JVIB

August 2004 • Volume 98 • Number 8

The Language System of Audio Description: An Investigation as a Discursive Process

Philip J. Piety

Abstract: This study investigated the language used in a selection of films containing audio description and developed a set of definitions that allow productions containing it to be more fully defined, measured, and compared. It also highlights some challenging questions related to audio description as a discursive practice and provides a basis for future study of this unique use of language.

Audio description, the practice of using language to give persons who are visually impaired (that is, those who are blind or have low vision) access to movies, television programs, and live events, has been practiced for more than 20 years. Also known as audiodescription, video description, and described video, this practice provides cultural, social, and educational benefits by allowing persons with visual impairments of all ages to experience a variety of cultural and educational texts that would otherwise be

inaccessible (American Foundation for the Blind, 1991). This article, which presents the first investigation of audio description as a language system, shows how audio description is both similar to and different from other popular forms of language use, including spoken and written discourse and human-mediated communication, such as sign language interpretation. It also shows it to be a fundamentally unique process that has much in common with other practices in which language is used to give persons who are visually impaired access to visual information, including audio textbooks and the Internet (Piety, 2003).

The first academic record of the concept behind audio description appeared in a 1975 master's thesis entitled *The Autobiography of Miss Jane Pitman: An All-audio Adaptation of the Teleplay for the Blind and Visually Handicapped* (Frazier, 1975), in which Frazier drew on some experimental audio productions from the early 1970s and theorized that information "could be inserted to increase listener comprehension" (p. iii). Today, audio description is not only a theoretical possibility but an established practice, and much of what Frazier envisioned is evidenced in current productions with audio description. In addition, although the producers of audio description today do not work with a formal set of standards, all those who are represented in this article draw from a common methodological history that began in live theater in 1982 (L. Goldberg, personal communication, fall 2002;

Packer & Kirchner, 1997; M. Pfanstiehl, personal communication, fall, 2002; J. Stovall, personal communication, October 24, 2002). Audio description has also been steadily expanding into new areas, including museum exhibits and live events, and internationally to many additional countries (C. Pfanstiehl, personal communication, fall 2002; J. Snyder, personal communication, summer 2002; Snyder, 2002). Furthermore, although research has shown that audio description is highly beneficial for both cultural inclusion and socialization (Packer, 1996) and can improve learning in educational contexts (Frazier & Coutino-Johnson, 1995, cited in Kirchner & Schmeidler, 2001; Katz & Turcott, 1993, cited in Kirchner & Schmeidler, 2001; Kirchner & Schmeidler, 2001), many important questions remain as to what type of production audio description creates; how, why, and which aspects of the practice are most effective; and how different approaches to audio description can be evaluated and tested.

This study is also part of a master's thesis (Piety, 2003) that investigated audio description as a process that depends on assistive technology but that is fundamentally a process of using human language. In this view, language is a facility that humans have evolved to utilize (Pinker, 1994) and that is formed by social practices (Halliday, 1978, 1985). Furthermore, it is important to note that persons with visual impairments, unlike many who are deaf, do not have a unique language. They are members of speech

communities that are made up mostly of people without significant visual impairments. The language that consumers of audio description use in daily life is thus shaped by the sighted world.

In pursuit of a greater understanding of the practices and effects of audio description, numerous viable paths for research exist. The range of options that are available to those who provide description in terms of which visual stimuli they will put into words, the words they will use and the way the words are assembled, and the interrelationships between the audio description inserted into gaps in the dialogue and other sound effects that could also be meaningful, as well as cognitive issues within the minds of consumers (the target area of this practice) are some of the perspectives that could generate meaningful research questions.

Furthermore, audio description operates under constraints that are imposed by the media that are used and the specific productions being described. For example, theatrical productions include variations in performance; television shows have generally short interdialogue opportunities for description to be placed; and films can include long stretches of time without dialogue in which information is contained in visual sequences, often with special effects, that, with their novelty and visual richness, present descriptive challenges. In addition, some genres, such as mysteries and musicals, impose additional challenges to the

describers. Despite the many issues related to research on this unique and important practice, there is no accepted theory or baseline set of definitions to support such research. This study was conceived as an initial step toward future research by developing a set of core structural and functional definitions derived from the analysis of the language used in several audio-described productions of a similar type.

