
LiveDescribe: Can Amateur Describers Create High-Quality Audio Description?

Carmen J. Branje and Deborah I. Fels

Structured Abstract: *Introduction:* The study presented here evaluated the usability of the audio description software LiveDescribe and explored the acceptance rates of audio description created by amateur describers who used LiveDescribe to facilitate the creation of their descriptions. *Methods:* Twelve amateur describers with little or no previous experience with audio description used the software LiveDescribe to describe a single episode of a 20-minute comedy show. Seventy-five reviewers who were blind, had low vision, or were sighted then rated the descriptions using a number of criteria, including overall quality and entertainment value. *Results:* LiveDescribe was found to be easy to use and useful. Three of the 12 describers produced descriptions that were rated as of good overall quality, 6 produced descriptions that were rated as of medium quality, and 3 produced descriptions that were rated as of poor quality. *Discussion:* These findings indicate that amateur description is feasible even with minimal training in either description itself or LiveDescribe. Audiences' preferences for description seem to be based on various characteristics of describers, such as the describers' vernacular and tone of voice and the length and timing of the descriptions. *Implications for practitioners:* If amateur description is indeed feasible, the quantity of audio descriptions that are available to the general public could be increased significantly. A great deal of informal description is already created by families and friends of individuals who are visually impaired through the "whisper method." If this description process could be captured and formalized through a tool such as LiveDescribe and shared through the Internet, many more descriptions could be made available.

Audio description, also known as video description or described video, is a process that has been developed to provide access to television, film, and theater content for viewers with visual impairments (that is, those who are blind or have low

vision). It provides a spoken description of visual content, including action sequences, costumes, and facial expressions (Fels, Udo, Diamond, & Diamond, 2006). Descriptions of these elements are inserted via a second audio channel

between the dialogue, so that there is no overlap between the description and the characters' voices in the show. The timing and precision of the description is critical because the available spaces in which to insert descriptions are often short in duration (fewer than five seconds) and infrequent. As a result, not all the important or relevant visual information can be described.

The formal process of description, in which procedures and processes are explicitly defined and published, appeared in the 1970s (Snyder, 2004). However, informal description, in which family members and friends describe the visual world to individuals who are visually impaired (Schmeidler & Kirchner, 2001), has occurred for many years. Formal descriptions are created using a process involving identifying spaces between the elements of the dialogue where descriptions can be inserted and then writing, recording, and editing a description script to fit within these spaces.

Informal description usually occurs live and without much preparation, writing or scripting, or specialized recording equipment. In addition, all description

tasks, including the composition and delivery, are performed by one person. In this article, paid professionals who work as describers and follow the formal processes and procedures prescribed by a "professional" community are differentiated from unpaid amateurs who are family members, teachers, or friends of people who are visually impaired, and who carry out description as part of their day-to-day interactions with these individuals.

There are a number of issues that impede the development of formal and informal audio description techniques and technologies. The Independent Television Commission (2000, p. 12) reported that it takes one describer one week of work to produce about two hours of described programming. This increase in production time for creating and deploying a piece of content can become a barrier to the creation of audio description.

Another but related issue is cost. Reports of production costs for audio description have ranged from \$1,500 per hour of content (personal communication with R. Trimbee, National Broadcast Reading Service, May 12, 2008) to \$4,000 per hour of content (Clark, 2007), meaning that a full-length movie could cost as much as \$10,000 to describe. Since producers of content are under constant financial pressure, the cost of audio description represents a large barrier in the creation of a high volume of described content. One method of increasing the efficiency of creating descriptions is to introduce more automatic processes, such as using programs that automate a portion of the process for the describer (see Branje, Marshall, Tyndall, & Fels, 2006; Gagnon, Foucher, Laliberte, Lalonde, & Beaulieu, 2006).

Editor's Note: The first author of this article is the developer of LiveDescribe, which is discussed in this article. Readers should note that the *Journal of Visual Impairment & Blindness* does not endorse products that are mentioned in the journal. *Authors' Note:* Funding for this research was generously provided by Canadian Heritage, the CCOP program, the Natural Sciences and Engineering Research Council, and the GRAND National Centre for Excellence. Many thanks to Lisa Copeland and Ryan Farnum for their technical and editing skills. Finally, we gratefully acknowledge all the people who participated in the Phase 1 and 2 studies.

