JVIB May 2006 • Volume 100 • Number 5

A Comparison of Alternative Narrative Approaches to Video Description for Animated Comedy

Deborah I. Fels, John Patrick Udo, Jonas E. Diamond, and Jeremy I. Diamond

Abstract: This study explored the feasibility of using a first-person narrative style for video description of an animated comedy, *Odd Job Jack*. It found that viewers who are blind find the first-person style more engaging, entertaining, and preferable but less trustworthy than the more conventional third-person description style.

The authors thank the Ontario Media Development Corporation for providing funding through the Pioneering Content program. They gratefully acknowledge the many contributors to this project, including Smiley Guy Studios, Don McKellar, Krystal Music Studios, and Ryerson University. John Stubbs provided much-needed advice, insight, and support from Audio-Vision Canada. Finally, the authors thank all the blind and sighted participants who gave their time and energy for the studies.

People who are blind or have low vision have only recently begun to enjoy greater access to television and video through Described Video Information (DVI). DVI provides spoken descriptions of the visual content (such as costumes, facial expressions, and sources of sound effects). It uses the existing audio pauses in the original audio track to insert spoken descriptions so that people who cannot see the visual elements are still able to understand the content (Canadian Radio and Television Commission, 2004; Centre for Accessible Media, 2003; Wall, 2002).

Early research on DVI demonstrated the value of DVI to people who are blind. Pettitt, Sharpe, and Cooper (1996) and Schmeidler and Kirchner (2001) compared participant-enjoyment and performance measures for content containing description to content that did not. The results indicated that the participants' experience with DVI was generally positive and enjoyable and that the participants preferred to watch television with DVI content than without it. However, people who were sighted were not included for comparison or for their reaction to DVI.

Peli and Fine (1996) examined the performance on multiple-choice questions of sighted viewers and those with low vision after watching segments of two documentaries. They found that the participants with low vision who watched the segments with DVI were able to answer significantly more questions correctly than were those who listened only to the audio portion (without DVI) of the same segments. They also found that the sighted viewers performed significantly better than did the low vision viewers using either DVI or the audio track-only version. However, Peli and Fine used conventional DVI and tested only two genres of programming--a nature documentary and a mystery. They did not assess the entertainment value of the DVI, only whether people could remember facts about each show. In our study, we focused on the impact of the style of DVI on engagement and entertainment factors, rather than on the ability of viewers to recall facts.

Television is a multifaceted medium that allows many types of stories to be presented in different ways. Narration is one important method that is used in television for storytelling. If a director uses a narrator to propel narrative events, she or he must decide between two main styles: a third-person (external or covert) style or a first-person (internal or overt) style (Chapman, 1978; Lothe, 2000). Wells (1998) suggested that narrated fictional animation and comics should use a first-person, rather than a third-person, narrator.

First-person narrators have a dual role; they exist within the plot as a character, and they narrate the story as their character in another place and time (Lothe, 2000). First-person narrators allow the audience to identify with the character and understand the inherent subjectivity of the narrator's version of the story. An advantage of this style is that it allows the viewer to interpret and criticize the narrator's version of events because the narrator is not given the same omniscient authority as in third-person narratives. The main disadvantage is that the audience may be receiving the story from a skewed or untrustworthy perspective and may be influenced by this one point of view.

Conventional DVI uses a third-person narrative style that is characterized by neutrality and noninterference in the story. An advantage of this point of view is that it is trustworthy, objective, and informative. The main disadvantage is that the narrator is not situated within the same fictional universe as the characters, meaning that he or she does not share the same emotional link as a first-person narrator would, which could detract from the overall feel of the story.

The standard practice for adding DVI is for the original production team to send the finished video to an outside service provider, who writes a separate

script, hires a narrator to record the script, and edits and intersperses the DVI within the original sound track. In creating the DVI script, the outside service provider follows some generally accepted practices. These practices include carefully observing the visual events, reducing what is seen into the most important information that can be conveyed in the time that is available, and objectively translating this information into appropriate words and phrases (Snyder, 2004).