Method

Research design

Because audio description is such a recently developed practice and has yet to generate a theoretical literature, the design of an inquiry that would best serve to characterize its nature required borrowing techniques that have been used in other analyses of language. This methodology draws on techniques that have been used in spoken discourse analysis, including the use of natural data (Chafe, 1994). The use of natural data has become an accepted method of understanding language through the ways it is used by highlighting actual productions of language, rather than hypothetical language presented out of context. The practice often involves the use of recorded episodes of the use of language that are then transcribed and analyzed. Transcription allows for the analysis of the language across time scales that are not practical in the moment-by-moment stream of language that listeners usually experience. Consistent with investigations of the use of

social language, elements were not randomly selected; rather, they were drawn from a continuous body of work and analyzed within the context of the other information (dialogue and visual and audio content) concurrent with the description (Lemke, 1998; Ochs, 1979; Schiffrin, 1994; Tannen, 1989).

Procedure

The unique nature of audio description guided the development of a new analytic procedure that was inspired by analyses of spoken discourse, but also took into account the fact that audio description involves language being inserted into another preexisting text. A key aspect of the analytic process is the ability to reconstruct, as much as the technology of print publishing allows, the information that a sequence of audio description is intended to convey. Naturally, different sequences of audio description are likely to be capable of carrying different types and amounts of information. Furthermore, important information for the consumer of audio description is carried through the dialogue in the original production and the other audio cues, such as environmental sounds and music, that appear to the consumer as an audio amalgam. But even though it contains these three elements—description, dialogue, and other audio cues—the audio portion of the described production alone would be inadequate for analysis because the describer also uses the visual portion of the original production to create the description. The analyst must keep in mind the two

different versions of the same production: the original version that the describer experiences and the modified one for the consumer.

To simplify the analysis and focus on understanding the different forms of the words of audio description, I used a simple transcription approach. First, I converted the original production (see Materials) to a digital video file. Next, I transcribed only the words of the description. The transcriptions were done for small continuous units of speech called "utterances" (described in the Results section) which were logged into a database and stored according to the time location in the original production and the length. This approach allowed me to analyze any utterance of audio description by its precise location in the original material. Then, to review any piece of audio description in the study, I looked at both the words of the description, as transcribed, and the original audio-video environment by using the time location stored in the database to access the digital video file.

The presentation of transcripts in this article includes the time signature, as well as a transcript of the original dialogue from the source material, as Transcript 1 illustrates. (Note that in all the transcripts, two dots indicate a slight pause.)

Transcript 1: From *Gladiator* (1:42:30)

1. Cassius: People of Rome . . on the fourth day of

Antioch . . we can celebrate . . the sixty-fourth day of the games

2. Describer: In the crowd, Maximus' servant Cicero looks around

3. Cassius: In his majestic charity, the Emperor has deigned this day to favor the people of Rome with a historical final match

4. Returning to the Colosseum after five years in retirement . . Caesar is pleased to bring you the only undefeated champion in Roman history . . the legendary Tigris the Gaul

5. Describer: The crowd stands as four galloping horses draw a chariot into the arena

6. Next to the driver, a gladiator salutes the crowd

7. He wears leather straps across his stocky chest and a metal helmet shaped like a tiger's head

8. On one of the underground ramps leading to an arena gate, Maximus swings a short sword

9. Proximo: He knows too well how to manipulate the mob

10. Maximus: Marcus Aurelius . . had a dream that was Rome, Proximo

The analytic process used the transcription database and original material to look at the range of different

types of descriptive options to develop a characterization of the language used in audio description. The transcript database can further be used for additional research on specific subvariations of audio description or the situations (scenes, things described, the locations of action) in which it is used, as well as to compare the way audio description is used with the studied productions with how it is used with other types of productions.

Methodological note: The features of transcription have been viewed as a methodologically significant factor (Ochs, 1979) because they constrain and privilege certain types of information. Because all productions in the corpus of audio description used a similar style of narrative speech that includes flat and consistent prosody, I used a simplified transcription system based on Tannen (1989). A common practice in the analysis of spoken discourse is to provide contextual information (for example, scene, movement, and gestures) that allows readers to better interpret the transcribed data. The transcription approach that I used does not use that level of detail because, since the source material is published, another researcher can recover that additional information directly by using the time sequence presented at the beginning of each transcript.

Materials

The body of language used in linguistic analyses is

often called a "corpus." Its selection is made on the basis of various criteria, including what is available (for ancient languages or languages with limited availability), certain conversational situations involving types of participants (for the analysis of spoken discourse), or certain types of written text (for other studies). The data for this study consisted of language used in 4 film productions containing over 23,000 words of audio description that were produced by three established describers, as shown in Box 1. (The selection of certain producers does not imply that they are the only established producers at this time or that some of the organizations that provide described productions that were not included in the study are less significant.)

