on, Miguel. It's OK.

- Miguel hesitantly walks through the magical entrance to the bridge and loses his glow. He notices the petals lighting up beneath his feet. Miguel grabs a handful of them, and they blow into the wind. Dante runs ahead.

[END PLAYBACK]

**JAMES HERNDON:**

All right. So once again, we have descriptions grounded in familiar terms. We have "the skinny boy looks ahead and gapes in wonder." The describer could have said Miguel looks ahead, but he, wisely, I think, takes this opportunity to ground us in something familiar, even more familiar than this particular character, a skinny boy looking ahead and

gaping in wonder.

The image feels more tangible and universal that way. And I think it calls attention to the listener. It calls attention for the listener to Miguel's apparent vulnerability in this situation. It also helps to ground us before the surreal bridge is introduced. He and his late family members-- that's another surreal image.

A massive bridge made out of glowing orange flower petals. Skeletons come and go. As they step onto the bridge, their glow disappears. These are all pretty surreal actions and images, but the narrator has grounded us in something very familiar before we literally step into the unfamiliar.

Now, what's different about this clip from the previous one, and what I really like about it, and I think is very important when creating audio description, is how it conveys a sense of what is real and what is illusory for the listener. The skinny boy, quote, "gapes in wonder." That tells us that what's happening onscreen is not entirely normal or common, even in in the world of this movie. Miguel walks through the magical entrance. That tells us we're moving from something that is ostensibly real, even if it's animated, into something that's surreal.

And finally, I think it does a great job of meeting the DCMP's appropriate criteria. We have that it's objective and it seeks simplicity, and it considers the intended audience. The vocabulary is simple and direct so the intended young audience could understand it. It doesn't make any value judgments, not even, I think, when Miguel is gaping at what he sees. It doesn't say the skinny boy looks ahead and gapes in wonder at the strangeness of it all. It's-- he just gapes in wonder. It's objective, and it's a simple and succinct description as well.

All right. Let's move on to our final clip here. This is a scene from the movie *Space Buddies*. I want to say upfront, there are some onscreen text happening in this one that is not read out loud. And interestingly, they are credits. And this is a children's movie. And I'm wondering if it is

because-- and I've noticed this in other children's movies-- credits are not read aloud if they are not timestamps like they are in *Okja*-- 10 years later, far from New York, or 8 years later.

So it may be that that is a growing trend in audio description for children's films. But anyway, we'll talk more after the clip.

[VIDEO PLAYBACK]

- Our view drifts through a star field to a space station outfitted with solar panels.

[CLASSICAL MUSIC]

A cratered moon looms in the distance.

[CLASSICAL MUSIC]

Light shines from a window in the bulky station. Inside, we glimpse a white dog in an astronaut uniform. A black spot surrounds the bull terrier's left eye. He talks to us.

- Dreams are like stars. You can't touch them. But if you follow them, they will lead you to your destiny.

[END PLAYBACK]

**JAMES HERNDON:**

Now, the first bullet I have here is the description content and voicing matched the style, tone, and pace of the program, like we've talked about with DCMP before. I love how matter-of-fact this description is of this outer space tableau that suddenly makes the deadpan reveal of a talking dog. This is such a good example of matching the tone to the content.

None of what's happening onscreen is presented as if it were extraordinary, and the audio description takes the same approach. It feels normal, familiar in the context of this movie. And it's what makes

the short description effective and a bit funny. But it's appropriate to the intended audience, as I've noted in the fourth bullet. It considers them.

This is a movie for young children. And they're likely to take this at face value, right? They're going to see this and think, yes, of course there's a dog in space. And this dog has something to say to me about dreams. So if you're a child with a visual disability, this audio description is probably going to be very effective.

The second bullet I have is was not tempted to fill every pause. This has had some of the longest silences out of any of the clips we've seen so far, right? But I don't think that extra description here would've given us a greater sense of immersion in this scene. If anything, the silences made it a bit easier to settle in.