The emerging wiki phenomenon has opened the door to new ways of computer-supported human-to-human collaboration. Web 2.0 is a concept in which users create, edit, and share multimedia content online (Bleicher, 2006; Tapscott & Williams, 2006). Using tools such as LiveDescribe and Web 2.0 concepts, it may be possible to capture, preserve, and make audio description public. This article presents the results of a study that examined the feasibility of using amateurs to create audio descriptions using LiveDescribe while maintaining an acceptable level of quality. The study was the first step toward determining the possibility of an audio description wiki.

LiveDescribe

LiveDescribe is an open-source software application, developed by the first author, that was designed to facilitate the creation of audio description for digital video content by amateur describers. The interface allows users to record, insert, edit, and manipulate their descriptions on a time line. Furthermore, LiveDescribe provides a graphical representation of the audio track that automatically identifies and highlights the nondialogue spaces that are available for inserting descriptions (see Branje et al., 2006, for a detailed description of the user interface and the discrimination algorithms).

The study

RESEARCH QUESTIONS

The primary objective of the study was to determine the feasibility of amateur describers creating usable audio description using LiveDescribe. Another objective was to begin to understand the characteristics and attributes of successful and un-

successful amateur describers. This knowledge will help guide the future creation of tutorials, guidelines, or other documents that are designed to train new and existing amateur describers as well as the future design of audio description software or hardware tools.

METHOD

Two phases of an exploratory study were designed to begin to answer the research objectives. In Phase 1, amateur describers created descriptions that were to be reviewed by viewers who were blind, had low vision, or were sighted in Phase 2. Such factors as overall quality, vocabulary quality, and audio quality were the focus of the Phase 2 evaluation. The research was approved by the Ryerson University Research Ethics Board.

Phase 1

During the first phase, 12 participants, 5 women and 7 men, created audio descriptions for an entire 20-minute episode of *The Daily Show* with Jon Stewart. *The Daily Show* is a 20-minute daily mock news program during which a comedian delivers funny commentary on current events, using video and audio aids. Of the 12 participants, 9 were aged 18–29, and 1 each were in the 29–39, 39–49, and 59–69 categories. Four had high school diplomas, 2 had college diplomas, 4 had university degrees, and 2 had graduate degrees. Only one participant had created audio description before, but not regularly. Four participants had heard of audio description before they entered the study, and the remainder had no experience with audio description.

On their arrival at the test facility, the participants were given a 17-question (15

forced-choice Likert scale questions and 2 open-ended questions) prestudy questionnaire, which was designed to capture demographic information, such as age, gender, and computer experience and any previous experience with audio description. They were also asked about their familiarity with and level of appreciation for *The Daily Show*. Next, training was provided for audio description, specifically, how to use LiveDescribe to create and record descriptions. This training lasted about 15 minutes and consisted of informing the participants of some of the conventions of description, such as attempting to keep descriptions within pauses of dialogue and avoiding describing visual elements that also have an audio cue (such as a telephone ringing). Once the training tasks were completed, the participants were asked to complete the task of creating a description of one episode of *The Daily Show*. Each participant was responsible for all the description tasks, from the creation or composition of the description to its delivery, emulating how amateur description is normally carried out. A video recording of the actions performed during this task was collected for each participant.

After the participants completed the description task, they filled out a poststudy questionnaire (12 forced-choice Likert scale questions and 3 open ended questions) that was designed to evaluate the usability of LiveDescribe and capture the participants' experience of describing the show. There were 6 questions on the ease of use and ease of learning of LiveDescribe and 9 questions on to the ease of performing the various description tasks, such as creating, timing, and editing the descriptions.

Phase 2

For each describer from Phase 1, an audio-only version of *The Daily Show*, including the created descriptions, was generated and divided into five separate clips of about three to five minutes long and placed online. Each clip was divided at an appropriate point, such as during a commercial break or at the beginning of a new sketch or segment of the show.