This corpus represents a focused sample of audio description. It does not include theatrical productions, television shows, live events, or museum exhibits. All these other genres are important, and their inclusion in future studies would be desirable, especially since they may show variations on what this study found. The benefit of focusing on one genre, as this study did, is that within the limits of one study, the analyst is able to delve deeper into the genre and gain a more holistic view of a described production by reviewing each production from the beginning to the end along the same path that a consumer of audio description takes.

Results

This study produced both primary and secondary results. The primary results consist of definitions of the structural and functional components of audio description. These definitions provide common terms of reference and allow for a range of inquiry at a finer level of detail than was previously practical. They clearly indicate that the language used in audio description is not entirely the same as spoken or written language. Below the level of the word, at the morphological and phonological levels, this language does not seem distinctive. Above the subword level, however, it can be seen as having four distinctive structural components—insertions, utterances, representations, and words—which are defined in the following sections. The secondary results—some stylistic comparisons and statistics that provide additional preliminary insights into other dimensions of the practice—are presented at the end of this section.

Primary results

Insertions

The term *insertion* was developed for this study to indicate the essential element of audio description, the insertion of language into specific places in the production. The insertion is defined for audio description as a contiguous stretch of description that is uninterrupted by other significant audio content, such as dialogue. For example, Transcript 1 contains two

insertions (line 2 and lines 5–8). Insertions, which are usually bounded by dialogue, can be either short, lasting only a few seconds, or much longer, lasting several minutes. The mean length of insertions in this corpus was 11.09 seconds, with several insertions in one film (*Gladiator*) lasting over 4 minutes.

Utterances

The term *utterance* has been used in both philosophy and linguistics to mean the unit of language that is actually spoken (as opposed to the more common construct, the sentence, which is not always a component of spoken discourse). It is from this established definition that this study adopted and used the term *utterance* (see Harris, 1951, for discussion). Utterances can be arranged by the describer in any way to fill the time available in the insertion. They can be as long as the insertion itself or much shorter. Figure 1 shows that most are short, with almost 60% lasting between 1 and 2 seconds and almost 30% lasting between 3 and 4 seconds.

Utterances appear to the consumer, then, as short snapshots of language that describe some visible features. They are strung together to fill the space between dialogue.

Representations

While the definitions of utterances and insertions are

based on the physical properties of what is said (segments of description bounded by pauses), representations are semantic units that force this analysis to take a 90-degree turn from form into meaning and what the describer is attempting to communicate. Although the aspect of meaning entails complications of both theories involving language and the practical issues of analysts trying to assign meaning to these statements in a reliable and replicable way, issues of meaning are essential to understanding what audio description is communicating. Despite the theoretical complications of evaluating meaning from language, understanding audio description as a language system requires investigating what types of information are seen in practice. The term *representation* was inspired by a definition used by Halliday (1985) to describe functional grammar that includes processes, participants, and circumstances. Representations in this study were categorized into a taxonomy that includes seven types of information:

1. Appearance: The external appearance of a person, place, or thing.

2. Action: Something in motion or changing.

3. Position: The location of description or of characters.

4. Reading: Written or understood information that is literally read, summarized, or paraphrased.

5. Indexical: An indication of who is speaking or what is making some sound.

6. Viewpoint: Relating to text-level information and the viewer as viewer.

7. State: Not always visible information, but known to the describer and conveyed in response to visual information.

Each category, in turn, contains subcategories, many of which are described next. This taxonomy should be viewed as initial and open to expansion with further research (see Piety, 2003, for a more detailed discussion of the taxonomy and Turner, 1998, for a taxonomy derived for a different purpose and with different data).