And that's important to remember-- when not describing can benefit. If you're already provided clear, concise description, like we've been talking about, then you can give the listener some breathing room. Let them listen to the standard audio track of the video along with everyone else if you've already set them up.

And third, I have the excellent sense of timing. This is important because some of the descriptions in here are staggered slightly. And that's OK. You may have noticed when the describer said, "a black spot surrounds the bull terrier's left eye." When we heard that, the black spot surrounding the terrier's left eye was not yet visible.

So the narrator is trying to provide a vivid description of what's happening onscreen before the dog begins its monologue. Because once the dog starts talking, the narrator doesn't want to talk over the dog in standard audio description. So in order to do that, the narrator has to bump up the audio description in the timeline to a point in time before the thing being described onscreen has fully appeared onscreen. And it's OK to do that.

So if any of you are wondering, can I get away with standard? I don't know. I might want to try extended, just because I'm not sure. Keep in

mind that it is OK to slightly shift forward your descriptions. You can start a description a little bit earlier if you need to, or even a little bit after if you have to.

You have a little bit of room. So don't feel as if everything has to happen in the exact moment. It's going to be frowned upon for you to talk over dialogue unless it's the least consequential dialogue-- you know, five characters all shouting over each other in a way that's going to be indecipherable to most anyone listening. But otherwise, it's OK for it to be staggered slightly in order to make it work in the time amount given.

Now, I've given you an overview of some principles of best ways to create audio description and some examples of clips that do it very well. But if you want to drill down at a very granular level about audio description in terms of examples with respect to legal standards and some technical ideas for plugins and other things that you can use for adding it yourself, I highly recommend checking out the 3Play Ultimate Guide Audio Description. You can find that on 3Play's site. It's a great resource.

All right. So now, as promised, this is your turn to write a bit of audio description. What we're going to do is, as I mentioned before, I'll show you a GIF file. And in the chat, go ahead and just write a bit of description. Don't worry about it being perfect. Just do what you can. And then Samantha will share with me some of your suggestions, and we'll talk about them. So without any further delay, let's move on to the first one. Take about 30 seconds.

**SAMANTHA SAULD:**

Someone says, a fuzzy monster puppet is inside a trash can. Another person says, Oscar dances and talks on the phone while turning pages. A person says, Oscar, a green, furry puppet, is in a trash can and rummaging through a book. Another one says, hunched over a garbage can in an alley, a fluffy green puppet with bushy brown eyebrows pages through a newspaper.

**JAMES HERNDON:**

Those are great. I really like, of course, number one, that you've

identified the character. I think it's very important in audio description to take a people or creature first approach and not get too lost in details. So I think you all did a great job in identifying the creature itself, if not by name, then by the fact that it's this shaggy, fluffy creature.

I like the sense of textures that you've all provided, too, about what the look and feel of Oscar might be. That's something that it's easy to take for granted as a sighted person. We all know-- those of you who are sighted listening who are accustomed to experiencing Oscar as this shaggy green puppet with an irritated facial expression, it's easy to overlook those types of things, or you take them for granted. So I really like that all of that is included. And I think that's excellent.

Let me show you what I came up with. I said Oscar the Grouch, a shaggy, green puppet with an irritated facial expression who lives in a trash can on Sesame Street, dials a number on a flip phone. I think in that GIF, it's not entirely clear what he's handling, because I think it's made of the same felt material like him.

So it's not a flip phone that any of us have handled before, so totally understandable that it's not clear what it is. If I had very little time-- if this-- say I had to-- say you have two seconds between dialogue-- I might say something much shorter, like Oscar the Grouch dials a number on a flip phone. But I think a lot of the descriptions I've heard would work very well too.