In this article, we compare the reactions of users who are sighted and blind to conventional and first-person DVI forms using one example--an animated situation comedy called *Odd Job Jack*, produced by Smiley Guy Studios (2004) in Canada. Since sighted users and those who are blind often watch the same content together, it is important to understand the impact of DVI on both user groups.

Video content

Smiley Guy Studios produces the animated series *Odd Job Jack*, which airs on the Comedy Network (Smiley Guy Studios, 2004). *Odd Job Jack* is about the misadventures of a man, Jack Ryder, in temporary employment and Jack's quest to get a full-time job. In the episode, "American Wiener," used in this research, Jack works as a wiener-mobile driver and drives across the United States and Canada to enter a special box of wieners in an international wiener contest.

In this research, DVI production was carried out during the creation of *Odd Job Jack*. The script writer, director, and sound team were involved in creating and producing the DVI. The creative team decided that a first-person narrative approach would be the best way to convey the visual comedy that was presented on-screen.

The main character, Jack, provides a first-person oral account of the important visual constructs in the episode. He narrates in the past tense and includes all the subjective and emotional tones that would be expected of that character. For example, he speaks dialogue, such as "In Detroit, they stole all of our wieners!" in a disgusted tone. For a detailed account of the process and the scriptwriter's and director's perspectives, see Fels, Udo, Ting, Diamond, & Diamond (in press).

For comparison, one episode of *Odd Job Jack* was described by Audio-Vision Canada using its conventional National Broadcast Radio Service (NBRS) third-person narrative technique. The NBRS technique involves inserting accurate DVI as close to the visual equivalent in the story as possible.

Method

Seven subjects who were blind (three male and four female, who ranged in age from 30 to 39) and seven sighted individuals (three male and four female, who ranged in age from 20 to 29) provided self-reports for the study. Of the participants who were blind, two men made up one group, one man and two women comprised the second group, and two women participated individually to accommodate their scheduling preferences. Of the participants who were sighted, two men and one woman made up one group, one man and one woman were in the second group, and the remainder participated as individuals.

All the participants who were blind watched four or more hours of television per week and frequently rented or borrowed movies only with DVI. However, they usually watched television without description, since little of it is available. The participants who were blind stated that video description was important to their enjoyment and understanding of television shows, particularly those with less dialogue. They all found it important to watch television or movies independently of others. Of the sighted participants, four said that they watched one to three hours of television per week, and three watched more than four hours of television per week. All but one rented movies occasionally or often. In addition, only one sighted participant had been exposed to video description.

To maintain interest and manage the time constraints, we asked the participants to view four different and representative video clips with lengths of two to three minutes each. Specifically, the first clip introduces the episode by showing enticing scenes from later in the episode and the credits. The second clip shows a series of short scenes of Jack and his friends selling wieners in different cities. The third clip involves a scene in which Jack and his friends drive the wiener mobile through a large donut hole and the sexually oriented antics they experience. The fourth scene is a fight scene between Jack and another character.

The order of the clips remained the same for all the participants to ensure that the narrative structure flowed logically. Each clip was shown at least twice, once for each style of DVI. The order of viewing each clip was randomly determined and different for each participant or group of participants. Some blind and sighted participants asked to view the original nondescribed version of the *Odd Job Jack* clips to have a frame of reference regarding what was described versus what was contained in the original dialogue, particularly for the version with first-person narrative. Following the viewing

of each set of clips, a discussion ensued about the positive and negative aspects of each style of DVI, the participants' preferences, and the participants' understanding of the content.

All the clips were viewed on a Sony Trinitron 24-inch television that was connected to a Macintosh Powerbook G4 computer. The sound was played through a set of speakers that were connected to the computer and placed within the participants' reach. The volume levels could be adjusted to the participants' level of comfort.

DATA ANALYSIS

The data were collected using a questionnaire or interview format for the pre- and poststudy surveys. The discussions and dialogue that occurred during the viewing of the clips were recorded using a notetaker and videotape.