Appearance. In some ways, appearance is the antecedent of all the other types of representation because all representations require an appearance of something in the original production to be realized in the description. Representations of appearance provide information about the direct visual properties of something in the production, including luminance, color, size, and shape. Appearance is usually realized through adjectives and the nouns they modify, as some examples (in italic) from Transcript 2 illustrate:

Transcript 2: From *The Gift of Acadia* (1:06)

1. Describer: A young woman in *a blue shirt and*

shorts lies on her back on a rocky ridge overlooking the sea below

2. She is reading a book

3. Narrator: The value of diversity and in diversity, harmony

4. Describer: A *small brown fawn* looks at us, twitching his left ear

5. A *black-headed loon* drifts by

6. A *thin black dragonfly* on a *green leaf* opens and closes its wings

7. Two *little orange-breasted baby robins* wiggle their heads

8. Under water, two *white-sided dolphins* swim smoothly side by side

9. On the *quiet* surface of the sea, two *black triangular dorsal fins* emerge, then curve back down under water

External appearance can also be conveyed with prepositional attachment, as Line 1 shows, and through adverbial phrases. It follows that if a consumer is interested in visual information—in what things look like normally or in certain situations—the information would often be provided through representations of appearance.

Action. Most utterances in this corpus are based on some form of action. Actions include gestures, movements, and activities, and they can act as the core representation around which other representations are clustered. Transcript 3 contains a typical set of action-oriented utterances (the actions are indicated in italics).

Transcript 3: From *Gladiator* (40:26)

1. Maximus: At least give me a clean death . . a soldier's death
2. Describer: One guard *moves* behind Maximus
3. Then *rests* his sword point on the back of his neck
4. Maximus *bows* his head as the guard *raises* the sword
5. Maximus *leaps* up and *butts* the guard off-balance, then *catches* the blade and *spears* him in the throat
6. *Spinning,* he *chases* the second guard, whose blade sticks in its scabbard
7. Maximus: The frost . . sometimes it makes the blade stick
8. Describer: With bound hands Maximus *slices* the sword across the guard's face
9. Nearby, two other praetorians *sit* on restless

horses

10. One *gallops* into the clearing, then *twists* in his saddle

11. A sword *flies* at him end over end

12. It *buries itself* in his back

13. Maximus *steps* out from the trees *glaring*

14. Maximus: Praetorian

The two insertions in Transcript 3 contain action in every utterance. This sample shows some of the different ways that action can be presented. Line 2 shows an action that relates the position of one person to another. Line 3 indicates an action with an object (sword point) and the location of the object. Line 4 is an example of simultaneous actions, and Line 5 contains a sequence of actions that are presented in a list form. Line 6 represents one action (spinning) as part of another action (although it is likely that the meaning intended is two actions in sequence). Lines 8 and 13 describe the manner of an action (with bound hands, glaring), and Lines 11 and 12 indicate an action in which the agent is inanimate.

Position. Another type of representation that is often associated with actions identifies the positions or locations of the information that is being described. Positional representations can act as action setters or

scene shifters (indicated in italics), as Transcript 4 shows:

Transcript 4: From *Gladiator* (1:42:30)

1. Cassius: People of Rome . . on the fourth day of Antioch . . we can celebrate . . the sixty-forth day of the games

2. Describer: *In the crowd,* Maximus' servant Cicero looks around

3. Cassius: In his majestic charity, the emperor has deigned this day to favor the people of Rome with an historical final match

4. Returning to the Colosseum after five years in retirement . . Caesar is pleased to bring you the only undefeated champion in Roman history . . the legendary Tigris the Gaul

5. Describer: The crowd stands as four galloping horses draw a chariot *into the arena*

6. *Next to the driver* a gladiator salutes the crowd

7. He wears leather straps across his stocky chest and a metal helmet shaped like a tiger's head

8. *On one of the underground ramps leading to an arena gate,* Maximus swings a short sword

9. Proximo: He knows too well how to manipulate the mob

10. Maximus: Marcus Aurelius . . had a dream that was Rome, Proximo

The information on location works as an action setter and relates characters to each other and the setting, as is shown in Lines 2, 5, and 6. Scene shifting occurs when a complex scene contains multiple perspectives that are alternatively presented to the audience. The scene shifts the viewpoint of the audience, but does not advance the action of the movie to a new scene. In some ways, it is similar to a flashback or dream sequence that allows for a suspension of action. Line 8 in Transcript 4 presents an example of this scene shifting; the main scene is in the Colosseum before a gladiator match, but attention has shifted to a quiet spot below the arena. Although all information on location seems important to viewers who cannot access the visual component of the text, these scene-shifting descriptions seem especially important because they allow viewers who cannot see the change in context to be able to comprehend the action, as a complex scene that is typical of the climaxes of modern films, unfolds.

Reading. Reading occurs when some language or recognizable symbols come on the screen and are literally read "as is" by the describer. Reading often comes at the beginning and end of movies when there are credits and titles. It also frequently appears throughout some movies in various forms. In Transcript 5, Line 4, a set of words are introduced to

indicate the location of the movie's action (a verb of introduction is underlined and read words are in italics).