All right. Let's do one more. This one is going to be a little bit trickier. This is the singer Bjork. I chose a GIF from one of her music videos because they're known for being a little bit strange. And I thought that it would be interesting to take a GIF of something that would be a little more visually uncommon and see what you all came up with. So let's take another 30 seconds or so, and then Samantha, let me know what people came up with.

**SAMANTHA SAULD:**

So some of the responses are, a woman wearing a mask plays a flute while standing in a field. Bjork sits in a surreal landscape playing flute

amongst floating monsters. Young woman in flowing gown-- sorry-- plays flute in a fantastical environment with floating objects. Another one says, Bjork plays a flute. She wears a lush pink dress with a skull mask in a field. Objects surround her in the background.

**JAMES HERNDON:**

Those are great. I love how most of them have called attention to what she's doing, which is playing the flute. I think her action is important in this scene. She's not just sitting there resting in this grass. She's doing something, and that action is described. I think that's important, and I love that that is called out.

I think in this case, these background creatures are very prominent, and I don't think they could be called secondary. So I like that they've been called out. I think that's great. And I think her dress-- in this case, I don't think it's at all superficial to call attention to it, because I think it is so puffy and prominent and matching of the color scheme that's happening that it feels significant to me, at least, to the feel of what's happening to call attention to the look of it.

And if I were listening to this without being able to see it, I would be very curious to know what Bjork was wearing while such a strange song was playing. And given the other strange elements in the video, I would be very curious to know what she was doing and wearing while all of this was happening. So I think those are great.

What I came up with was "the musician Bjork plays the flute while wearing a puffy pink dress at the edge of an island cliff at sunset while magical periwinkle and sapphire cephalopods undulate in the air nearby." And again, if I had very little time, just a quick crunch, that I had to say something quick, I might come up with something like "Bjork plays the flute outside for magical creatures."

All right. We've got a little bit of time for questions. Samantha, does anybody have any questions? I'll do my best to answer them.

**SAMANTHA SAULD:**

Yep. So let's get started. The first question is, are there media player differences that captioners or web designers need to keep in mind? For

example, do some players support extended audio description better than others?

**JAMES HERNDON:**

There are going to be differences. I would recommend working with-- if you're working with your employer, I would recommend finding which one they are going to refer to or which one they would prefer.

I would generally recommend going with one of the larger supported players, like a YouTube or Vimeo type of a player, just because there's going to be-- and honestly, at this point too, I think it's also a good idea to-- I think it's becoming more common to simply just create a secondary version of the video without the toggle.

I know it's becoming more common to toggle audio descriptions on and off, like with captions. But I think if you are in an environment where you don't have video technology to do that, or you are limited to technology that doesn't have-- I'm sure you've all maybe seen embedded videos where you have that AD option to turn them on and off. If you're in a place where you don't have that, it's perfectly OK to provide a second embedded video in which audio descriptions are available and just share two versions of the video.

**SAMANTHA SAULD:**

Great. Thanks. The next question is, will you discuss the use of a second person to voice the audio description versus one voice?

**JAMES HERNDON:**

If I understand the question correctly, they are asking, using the word "you" in the description where "you are doing this. You see"-- or essentially, using the word "you," similar to what I mentioned earlier about using the word "we" in "we see." I think generally, it's best to avoid pronouns altogether unless they're absolutely essential. And that includes any kind of gendering of the point of view of who's seeing it. And it's really generally best to take that *Black Panther* approach of just complete sentences, but describing the action.

Because I think the more-- the more you talk about what is happening onscreen in the sense of you and we or even I and so on, the more you draw attention to the narration as a character experiencing the video

rather than immersing in the video and the service of providing immersion in the video, which is what audio description needs to do. So I would say if you absolutely have to-- and I can't think of an example of what a situation for it might be.

But I would say if you have to use a pronoun of some kind, "you" is certainly preferable to others. But I would strongly recommend scripting in such a way that you can avoid pronouns altogether, because I think they create a sense of distance that detracts from the sense of immersion in the presentation and reminds the person that they are someone hearing audio descriptions rather than somebody immersed in a video the same way that a sighted person is.