Thematic outcome measures for the video commentary were derived by two independent reviewers and then grouped into eight measures by agreement. These measures, shown in Box 1, were used in matrices of the type proposed by Miles and Huberman (1994). Nonparametric analyses were used in this comparison because of the relatively small number of measure counts and participants. A descriptive and qualitative presentation of the results is also provided.

To determine the interrater reliability of the video analysis procedure and the operational definitions, two evaluators were instructed in the video analysis procedure and trained in the coding categories. Each evaluator then independently rated the same two randomly selected video data sets, and their ratings correlated. The single-measures intraclass correlations for all categories were 0.6 or better, which indicates that interrater reliability was good. All the subsequent analyses were conducted by a single evaluator.

Results

In this article, the measures of greatest interest are the information-comfort, opinion, and preference measures, since these measures are related to the participants' subjective impressions of the quality of the information that was provided in each type of description. A Mann-Whitney nonparametric analysis was conducted between the blind and sighted groups for the total counts for all measures. A significant difference between these groups was found for the preference category, [$U(28) = 56.5$, $p < .05$]. No other significant difference was found for any category. Table 1 shows the total frequency counts and the mean and standard deviation (*SD*) for each participant group for all measures of interest.

For the participants who were blind, the positive-opinion measure had the greatest number of comments--75 (mean = 9.4, *SD* = 3.5)--and the negative-opinion category had the fewest--7 comments (mean = 0.9, *SD* = 1.0). Similarly, for the sighted participants, the greatest number of comments--67 (mean = 8.4, *SD* = 5.1)--occurred for the positive-opinion outcome measure, and the fewest comments--17 (mean = 2.1, *SD* = 1.5)--occurred for the negative information-comfort measure.

A Mann-Whitney test was also carried out that compared the conventional and first-person versions in all categories for each group of participants. A significant difference was found between the conventional and first-person versions for the negative measure of information comfort [$U(14) = 4.0$, $p < .05$] for the blind group. For the sighted group, there was a significant difference in preference between the first-person and conventional descriptions [$U(13) = 5.5$, $p < .05$]. The frequencies of each category are presented in Tables 2 and 3.

Mann-Whitney tests were also conducted between the blind and sighted groups for each style of description. For the conventional version, no significant differences in comments were found between the two groups. However, for the first-person version, there was a significant difference between blind and sighted groups in the preference category [$U(20) = 20$, $p < .05$]. Finally, a Kruskall-Wallis analysis was carried out to determine whether there were any significant differences that were due to the age differences between the groups. No significant differences were found.

DESCRIPTIVE ANALYSIS: CONVENTIONAL STYLE

As Table 2 shows, for the participants who were blind, the greatest number of comments for the conventional type of description occurred for the -positive-opinion category (total = 29 comments, mean = 7.3, *SD* = 2.5), followed by the positive information-comfort category (total = 23, mean = 5.8, *SD* = 2.4). For the sighted participants, the greatest number of comments also occurred in the positive-opinion category (total = 21, mean = 5.3, *SD* = 4.3), followed by the negative-opinion category (total = 16, mean = 4.0, *SD* = 2.2).

For the participants who were blind, the smallest number of comments (total = 2) occurred for the negative-opinion and the negative information-comfort (total = 4) categories. In addition, there were only five occasions when a blind participant expressed a preference for the conventional style of description. For the sighted participants, the fewest number of comments occurred in the preference category, for which there were only four occasions

when a participant expressed a preference for the conventional description.

DESCRIPTIVE ANALYSIS: FIRST-PERSON STYLE

For the blind group, the greatest number of comments for the first-person style of description occurred in the positive-opinion category (total = 46, mean = 11.5, *SD* = 3.1), followed by negative information-comfort category (total = 24, mean = 6.0, *SD* = 3.4). For the sighted group, the greatest number of comments also occurred in the positive-opinion category (total = 46, mean = 11. 5, *SD* = 4.1). For the blind group, the fewest comments (5) occurred in the negative-opinion category. For the sighted group, the fewest comments occurred in the negative-opinion category (7) and the negative information-comfort category (8).

