2005

Emotive Captioning and Access to Television

Deborah I. Fels
Ryerson University, dfels@ryerson.ca

Daniel G. Lee
Ryerson University, d1lee@ryerson.ca

Carmen Branje
Ryerson University, cbranje@ryerson.ca

Matthew Hornburg
Marblemedia, Inc., matt@marblemedia.com

Emotive Captioning and access to Television

Deborah I. Fels
Ryerson University
dfels@ryerson.ca

Daniel G. Lee
Ryerson University
d1lee@ryerson.ca

Carmen Branje
Ryerson University
cbranje@ryerson.ca

Matthew Hornburg
Marblemedia Inc.
matt@marblemedia.com

ABSTRACT

Closed captioning has been enabling access to television for people who are deaf and hard of hearing since the early 1970s. Since that time, technology and people's demands have been steadily improving and increasing. Closed captioning has not kept up with these changes. We present the results of a study that used graphics, colour, icons and animation as well as text, emotive captions, to capture more of the sound information contained in television content. deaf and hard of hearing participants compared emotive and conventional captions for two short video segments. The results showed that there was a significant difference between deaf and hard of hearing viewers in their reaction to the emotive captions. Hard of hearing viewers seemed to enjoy them and find them interesting. deaf viewers had a strong dislike for them although they did see some potential for intermittent use of emotive captions or for use with children's programs.

Keywords

Closed captioning, emotions, television production processes, deaf and hard of hearing

INTRODUCTION

While many factors contribute to the moulding and propagation of North American culture, television and film are, arguably, the most prominent and pervasive media, influencing, informing and entertaining the audiences they attract. Whether this influence is strong or weak, these two media give individuals access to information that is readily and easily available. The majority of North Americans have little difficulty accessing the medium's message, as it is presented using audio and visual content. People who are deaf or hard of hearing, however, experience limited access to these media and, because of this, are unable to fully participate in the common cultural experiences that television and film offer. Technologies that facilitate the transmission of audio content to people who are deaf/hard of hearing need to be more adequately developed to allow these individuals to enjoy television and film. The current North American technology, closed captioning, does provide some access to the sound information. Closed captioning has been available since 1980 to deaf and hard of hearing viewers who purchased special decoder boxes that attached to their televisions. Since 1993, due to the U.S. Television Decoder Circuitry Act, decoders are available in all televisions produced in North America (Robson, 2000). However, the specifications for closed captioning used with these decoders are limited and do not sufficiently meet all needs of the community.

Closed captioning for analogue television (EIA-608 standard, Abrahamian 2003)) is one of the mechanisms used to provide access to television and film (Field 2000). The current form of closed captioning uses simple text-based format, with a modified character set built into the television decoder chip. It consists of a single mono-spaced font, a single font size, a short set of text colors, and a black background (VITAC 2003). Initially, the system was limited to only white uppercase letters. Since this original specification, several options have been added, permitting the use of mixed case letters, a small set of text colors, along with a few special characters (e.g., music notes). These additions, however, are frequently absent, as users have become habituated to and prefer uppercase captioning and a white font. It is only recently that caption writers have developed an interest in using mixed-case words.

(Jordan et al. 2003) suggested that although audiences were generally satisfied with captioning quality for analogue television, missing words, spelling errors and captions moving too quickly caused dissatisfaction. Instead of using only text to provide all of the sound descriptions, it may be possible to use graphics. Graphics could be used to decrease the amount of text-based captioning required and, consequently, reduce the rate at which the captions move. Graphics could also be used to capture some of the sound information that cannot easily be described using text (often because there is no time or space in text captions for these additional descriptions).

The captioning standard for digital television (CEA-708) contains provisions for using alternative fonts, colours, caption positioning and other options related to text-based enhancements (Robson 2000). However, there have been few research initiatives or actual CEA-708 implementations investigating ways of incorporating enhancements such as colour, and font. Few available captioned examples exist that demonstrate even simple enhancements for broadcast content.

One way of improving access to broadcast content for people who are deaf or hard of hearing is to translate the missing data into an alternative visual and/or tactile modality. (Oviatt 1999) suggests that simple translation is insufficient to convey content from one mode to another. For example, speech input is better at conveying time-dependent descriptions while gesture-based input is optimum for specifying spatial information such as location of an object in a room. Attempting to express spatial information with text (as speech or written form) is less efficient, more error-prone and requires more descriptors and interpretations. While these examples from the literature are restricted to input modalities, we can apply the same caution to multi-modal output such as closed captioning embellished with graphics.

To understand the full message that is being communicated with existing closed-captioning models, the deaf or hard of hearing caption viewer must rely on visual-only cues such as body language and gesture, and combine this information with the words and short text descriptors shown within the captioning. (Mehrabian 1968) suggests that human-to-human communication consists of seven percent words, thirty-eight percent non-verbal linguistic modifiers (paralanguage), and 55% visual cues or gestures. While television viewers are not active participants in the human conversation of a television program, they are the intended receivers of this communication. It is through the interaction between humans on-screen, background sounds and music that much of a television program's semantics are conveyed. Without continual access to conversation modifiers and access to the paralinguistic components of human conversation, people who are deaf and hard of hearing may misinterpret or misunderstand the semantics of television and come away from their viewing unsatisfied.