Seventy-five participants (6 of whom were sighted, 25 of whom had low vision, and 44 of whom were blind) reviewed the online descriptions and rated various aspects of the descriptions, such as overall quality, vocabulary level, and style of delivery. They were asked to complete a total of six surveys. The first consisted of 11 forced-choice Likert scale questions and 2 open-ended questions that collected demographic information, such as vision status and age. This questionnaire also asked how many hours of television the participants watched each week, how often they went to the movies, and their experience with audio description.

The second set of surveys consisted of the same survey administered five times, one for each of five clips with five different describers randomly assigned. This survey consisted of 11 forced-choice Likert scale questions and 2 open-ended questions that asked about the overall quality, vocabulary level, audio quality, and style of the description clip that had just been reviewed. The open-ended question asked the participants to provide general positive and negative comments about the description.

DATA ANALYSIS

Data from the Phase 1 questionnaire were collected from 12 Phase 1 participants.

Likert scale responses from the questionnaire were coded, with 5 as the highest positive value (very easy or very useful) and 1 as the lowest negative value (very difficult or very useless). Data from the Phase 2 questionnaire were collected from 75 participants who reviewed 287 clips. Although some participants did not finish the entire survey, the clips they reviewed were used in the final analysis.

For the Phase 2 data analysis, a Bonferonni adjusted significance level of .01 was used to safeguard against the increased probability of Type I errors resulting from multiple tests of statistical significance on the same data set. Although the variables are reported with a $p < .05$ significance level to illustrate trends that may warrant further study, variables that passed the Bonferonni adjusted significance level of $p < .01$ are highlighted with an asterisk.

Results

PHASE 1

Usability factors

Of the 12 participants, 11 reported that learning to use LiveDescribe was easy, and the remaining participant reported that it was not easy or difficult. All 12 participants reported that using the computer was easy. Nine participants reported that learning to use software in general was easy, and 3 reported that it was neither easy nor difficult or was difficult. Using the LiveDescribe time line, users can change or edit the “description boundaries” or segments of the video that the system has automatically determined to be free of dialogue. Eight of the 12 participants thought that editing the description boundaries was easy, 2 thought

it was difficult, and 2 thought it was neither easy nor difficult.

Six participants reported that finding space to insert a description was easy or very easy, and six reported that it was difficult. Seven participants reported that navigating through the video was easy, two were neutral, and three reported that it was difficult. Eight participants reported that writing descriptions into LiveDescribe was easy, two were neutral, and two said it was difficult. Eight participants reported that recording description was easy, one was neutral, and three reported that it was difficult. Eight participants reported that understanding the graphs was easy, and four reported it was difficult.

The participants rated the usefulness of the recording functions, graphical time line, automatic space detection, timer for recording, and the writing functionality. For each category, at least 10 of 12 participants reported them as useful for their creation and recording of audio descriptions. Five participants reported that the running list of descriptions was useful, 6 were neutral, and 1 reported that it was useless.

Description factors

Four of the 12 participants thought that learning to describe was easy, 2 were neutral, and 6 thought it was either hard or very hard. Three participants thought that deciding what aspects of the show to describe was easy, 1 was neutral, and 8 said that it was either difficult or very difficult. Two participants reported that choosing the words to use for description was easy, 1 was neutral, and 9 thought it was difficult. Ten participants reported that understanding the video was easy, and 2 were neutral. All the describers in Phase 1 of the study were able to

Table 1

Summation of significant variables of when grouped by the describers. These variables were found to have significant differences among the describers.

Variable	df	F	Sig.	Eta ²
Overall quality	11	4.38	0.00*	0.15
More humor	11	3.76	0.00*	0.13
More information	11	1.89	0.04	0.07
More entertainment	11	2.11	0.02	0.08
Vocabulary level	11	4.01	0.00*	0.14
Style of delivery	11	2.93	0.00*	0.11
Listen to more	11	3.38	0.00*	0.12
Compared to professional	11	3.00	0.00*	0.11

* $p < .05$.

complete the task of description even though they all had little or no exposure to audio description prior to their participation.