There were 16 occasions (mean = 5.0, *SD* = 4.3) when the participants who were blind expressed a preference for the first-person description over the conventional version. In contrast, there were 28 occasions (mean = 7.0, *SD* = 0) when the sighted participants expressed a preference for the first-person description.

In addition, Kruskall-Wallace tests were conducted to determine whether there were any differences in responses between clips for each group and the groups combined. No significant differences were found between any of the four clips for both description styles.

Discussion

One of the most surprising results in this study is the quandary presented by the preference and negative information-comfort counts for both groups. Although the participants who were blind made a relatively high number of negative comments (24 negative information-comfort comments) regarding the trustworthiness of or their discomfort with the accuracy of information being provided in the first-person version of the video clips, they preferred that version three times as often as the conventional version. In addition, they had the highest number of positive-opinion comments and negative information-comfort levels for the first-person version of the show.

Two participants who were blind stated that the first-person version was like a "radio drama," and two suggested that is was like the character was "speaking to you." Another participant who was blind reported that "It was relaxing and fit with the style of the show." This participant also reported that it was difficult to switch between the comedic style of the content and the more "newscaster-like" conventional version of narration. One blind participant stated that she did not need to know all the information, but that

her priority in watching television was entertainment. Most participants who were blind stated that the first-person version was more entertaining and fun and reflected the nature of the show.

Two participants who were blind suggested that it is important for video description to be "audiocentric" (people thinking auditorily, rather than visually). They noted that the first-person narrative version provided a more audiocentered approach because it added elements of emotion and storytelling. Tabata (1995) indicated that the emotional content and subjectivity of first-person narratives is indicative of a more oral style of storytelling, perhaps fitting an audiocentric approach to DVI.

All but one blind participant, however, stated that they were uncertain or suspicious of the accuracy of the first-person version and suggested that this version was not providing all the "important" information or description and could not be trusted. In addition, three participants who were blind suggested that the first-person version provided only the main character's point of view and could influence the viewer. The uncertain and subjective nature of the first-person narrative form, discussed by Lothe (2000), seems to apply here as well. The purpose of the first-person narrative is to provide a more intimate and one-sided view of the story. This narrative is intended to provoke uncertainty and subjectivity.

It also appears that the discomfort that the participants who were blind experienced with the first-person narrative may have been due to their comfort with what was familiar (the third-person narrative), rather than to their doubts about the accuracy of the first-person narrative. The highest number of comments by the blind group regarding conventional description occurred in the positive-opinion category (26 comments, or about half the number of comments for the first-person version). The next-largest number of comments were in the positive information-comfort category (23 comments), indicating that the participants who were blind were probably comfortable with the style and accuracy of the information that was presented in the conventional version.

It was difficult to determine what the participants who were blind meant when they referred to the "importance" and "accuracy" of the description of content conveyed through DVI because the quantity and type of description are always decided by the writer-describer, not the viewer. These participants' comments revealed an interesting element of "trust" in the description, in which the narration is expected to be outside the content, not part of it. Further study is required to understand how people who are blind learn to trust the quality and quantity of information that is provided through

DVI, regardless of the style of presentation.

Comparing the conventional and first-person DVI scripts, we found that the vast majority of descriptions occurred at the same time and described the same events except for the introduction clip. The conventional version described the credits and the fact that the main character changed uniforms during the introduction (for example, "Don McKellar as Jack Ryder and" "Teress Morton as Betty Styles"). The first-person version provided background information for the show, such as the name of the main character and the purpose of the show (for instance, "My name is Jack Ryder, and I work for the Odd Jobs Temporary Employment Agency. My career counselor is the lovely Miss Betty Styles" (said with an adoring tone of voice). Most of the participants preferred the first-person introduction over the credit information because it was more engaging. Although this approach to describing credits may have been preferable, copyright and legal requirements of the film industry may prevent credits from being introduced this way. More research is required to determine how information on credits can be appropriately combined with the context and spirit of DVI.