Transcript 5: From *Gladiator* (47:49)

1. Juba: Better now? Clean, you see
2. Describer: Maximus lowers his lolling head back onto the wagon
3. Later the caravan approaches a congested desert town
4. Words appear, *Zucchabar*, *Roman Province*
5. A crude amphitheater dwarfs the surrounding red clay buildings

Transcripts 6 and 7 also show the describer reading signs that are part of the sets, rather than just screen text.

Transcript 6: From *L.A. Story* (3:38)

1. Describer: He rides in a park with other stationary bikers
2. A sign reads "*stationary bike riding park . . . no running*"

Transcript 7: From *L.A. Story* (19:00)

1. Large white-lettered signs reading " *now* " hang on the wall

2. Blue lights bathe the hip shoppers

In a manner similar to the way in which the speech of a person is reported or constructed in conversation (Tannen, 1989), the information that is being read is introduced through a verb of introduction, such as *read, reading, says, flashes,* or *appear,* that may indicate the manner in which the words are displayed on the screen as well as the content to come.

Indexical. Indexical or deictic information is information whose meaning can be determined only from the context (Levinson, 1983). In conversation, words like *here* and *now* provide direct meanings for speakers and hearers, but understanding the meanings requires an understanding of the place and time in which the conversation is situated. Although deictic content is common in conversation, I found only a few of these indexical representations (underlined) in the audio descriptions, revealing that audio description is not usually dependent on context. In Line 3 of Transcript 8, the describer indicates what object the character in Line 2 had just mentioned (in italics). In this case, to recover the meaning of this piece of description, the prior dialogue is required.

Transcript 8: From *A Star Is Born* (1:58)

1. Boy: Mush, that's what it was, just a lot of

mush . . there wasn't anybody killed in the whole thing

2. Father: Oh well, then, I'll stick to these . . these don't talk

3. Describer: *Looking at pictures*

4. Boy: That big cluck Norman Main was in the picture tonight

Transcript 9 shows another form of indexing in which the describer indicates who the next speaker is. In Line 2, the name Quintus (in italics) is said by the describer, and from accessing the video portion of the source text, it is clear that this statement identifies a character as the speaker.

Transcript 9: From *Gladiator* (6:10)

1. Describer: Across the battlefield at the edge of the forest, hundreds of barbarians wave their swords

2. *Quintus*

3. Quintus: Load the catapults

Viewpoint. Representations of viewpoint relate to what the viewer would perceive as affecting the entire visual field or text. They include scene changes or shifts, screen, and special effects. Scene changes are commonly indicated with the marker *now* or *later* (as indicated in italics in Transcripts 10–12):

Transcript 10: From *Gladiator* (1:15:29)

1. Describer: *Now* in the palace, a blurred face *comes into focus*

Transcript 11: From *Gladiator* (12:00)

1. Describer: Surrounded by flames, hundreds of men battle *in a blur of muted color*

Transcript 12 also shows a kind of scene shifter because at this point in the movie, a number of different screen effects appear in succession. *Next* indicates a change, but in this case not necessarily a formal change of scene.

Transcript 12: From *L.A. Story* (1:50)

1. Describer: *Next, a montage* of funky LA architecture

State. Description sometimes provides information that is not visually evident but is available through the describer's knowledge of the text. This information is presented by describing the identity or name of a character or place, providing relational information about entities that are visible, providing information about internal states (including emotions and intention), and specifying time.

Transcript 13 shows the naming of places. Although

locations are named in the movie (as Transcript 5 shows), in this example, the location, *Imperial Rome* , is not. The information on the location of the action and the buildings that are present was added by the describers (as indicated in italics).

Transcript 13: From *Gladiator* (58:42)

1. Describer: As they look at the stands that encircle them, the arena seems to spin like a carousel, blurring the cheering crowd

2. Now, *Imperial Rome* stretches far below

3. A flock of birds soars over the *Circus Maximus* and the *Colosseum*

Transcript 14 shows both the naming of a character and the relationship of the character to another character in the same utterance. This type of naming seemed to occur more with minor characters than with main ones. Transcript 14 also shows a common revelation of a shift in time, with *later*, which indicates that the shift is later in the script. Because movies can contain flashbacks, when a scene changes, viewers may not always know immediately that the scene has changed. The use of *later* identifies it as a change that is further in time.