PHASE 2

Four participants reported that they never watched television, 25 reported that they watched television 1–5 hours per week, 20 reported that they watched 6–10 hours per week, 14 reported that they watched 11–15 hours per week, and 12 reported that they watched more than 15 hours per week. Twelve participants reported that they never went to the movies, 30 reported that they went once a year, 28 reported that they went once a month, 4 reported that they went once a week, and 1 reported going more than once a week.

The participants were also asked how often they used audio description when they watched television or went to the movies. Twenty-two participants reported that they never used it when watching television, 18 reported that they usually watched television without it, 28 reported they sometimes used it when watching television, 8 reported that they usually used it when watching television, and only 1 reported always using it when

watching television. One also reported never watching television. In addition, 30 participants reported that they never used audio description when attending the movies, 10 reported that they usually did not use it, 14 reported that they sometimes used it, 3 reported that they usually used it, and 5 reported that they always used it. Thirteen reported that they did not go to the movies.

A one-way analysis of variance (ANOVA) was conducted on the Phase 2 survey data to assess differences in the opinions of the reviewers in Phase 2 and the describers in Phase 1. Describer D12 received the fewest number of reviews (19), and Describer D8 received the most reviews (34), while most describers received between 23 and 27 reviews. A significant difference among the describers for all the variables except audio quality was found. Table 1 shows the results of the ANOVA for all significant variables to a level of $p < .05$.

Overall quality was used as the primary measure of the quality of the descriptions throughout the data analysis because it provided a summary judgment. If we consider good describers to be rated above a mean of 3.5 (good and very good), poor

Table 2**Summation of the results among the vision status groups for all significant variables that were found to have significant differences when grouped by the vision status of the reviewer.**

Variables	df	F	Sig.	Eta ²
Overall quality	2	3.081	0.047	0.06
Added humor	2	5.660	0.004*	0.02
Added entertainment	2	4.820	0.009*	0.02
Audio quality	2	6.377	0.002*	0.03
Listen to more	2	4.033	0.019	0.04
Compared to professional	2	6.928	0.001*	0.08

* $p < .05$.

describers rated as below a mean of 2.5, and neither good nor poor (neutral) reviewers as having a mean between 2.5 and 3.5, three describers (D12, D10, and D9) were rated as good, 3 describers (D11, D8, and D4) were rated as poor or least preferred, and 6 describers were grouped in the center of the preference scale.

Further illustrating the effect of the describer on the rating of quality are the results from the compared-to-professional question. For this question, a rating of 5 meant that the participants thought the description was much better than professional description, and a rating of 1 meant that they thought it was much worse than professional description. Similar to the results for overall quality, Describers D12 ($M = 2.63$, $SD = 1.30$), D10 ($M = 2.25$, $SD = 1.07$), and Describer D9 ($M = 2.43$, $SD = 1.16$) achieved the highest rating for compared to professional. This result suggests that the level of quality of the top describers was rated as only somewhat worse than professionally produced description.

A Tukey HSD post hoc analysis showed that the differences were primarily between the worst and best describers. There is significance between the mean values for overall quality between De-

scriber D12 (the highest-rated describer) and Describer D11 ($p = .018$), D8 ($p = .011$) and Describer D4 ($p = .045$) (the lowest-rated describers), and between Describer D10 (one of the highest-rated describers) and Describer D11 ($p = .031$) (one of the lowest-rated describers).

A one-way ANOVA showed that significant differences were reported among the five clips for all reviewer variables except compared to professional (see Table 2 for the results of the ANOVA). Post hoc tests revealed a significant difference between Clip 2 and Clips 1, 4, and 5. Clip 2 had a significantly lower review for overall quality ($M = 2.31$, $SD = 1.11$) than did Clips 1 ($M = 3.06$, $SD = 1.14$), 3 ($M = 2.62$, $SD = 1.48$), 4 ($M = 3.02$, $SD = 1.45$), and 5 ($M = 3.26$, $SD = 1.17$). Clip 2 had the lowest mean value, while Clip 5 had the highest. This trend may suggest a chronological effect on the perception of quality. In addition, the describers did not produce as much description for Clip 2 as with the other clips, which may have affected the viewers' perception of quality.