Although there were no significant differences in the responses to the various clips, there were some noticeable trends. Clips 2 and 4 in the first-person version had the most positive responses. For the participants who were blind, the first-person version of Clip 4 had the highest number of positive comments and preference statements of all the test clips for both types of description (22 positive comments and 5 preference statements), followed by Clip 2 of that version (16 positive comments and 8 preference statements). Clip 4 was a fight scene, in which strong emotions and excitement were expressed by the characters in the first-person version. The participants who were blind made such statements as "The first-person description is better than the conventional by a long shot" and "This is really funny, and I am not rushed" for Clip 4. They seemed to enjoy having the emotional and exciting elements for this particular clip portrayed in the video description. However, only one clip elicited this much response from the participants. Additional research is required to understand the desire of the participants who were blind for emotional elements to be introduced into video description.

Some interesting data appeared for the negative results, although the number of comments was relatively small. The majority of negative comments occurred for the first-person version of each clip. For example, there are 10 negative comments for the first-person version of Clip 3, but none for the conventional version for the blind participants. In addition, the number of negative comments in general for the conventional version of all clips was considerably lower than the negative comments for the first-person version.

This discrepancy may be the result of two conflicting factors: (1) the level of comfort, trust, and experience that the participants who were blind had with conventional video description and (2) the desire for fun and entertaining content that seemed more apparent in the first-person version.

The conventional version is the typical version that viewers who are blind have had with description; thus, the blind participants may not have felt compelled to make positive or negative comments because they are used to that version. The first-person version offered a new approach, and therefore the participants who were blind were perhaps more willing to make comparisons and provide comments because it was different from their typical viewing experience.

The results of the study point to a model of DVI that uses more than one type of narrative style and indicate that one style may not suit all types of content or all users. In fact, the creators of television content successfully use different narrative styles. It would seem logical that different DVI styles may also be possible and successful.

Limitations

Although the study provided some interesting results and trends, some important limitations need to be noted. The first limitation is that the number of participants was low--seven per group. Additional studies with a greater number of participants would allow trends in the data to appear with more certainty.

Another limitation is that only a few participants in the sighted or blind groups fit the age of the target market (19-34-year-old men) for the particular content used in the study. Hence, some of the negative or positive comments could have been due to the participants' reaction to the content itself, rather than to the quality and quantity of description. Two of the older participants who were blind stated that they did not care for the sexually explicit content of one particular clip, but that they still believed that they could evaluate the description. Recruiting a large number of participants in the appropriate age group would be important (but likely difficult) for future work.

Six participants suggested that the first-person narrative approach might not work in all genres, but that it seemed particularly suited to comedy and drama. As Pettitt et al. (1996) reported, mystery and murder dramas are the most difficult for viewers who are blind to watch and understand. In our study, only one subgenre, an adult animated comedy, was used. Different genres (such as mystery, drama, comedy, and documentary), types of content, and production companies need to be engaged to examine different

narrative styles (first person and third person) of DVI and to explore more fully the impact on the entertainment value, trustworthiness, and appropriateness of these differences for various groups of viewers who are blind.

Only four short clips were used for the study to provide some initial data on whether using alternative narrative in video description is acceptable to users. The participants did not watch full-length episodes with either type of description and were not able to gain experience over a longer, multiweek period. Longitudinal studies are required to understand the impact of alternative forms of DVI on blind and sighted participants and to determine whether there is a significant novelty effect.

This study provided the initial exposure to first-person video description for all participants, providing a bias in the study toward conventional video description. A more balanced study would also include participants who had little or no experience with any form of video description.

Finally, having the production company involved in this research allowed us to avoid potential issues arising from copyright and DVI and does not recognize the current reality of the outsourcing of DVI production. However, another potential research opportunity is to explore the industry's receptiveness to being more involved in DVI so as to retain more control over DVI processes and copyright.