Transcript 14: From *L.A. Story* (7:50)

1. Describer: *Later,* in his girlfriend Trudy's

apartment

Transcripts 15 and 16 present examples (in italics) of the description and evaluation of a character's internal state. Transcript 15 is an example of the state baldly described, while Transcript 16 shows a representation of state embedded in an action. This embedding may also be called an "evaluative description."

Transcript 15: From *L.A. Story* (90:38)

1. Describer: Slowly, "conditions clear" is spelled over the screen

2. *Content* , Harris smiles

3. In an aerial view, other digital road signs along the highways echo the same message

Transcript 16: From *L.A. Story* (25:36)

1. Describer: Now a deluge of mail shoots through the letter slot in Harris's front door

2. From the kitchen he *irritatedly* kicks the wastebasket underneath the opening, where it catches the streaming mail

A variation of the description of a character's internal state may be indicated by including "appears" preceding the evaluative phrase.

Words. The words used in audio description are an

extremely restricted set of the words that are used in spoken or written discourses. While most language use deals with information that is not present at the time of speaking, including past and future events and possible conditions (Chafe, 1994), the language used in audio description relates only to what is actually occurring on the screen at the time or close to the time that the words appear. So, unless they are part of something that is included in a representation of reading, there should be no words indicating conditions, past or future states, or any number of other valid language constructs that do not reflect the immediate reporting required for audio description.

Secondary results

In addition to the structural and functional description of audio description that was the focus of this study, the process of analyzing audio description from different providers, and transcribing a corpus with time-signature information, some secondary results are available, including the ability to compare different descriptive styles and to gain summary statistics about the content of the descriptions.

Comparison of styles

Box 2 illustrates some of the different ways to represent the same type of information. In the first option, each line of description is an action that is taken by an actor who is identified at the beginning of

the utterance, while in the second, there is a more varied approach. This is but one example of a comparison or analytic approach that is supported by the definitions in this study. See Piety (2003) for other examples.

Summary statistics

The database of transcripts can also be used to access summary information about the productions. As Table 1 shows, some productions include frequent short insertions with only a few utterances each, and some include comparatively longer insertions. In *Gladiator*, several insertions were over four minutes long, which amounts to one speaker, the voice of audio description, continuously occupying the audio stage for an extended period.

Discussion

The study was largely a structural and functional mapping of the systematic use of language in audio description. It provided evidence that audio description is a distinctive way to use language whose forms and functions are shaped by that use. This understanding opens up a range of questions that can be addressed within the framework of a more complete (although certainly inchoate) understanding of the practice. In this section, I revisit three issues that were raised earlier in this article to provide examples of the different directions that further analyses can take.

What type of production does audio description create?

When audio description is inserted into a production, a type of production that is different from the original one is created. It is different from the original production because of both the insertion of descriptive language and because for the typical consumer of audio description, the visual content is not fully used even though it is still present. Whether this new production qualifies as a derivative work in the legal sense is questionable. In a physical sense, a new product is created. But this new production can be viewed as a physical repackaging that serves to synchronize the descriptive content. The describer, like a sign language interpreter, is in a position to comment on but not change the original production. Although the describer has the ability to add any type of content, this article has also shown a series of representations that are inserted only between the dialogue of a character and the narration and that relate directly to the visual information in the film, rather than introduce new characters or action. Furthermore, even within the single genre of films, there is a range of different characteristics that the audio description can have.

Understanding the impact of audio description in terms of the amount of time it takes in a production and the way in which that time is distributed is only one form of analysis. It follows that a production with such large

amounts of time without dialogue could be challenging or frustrating to a consumer who does not have access to the visual content. Another form of analysis is to look at whether the essential information is contained within these descriptive sequences to understand if it is necessary to comprehension. In the two productions with long descriptive sequences (*Gladiator* and *L.A. Story*), essential scenes with significance for the plot were found to be conveyed entirely without dialogue, so the descriptive content provides important information with which to comprehend the text, suggesting that *the description comes where it is needed the most.*

How is audio description effective?

It is extremely difficult to understand with certainty the effect that any unit of language has within the mind of a listener. It is more so with audio description, which is transmitted to an audience that is anonymous to the describers. Despite these general issues related to understanding the effectiveness of language reception, there are some important factors that can be used to assemble a conceptual model. This model can be built on the foundation that persons with visual impairments, whether their condition is congenital or adventitious, are members of the same speech communities as are sighted persons. The fact that they are members of the same speech community means that consumers of audio descriptions will have used visually based word meanings in conversations with

sighted interlocutors and hence should require little or no special language consideration. This aspect is supported with research on congenitally blind children whose use of visual language was appropriate, although their inferences were restricted (Warren, 1994; Wyver, Markham, & Hlavacek, 2000).