There was a significant correlation between ratings of overall quality and vocabulary level, $r(278) = 0.693$, $p < .05$, and style, $r(278) = 0.764$, $p < .05$). There is a pattern similar to that for ratings of

overall quality, with D9, D10, and D12 reporting the highest average ratings for vocabulary level and style; D4, D8, and D11 reporting the lowest ratings; and the remaining describers with neutral ratings (neither good or poor).

Discussion

PHASE 1

Usability

Although most of the describer participants found learning to use LiveDescribe to be easy, they struggled with the actual task of describing what they saw. The majority of participants indicated that most of the LiveDescribe functions were useful, indicating that the functional needs of these amateur describers were likely facilitated by LiveDescribe. The one function that seemed to be less useful was the running list of descriptions, perhaps because not enough descriptions were produced for the 20-minute show, so that it was not necessary to look through a list of them or because the information in each description was self-contained enough that there was no need to look through them for review or reediting. Further research on the usefulness of this function for more complex or lengthy content is warranted before it can be modified or removed.

The participants reported that finding space to insert a description and navigating through the video were more difficult to use than were other aspects of LiveDescribe. Although LiveDescribe provides an indication of the location of periods of nondialogue, the algorithms are only about 85% accurate, which may make it difficult to find the exact point for the insertion of a description. Work is

currently under way to incorporate new speech-discrimination algorithms to improve the accuracy of the system.

Navigation through the video in LiveDescribe is achieved by clicking and dragging a small pointer that also serves as a position indicator. Because of the small size of the pointer, a high degree of fine motor control is likely required to select and move the pointer accurately. One possible solution to address these issues is to create a “snap to” feature, where the mouse cursor automatically snaps to the pointer if it comes within close proximity.

Description task

Description appears to be a difficult task, since the majority of the participants in Phase 1 reported that many aspects of description, such as choosing which aspects of a show to describe or which words to use, were difficult to accomplish. This finding is congruent with previous work that suggested that creating descriptions is a cognitively demanding task that can be difficult to carry out (Fels et al., 2006). However, all the participants were able to complete the descriptions for the show even with a short period of training. We expect that with practice, the description task would become less difficult for an amateur describer but still remain a relatively demanding task.

PHASE 2

Most participants had some exposure to professional description only through television or movies, which could suggest that there is a need for more description and tools that can facilitate its creation by professionals and amateurs. Several other findings arose from the analysis of the

Phase 2 data. First, it seems that in the study, the describers could be categorized as “good,” “medium,” or “weak” on the basis of the overall quality of their descriptions. This finding begins to answer one of the main research objectives regarding the feasibility of the process of description by amateur describers. Indeed, at least 3 of the 12 describers from Phase 1 were able to create descriptions that the majority of the audience members rated as being of good quality.

Similar ratings patterns and the significant correlations between overall quality and vocabulary level and style of delivery suggest that judgments of overall quality were positively influenced by the participants’ evaluations of vocabulary levels and the style of the delivery of the descriptions. For example, if the vocabulary level or style of delivery or both are rated as low, then the overall quality will be judged low as well. Further study with additional participants is needed to establish the effect of each description factor on viewers’ judgments of quality and whether there are any other unmeasured factors that could also contribute.

One important step in discovering how amateur describers make good-quality audio descriptions is to examine the common characteristics that arise in the process. At first glance, Describers D12, D10, and D9 appeared to have had a limited number of description-related characteristics in common. Describers D12 and D10 were female, and Describer D9 was male. Describer D12 was in the 30–39 age category, while Describers D10 and D9 were in the 19–29 age category. Describers D10 and D9 were familiar with *The Daily Show* and liked it, whereas Describer D12 had never seen

the show before. While Describers D9 and D10 watched more than 15 hours of television a week, Describer D12 watched none. These differences may suggest that describers’ ages, education, television-viewing habits, and familiarity with the program are not important factors that influence the quality of the descriptions that are produced.