Conclusion

As was illustrated by the participants' evaluations, the first-person-description approach shows promise. The participants liked the first-person version of *Odd Job Jack*, especially after their first exposure to it. They suggested that it was fun to watch and made their experience with the show interesting, even though it was less trustworthy. This reaction occurred even though the participants who were blind were more comfortable and experienced with the conventional third-person narrative approach to DVI. The results seem to indicate that it is possible to consider alternative forms of DVI, rather than the single form that is currently used.

We also found that the entertainment value, as well as the trustworthiness, of the DVI information was important to viewers who are blind. However, it appears that there may be some trade-offs between the entertainment value and the perceived accuracy and completeness of the description. Future research needs to explore how well all forms of DVI match viewers' expectations and needs. In addition, further studies using different genres and longer amounts of content over multiple sessions are required to gain a better

understanding of the long-term reactions of viewers who are blind and how well various forms of DVI fit different genres. The audio track of *Odd Job Jack* that was described and an audience-reaction survey can be viewed at the web site <www.theclt.org/~oddjobjack/jackonweb/index.htm>.

References

Canadian Radio and Television Commission. (2004). *Broadcasting public notice CRTC 2004-2* [Online]. Available: http://www.crtc.gc.ca/archive/ENG/Notices/2004/pb2004-2.htm

Centre for Accessible Media. (2003). *Rich media accessibility* [Online]. Available: http://ncam.wgbh.org/richmedia/strategies/XAD.php

Chapman, S. B. (1978). *Story and discourse: Narrative structure in fiction and film*. Ithaca, New York: Cornell University Press.

Fels, D. I., Udo, J. P., Ting, P., Diamond, J., & Diamond, J. D. (in press). *Odd Job Jack* described--A first-person narrative approach to described video. *Journal of Universal Access in the Information Society*.

Lothe, J. (2000). *Narrative in fiction and film*. Oxford, England: Oxford University Press.

Miles, M. B., & Huberman, M. (1994). *Qualitative data analysis.* Thousand Oaks, CA: Sage.

Peli, E., & Fine, E. M. (1996). Evaluating visual information provided by audio description. *Journal of Visual Impairment & Blindness, 90,* 378-385.

Pettitt, B., Sharpe, K., & Cooper, S. (1996). AUDETEL: Enhancing television for visually impaired people. *British Journal of Visual Impairment, 14, 48-52.*

Schmeidler, E., & Kirchner, C. (2001). Adding audio description: Does it make a difference. *Journal of Visual Impairment & Blindness, 95,* 198-212.

Smiley Guy Studios. (2004). *Odd Job Jack*--Episode 13, "The American weiner." Toronto: Author.

Snyder, J. (2004). *Audio description: The visual made verbal* [Online]. Available: http://www.audiodescribe.com/article1.html

Tabata, T. (1995). Narrative style and the frequencies of very common

words: A corpus-based approach to Dickens' first person and third person narratives. *English Corpus Studies, 2*, 91-109.

Wall, J. S. (2002). TV broadcasting for all. Paper presented at the Open Workshop on Standarization in the Field of Subtitling. Seville, Spain.

Wells, P. (1998). *Understanding animation.* New York: Routledge.

Deborah I. Fels, Ph.D., *associate professor, School of Information Technology Management, Ryerson University, 350 Victoria Street, Toronto, ON, M5B 2K3, Canada; e mail: <dfels@ryerson.ca>.* ***John Patrick Udo, M.A.,*** *research assistant, School of Information Technology Management, Ryerson University; e mail: <judo@ryerson.ca>.* ***Jonas E. Diamond, M.B.A.,*** *chief executive officer, Smiley Guy Studios, 444 Bathurst Street, Studio 2, Toronto, ON, M5T 2S6, Canada; e mail: <jonas@smileyguy.com>.* ***Jeremy I. Diamond, B.A.,*** *writer, Smiley Guy Studios; e mail: <jeremy@smileyguy.com>.*

⠠⠃⠗⠇ Download braille-ready file

Download ASCII text file

Previous Article | **Next Article** | **Table of Contents**