Second, the representations of audio description are only one source of information for a consumer who is actively developing a personal representation of the production in his or her own mind (J. Stovall, personal communication, October 24, 2002). The consumer's cognitive process can be viewed as an active process in which representations of audio description, dialogue and other auditory cues, previous information established in the production, world knowledge, and inferences are continuously integrated to build and rebuild a representation of the scenes that are visually realized on the screen. In short, being a consumer of audio description can be viewed as an active cognitive process.

How can audio description be evaluated?

The qualitative issues regarding audio description were not specifically addressed in this study and present a large and significant challenge for both practitioners and researchers of the practice. A sign language interpreter takes units of language that are presented in one physical mode and presents them in another mode, often with a different language form (Lucas, 1989;

Valli & Lucas, 2001). In audio description, though, the describer is representing visual information that is usually a primary and parallel source of information by using language that is linear and a secondary source of information (Bateson, 1972). The describer's role, like that of the sign language interpreter (Metzger, 1999), can never be transparent. Critiquing any approach to audio description may appear easy, but it involves a number of significant factors. Because there is virtually an unlimited number of alternatives and because there are so many factors at play—from the informational requirements of the production to the acoustic and aesthetic issues to consider—determining which of the multitude of alternatives may be optimal requires considering a range of factors. Furthermore, because audio description is inserted within gaps in the dialogue that exist independently of the audio describer, describers are often constrained by the time that is available in choosing which representations to make and how to make them.

Every production of audio description includes two essential elements: the information that is being described and how that information is represented. Box 2 presents a sample of different styles of description. A detailed review of these or other descriptive approaches could consider many questions. Does a production that uses a variety of representational approaches increase the cognitive demand on the consumer or make the process more interesting and engaging or both? Do certain representational

approaches provide the consumer with or deprive him or her of important cognitive opportunities? Within different types of productions, should there be certain proportions of different types of representations; for example, in a historical drama, should there be types of representations of appearance that provide a certain overall impression of the period? Should the representations reflect the needs of certain types of consumers, such as those who are congenitally blind, who may be interested in facial expressions (L. Miller, personal communication, November 20, 2002) and other conversational cues? As a scene unfolds, should there be types of representations with priority that allow for the activation of schemas (Schank & Abelson, 1977) and expectations of interaction (Tannen, 1993)? As one looks at the larger cultural dimensions of the images, gestures, and symbols that are used in cultural productions, are there ways to translate the images that predominate in the visual culture (Barthes, 1957) into the world of audio description? Do the factors that influence good and appropriate description vary as the age and history of the consumer changes?

At present, there is neither a voice of the consumer nor a body of empirical research to indicate which approach may be better or if some other way of constructing representations would be more effective or efficient in helping consumers to build a relevant conceptual model of the production. The small group of researchers who have studied audio description have

yet to consider which styles create better cognitive opportunities for those who learn from, as well as are entertained by, this unique language form.

This article has been able to present only highlights of a study that presented the first description of the language used in audio description. It should not be viewed as conclusive or prescriptive because it comes at a time when the field of audio description is still new, and it is not intended to indicate what audio description should be across various genres but, rather, what audio description is today. The base set of definitions provided can be used to structure studies of human subjects, to compare different describers and the challenges that different genres present, and to assist in the development of guidelines and future refinements of this diverse language system. In addition, many of the issues and challenges that face audio description are faced by other practices, so-called visual assistive discourses (Piety, 2003), that attempt to make other texts, such as textbooks and hypermedia, accessible to persons with visual impairments. I believe in the fundamental importance of audio description for persons with visual impairments and that further study in this field is needed. Accordingly, I will make the entire corpus used in this research available to any other researcher who is interested in using it for similar purposes.

References

American Foundation for the Blind. (1991). *A picture is worth a thousand words for blind and visually impaired persons too: An introduction to audiodescription* [Brochure]. New York: AFB Press.

Barthes, R. (1957). *Mythologies.* London: Jonathan Cape.

Bateson, G. (1972). *Steps to an ecology of mind.* Chicago: University of Chicago Press.

Chafe, W. (1994). *Discourse, consciousness, and time.* Chicago: University of Chicago Press.