Although there was little demographic similarity among the three top-rated describers, there were important similarities in the types, styles, and number of descriptions they produced. All three top-rated describers had a similar number of descriptions (35 for D9, 25 for D10, and 26 for D12) and total description lengths (122.79 seconds, 113.96 seconds, 165.73 s for D9, D10, and D12, respectively). This finding suggests that these describers discovered (likely by accident) the length, number, and positioning of descriptions that would be most appreciated by the audience. Furthermore, all three describers seemed to enjoy the process of description, since they rated their experience of creating descriptions as good. For example, Describer D12 stated: “The task was not a difficult one; if anything, I enjoyed it a lot.”

There were also some stylistic similarities among the high-rated describers. Describer D12’s style closely followed that of the traditional describer in that she used a third-person narrative form, and her enunciation was clear. However, her tone was bright and cheerful, and it seemed to match the overall tone of the show. Her descriptions were thorough and detailed and rarely interfered with the existing dialogue. For example, when Jon Stewart did an impression of George Bush, she said “Jon slouches over with

hands outstretched in an unsure fashion while squinting his eyes” instead of identifying him directly. This more subtle delivery allowed the audience to come to their own conclusion that Jon was imitating George Bush, which is believed to be a critical part to the humor.

Describer D10, also a woman, maintained a minimalist style, rarely describing over dialogue and describing only the minimum amount required. For example, during one scene, the image of a tiger wearing a pimp’s costume was described by Describer D10 as “The Exxon Tiger has some ‘bling-bling,’” rather than trying to describe all the elements of the costume. Using the term *bling-bling*, a pop culture reference to flashy jewelry, is a term with a light-natured feel that closely matches the style of the show. Describer D10 also had a soft, nonintrusive tone of voice, which may also have been a contributing factor to her high quality score.

Describer D9, a man, also maintained a minimalist style and, like Describer D10, described in a soft, nonaggressive, nonintrusive tone of voice. Describer D9 rarely spoke over the dialogue and, for the most part, was able to fit descriptions with the spaces that were available. Like Describer D10, D9 also used pop-culture references to help shorten his descriptions; for example, he referred to Puff-Daddy, a pop-culture figure, when describing the Exxon Tiger’s pimp costume.

Describers D4, D8, and D11 all seemed to fit in the poor describer category because their mean ratings of overall quality were below 2.3. One interesting characteristic that was common among them was that they all spoke with a noticeable accent. Describers D4 and D11 spoke

with foreign-language accents, whereas Describer D8, while a native English speaker, spoke with an Australian accent. This result, although not conclusive, suggests that description audiences may prefer describers who speak without an accent in the language used for description or, more specifically, would like to hear a describer speaking in the same vernacular as their own or that of the majority of the characters in the show they are watching. However, because of the relatively low number of describers with a discernable accent, it is difficult to determine whether this was the most important factor that contributed to the poor-quality ratings of this group.

Other factors that seem to have contributed to the levels of perceived quality were the length of the description and the total number of descriptions. Describers D8 and D4 had lower-than-average total description lengths (53.04 seconds and 34.72 seconds) and a lower total number of descriptions (18 and 12), whereas Describer D11 had descriptions that were the longest (468.80 seconds) and had a high total number of descriptions (60). This result suggests that extreme description lengths (too short or too long) and the total number of descriptions (too few or too many) may result in descriptions that are problematic for audiences. Describers D8 and D4 likely missed many possible description opportunities, while Describer D11 likely described over the dialogue or described when it was not necessary, causing the low rating of overall quality.

Limitations

Although this study indicated some promising findings, there were important limitations. One limitation is that the study

examined only audio descriptions that were created for a single 20-minute show from one genre. Further studies must be conducted to examine whether descriptions of an adequate quality can be created for different shows in the same genre, as well as for plays, movies, music videos, or videos from other genres, and the impact that those descriptions and types of programs have on audiences.

Another limitation of the study was that because only one episode of one show was used, the novelty of the concept or of the show or both may have influenced the interest and motivation of the describers to complete their description tasks and the audience's willingness to give positive or negative ratings of quality. The characteristics of descriptions that may have been desirable for this single show may not be desired once the novelty effect wears off or once audiences become more familiar with the show. Furthermore, the reviewers were exposed to five different describers during one viewing of *The Daily Show*. Had each reviewer been presented with the version from only one describer, the quality rating may have changed. To examine these questions, we recommend that a much larger longitudinal study involving many different program genres and describers be conducted.