**SAMANTHA SAULD:**

Great. Thank you. The next question is, how would you suggest handling audio description for a video of a dance performance?

**JAMES HERNDON:**

That's a great question. People have done that a number of ways. And I would recommend checking out-- if you have Netflix, I would recommend looking at the audio descriptions for the most recent Thom Yorke solo performance called *Anima*. It's about 15 minutes long. It's directed by Paul Thomas Anderson. And it is essentially a 15-minute dance performance.

There are parts of it that are very abstract modern dance. And I think that-- I was interested to hear the audio description to see how they described it. They described the moves very literally, and they are accurate in that sense. But I found myself wondering how-- is there a way to accurately describe what's happening onscreen here, but also convey an emotional sense of what it feels like to see these bodies in motion in this way?

And again, I realize that is not the goal of audio description. It is, like I said, simply to describe what's happening onscreen and let the person who is hearing the audio description produce their own mental image of it. But there are other moments in the video where I think the descriptions are fantastic.

So I don't think I can answer that from the perspective of best practices because I have not seen enough audio descriptions of dance performances to be able to answer that with authority. But in terms of where things seem to be moving, I would highly recommend watching that Thom Yorke *Anima* performance on Netflix, because I think it represents a modern popular dance performance, essentially, that does have audio description and would give you a sense of how it's being done at the highest professional level. So check that out.

**SAMANTHA SAULD:**

The next question is, any general comments or advice for creating audio descriptions for educational videos versus a performance?

**JAMES HERNDON:**

So it really depends upon the type of educational video. When I create educational videos with screen captures, I always try to do so in the way that I described during that work order discussion for the video, the gentleman who used extended audio description where he paused and talked about simple work order and so on. I would-- I very carefully try to only show onscreen what I am talking about and try not to show more than I need to.

And I think if you can provide a broader context at the same time, then that's great. But it's very challenging to do. I would recommend checking out some PBS programming and see how they have done it. I don't think there would be-- if you ask somebody what their general guidelines are for it-- I'm suggesting PBS for you because I think the general guidelines you'll get for education are going to be some of the same DCMP-type guidelines that I mentioned before.

But I think in terms of giving you something you can get a handle on, I would recommend check out some of what you've seen and what you might find on PBS and see how they have done certain things. And maybe also check out-- let's see-- something to the effect of-- I bet I can find a couple other examples. Shoot me an email, [james@equalentry.com](mailto:james@equalentry.com).

And I'll respond to that with a couple of others. I can't think of what they

might be right now, but I'll send you a couple of videos and see if I can send-- I'll send you the URLs, and hopefully this can give you an idea of what some people have done.

**SAMANTHA SAULD:**

Thanks. So we only have time for one more question, although we have so many. So unfortunately, we're just going to go through one right now. And that question is, does extended audio description take away from the viewing experience?

**JAMES HERNDON:**

It's a tricky question to answer. If it is a movie, yes, I think it does, because the sense of suspense and timing, especially in a thriller or a comedy-- you're not providing an equivalent experience because timing is everything in those types of movies. And if there's constant pausing and saying, here's what's going on, you're totally jamming that mechanism.

Now, if this is a business presentation and you're working with PowerPoint slides, or if this is an instructional video, I don't think it's ruining the experience. I would still strive for standard just because of the time commitment. You're asking somebody to give more of their time to listen to an extended audio-described video as opposed to a standard length video that a sighted person can watch. But in terms of, is there something wrong with it, absolutely not. Extended audio description is perfectly, perfectly legit and very useful and totally fine.

**SAMANTHA SAULD:**

Great. Thank you. So thanks, everyone, for joining. And thank you to James for a great presentation. Keep an eye out for an email with a link to view the recording and slide deck. And I hope everyone has a great rest of the day.

**JAMES HERNDON:**

Thanks for joining, everyone.