Franzoni, D. H. (Producer), & Scott, R. (Director). (2000). *Gladiator* [Motion picture]. Glendale and Hollywood, CA: DreamWorks Pictures and Universal Pictures. (Audio description by Descriptive Video Service.)

Frazier, G. M. (1975). *The autobiography of Miss Jane Pitman: An all-audio adaptation of the teleplay for the blind and visually handicapped.* Unpublished master's thesis, San Francisco State University.

Halliday, M. A. (1978). *Language as a social semiotic.* Baltimore, MD: University Park Press.

Halliday, M. A. (1985). *An introduction to functional grammar* . New York: Arnold.

Harris, Z. (1951). *Methods in structural linguistics.*

Chicago: University of Chicago Press.

Jackson, M. (Director). (1991). *L.A. story* [Motion picture]. Culver City, CA: Artisan Entertainment. (Audio description by Descriptive Video Service.)

Kirchner, C., & Schmeidler, E. (2001). Adding audio description: Does it make a difference? *Journal of Visual Impairment & Blindness, 95,* 197–212.

Lemke, J. (1998). Analyzing verbal data: Principles, methods, and problems. In B. Fraser & K. Tobin, (Eds.), *International handbook of science education* (pp. 1175–1189). Dordrecht, the Netherlands: Kluwer Academic.

Levinson, S. (1983). *Pragmatics.* Cambridge, England: Cambridge University Press.

Lucas, C. (1989). *The sociolinguistics of the deaf community.* New York: Academic Press.

Metzger, M. (1999). *Sign language interpreting: Deconstructing the myth of neutrality.* Washington, DC: Gallaudet University Press.

National Park Service (Producer/Director). (2000). *The gift of Acadia* [Videotape]. Washington, DC: National Park Service. (Audio description by Metropolitan Washington Ear.)

Ochs, E. (1979). Transcription as theory. In E. Ochs

& B. Schiefflin (Eds.), *Developmental pragmatics* (pp. 43–72). New York: Academic Press.

Packer, J. (1996). Video description in North America. In D. Berger (Ed.), *New technologies in the education of the visually handicapped: Proceedings of a conference held in France, June 10–11* (pp. 103–107). Montrouge, France: John Libbey Eurotext.

Packer, J., & Kirchner, C. (1997). *Who's watching? A profile of the blind and visually impaired audience for television and video.* New York: Information Center, American Foundation for the Blind.

Piety, P. (2003). *Audio description, a visual assistive discourse: An investigation into language used to provide the visually disabled access to information in electronic texts.* Unpublished master's thesis, Georgetown University, Washington, DC.

Pinker, S. (1994). *The language instinct: How the mind creates language.* New York: HarperCollins.

Schank, R. C., & Abelson, R. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.* Hillsdale, NJ: Lawrence Erlbaum.

Schiffrin, D. (1994). *Approaches to discourse.* Malden, MA: Blackwell.

Selznick, D. O. (Producer), & Wellman, W.

(Director). (1937). *A star is born* [Motion picture]. United States: Selznick International. (Audio description by Narrative Television Network.)

Snyder, J. (2002). *Fundamentals of audio description.* Tacoma Park, MD: Audio Description Associates.

Tannen, D. (1989). *Talking voices: Repetition dialogue and imagery in conversational discourse.* New York: Cambridge University Press.

Tannen, D. (1993). What's in a frame?: Surface evidence for underlying expectations. In D. Tannen (Ed.), *Framing in discourse* (pp. 14–54) . New York: Oxford University Press.

Turner, J. M. (1998). Some characteristics of audio description and the corresponding moving image . *Information Today, 35,* 108–117.

Valli, C., & Lucas, C. (2001). *Sociolinguistic variation in ASL* . Washington, DC: Gallaudet University Press.

Warren, D. H. (1994). *Blindness and children: An individual differences approach.* New York: Cambridge University Press.

Wyver, S. R., Markham, R., & Hlavacek, S. (2000). Inferences and word associations of children with visual impairments. *Journal of Visual Impairment & Blindness, 94,* 204–217.

Philip J. Piety, M.A., *doctoral student, School of Education, University of Michigan; mailing address: 429 Branch Drive, Silver Spring, MD 20901; e-mail: <ppiety@umich.edu>.*

Previous Article | Next Article | Table of Contents

Search JVIB | JVIB Policies | Contact JVIB | Subscriptions | JVIB Home

If you would like to give us feedback, please contact us at jvib@afb.net.