To allow the participants to complete an entire episode of description for a single show, a program with a relatively low quantity and a complexity of description requirements was chosen for this study. Had a longer or a more visually intense and less dialogue-driven program been used, the describer may have not been able to complete the description task in a single session. In addition, audiences may have become frustrated or bored with a

long show that had poor description, which could have negatively affected their willingness to participate in this initial exploratory study.

Conclusion

The results of the two phases of this study indicated that it is possible for amateur describers to create high-quality descriptions for persons with visual impairments with little or no training in audio description techniques. Phase 2 showed that there were definite preferences for certain describers and their descriptions. These preferences seem to have been based on various characteristics of describers, such as the describers' vernacular and tone of voice and the length and timing of the descriptions. In addition, the audiences of persons who were blind, had low vision, and were sighted also showed a dislike for specific describers and their descriptions. There were few common or obvious factors, other than the describers' accents, that stood out as the cause of the low ratings. Further study is required to determine what influence specific characteristics of describers; audience factors, such as literacy levels, gender, and tolerance for an overlap of dialogue; genre preferences; television- or movie-viewing habits; and program factors, such as the length, genre, and complexity, that programs have on describers' ability to provide high-quality audio descriptions.

In Phase 1, it was found that the describers were able to complete the description task and that the tools in LiveDescribe that they used to assist in this task were easy to use and useful. It thus appears that LiveDescribe is able to support amateur describers in the tasks involved in producing audio descriptions.

Most of the describers thought that the process of description was difficult, although a few stated that it was easy and fun.

References

- Bleicher, P. (2006). Web 2.0 revolution: Power to the people. Retrieved from <https://wiki.umn.edu/pub/TELGrantCohortA/ListOfResources/revolution.pdf>
- Branje, C., Marshall, S., Tyndall, A., & Fels, D. I. (2006). LiveDescribe. In *Proceedings of the 12th Americas Conference on Information Systems: August 4–6 2006* (pp. 3033–3041). Atlanta, GA: Association for Information Systems.
- Clark, J. (2007). *The CRTC and audio description*. Retrieved from <http://joeclark.org/access/crtc/crtc-ad.html>
- Fels, D. I., Udo, J. P., Diamond, J. E., & Diamond, J. (2006). A comparison of alternative narrative approaches to video description for animated comedy. *Journal of Visual Impairment & Blindness*, 100, 295–305.
- Gagnon, L., Foucher, S., Laliberte, F., Lalonde, M., & Beaulieu, M. (2006). *Toward an application of content-based video indexing to computer-assisted descriptive*

video. Paper presented at the 3rd Canadian Conference on Computer and Robot Vision, Quebec City, Canada.

- Independent Television Commission. (2000). *ITC guidance on standards for audio description*. Retrieved from http://www.ofcom.org.uk/static/archive/itc/itc_publications/codes_guidance/audio_description/index.asp.html
- Schmeidler, E., & Kirchner, C. (2001). Adding audio description: Does it make a difference? *Journal of Visual Impairment & Blindness*, 95, 197–212.
- Snyder, J. (2004). *Audio description: The visual made verbal*. Retrieved from <http://www.audiodescribe.com/article1.html>
- Tapscott, D., & Williams, A. D. (2006). *Wikinomics: How mass collaboration changes everything*. New York: Wikinomics Portfolio.

Carmen Branje, M.M.Sc., doctoral candidate, Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Road, Toronto, ON, M5S 3G8, Canada; e-mail: <cbranje@mie.utoronto.ca>. **Deborah I. Fels, Ph.D.**, associate professor, School of Information Technology Management, Ryerson University, 350 Victoria Street, Toronto, ON, M5B 2K3, Canada; e-mail: <fels@ryerson.ca>.

Copyright of Journal of Visual Impairment & Blindness is the property of American Foundation for the Blind and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.