Viewer reading speed (Jensema et al. 1996) and vocabulary levels (Jensema 1998) also limit the quantity and speed of text presentation. Because of this, often there is only enough time and space within the captions to provide a verbatim translation of what is being said and ignores other aspects central to the successful understanding of the text's full meaning. Alternative display strategies such as graphical, animated and tactile displays are needed in order to convey the additional non-verbal information such as paralanguage, music and sound effects that are often omitted. (Fels et al. 2001) have begun to investigate the use of alternative displays for expressing emotion, music and sound effects for digital television. Their initial results are very positive as viewers who are deaf and hard of hearing seem to have a different and more satisfactory experience of the captioned video content.

This paper outlines the process used in developing a new form of captioning, "emotive captioning" that uses graphics, colour and animation to illustrate sound information. Reactions by deaf and hard of hearing viewers to the video samples containing emotive captions are compared with viewer's reactions to conventional captioning for the same content. The emotive captions are designed by the research team during the process of creating the video content. The film-making team participated in designing, selecting and approving the emotive captions for their content.

MODEL OF CAPTIONING – THE PROCESS

The model of captioning that was developed for emotive captions involved two important aspects: 1) using graphics to represent emotion and sound effects; and 2) working collaboratively with the production team to design and deploy the graphical captions. These activities were applied to 1.5 minute segments of two different television vignettes that were part of an eight-vignette series called Burnt Toast: Forever and Ever, and Traffic Jammed. Each segment uses operatic form to show two different comedic aspects of relationships. For example, Traffic Jammed shows a fantasy relationship building between two individuals as they are stuck waiting in traffic.

One of the difficulties with using graphics to convey sound information is deciding what information to convey, how to best convey it, and which graphics are most appropriate. This has the potential to become a complicated and overwhelming process with large databases of graphics and icons to select from, or requiring artists to create new graphics for each new captioning project. The captioning process now becomes an exercise in interpretation rather than verbatim translation. We wanted to simplify the emotive vocabulary and the process of identifying them. There has been considerable research in the study of emotions and the identification of a primary set of mutually exclusive emotional categories. (Acton 1998) provides an excellent review of this body of literature. We decided to begin by using six emotions that seem to be most common among all of these research results: fear, anger, sadness, happiness, disgust, and surprise. At the request of the creative director, two additional categories, sexy and love were added. An intensity rating between one and five could be added to each emotion to modify its intensity.

A Caption Markup tool, available for download at www.ryerson.ca/clt, was devised to allow the creative team to tag the script with the different emotions and their intensities. Its basic function is to display a text based file (.txt, .rtf) that can then be "marked-up" by users. Each member of the creative team (director, librettist, and the script writer) marked up the script with their interpretation of the emotive characteristics of the show.

Once the team marks up their scripts, the file is imported into the Rendering Engine tool developed in this project. The main function of the Rendering Engine is to automatically create graphical pop-on captions using pre-designed image files that are associated with each different emotion variation. Captions can be edited using the simple image and text (e.g., font style, size, and text colour) tools included with the Rendering Engine.

Once the captions for each segment were produced, an audience of deaf and hard of hearing individuals evaluated the segments in comparison to a version of each segment with conventional pop-on closed captions.

RESEARCH QUESTIONS

1. What are the user responses to emotive captions developed for two different video segments?
2. How do these compare with conventional captions?

METHOD

Six deaf, American Sign Language (ASL) users (two male, four female) and five hard of hearing/deafened/deaf users (two male, three female) participated in the study. An ASL interpreter was present during the session with the deaf users. The average age of the hard of hearing (HOH) users was between 25 and 34 years and the average age of the deaf users was between 45 and 54 years. Five deaf subjects participated as a group while the remaining subjects participated individually, or in pairs due to scheduling preferences.

Participants were asked to complete a pre-study questionnaire regarding television and movie viewing habits, caption preferences and difficulties, and levels of education.

All of the deaf participants watched between one and five hours of television per week, three participants indicated that they frequently watch television alone while the other two noted that they only sometimes or seldom watch television. Four participants reported that they communicate with others about the content they are viewing while watching television and two also discussed the show after it was over.

Three of the hard of hearing participants watched six to 10 hours of television per week, two watched five hours or less and one watched 15 to 20 hours per week. Four HOH participants watched television alone sometimes, one watched television alone always, and one seldom. One participant reported that he communicates with others about the content they are viewing while watching television and five also discussed the show after it was over, one did not communicate about the show at all with family and friends.

Nine participants reported that they always use closed captioning while watching television and two HOH participants reported using it occasionally. Eight participants liked the use of text for closed captioning. Seven participants reported that they liked the following aspects of the captioning on television: verbatim captioning, and placement on the screen. Six participants liked the size of text and colour (black and white). Four participants liked the rate or speed of display and only two people liked all the text capitalized.

Eight participants suggested that background music was lacking from closed captioning. Seven participants reported that emotions in the dialogue, and having text going at the same speed as the spoken word was lacking. Five participants noted that speaker identification was lacking from closed captioning and that there was inadequate information to time jokes/puns correctly. One person identified that there were too many spelling and display errors and that ends of sentences were missing.

Two 1.5 minute video segments were selected by the Burnt Toast director as representative of the emotional and sound variety of each vignette. Emotive and conventional captions were applied to these segments and submitted for user testing.

In order to maintain interest and manage time constraints, participants were asked to view three different versions of the two video segments. Each segment was produced with three different caption styles: conventional captions, emotive captions that were located in one consistent location (lower centre of screen), and emotive captions that were placed in different locations to show speaker identification (see Figure 1-3). The emotive captions consisted of graphics, colour and icons to represent the different emotions that were identified by the director, script writer and composer using the script markup tool developed in this project. The viewing order was randomized within each set and in the presentation order of the segments.

Following each viewing, participants were asked to comment on any positive and negative aspects of each caption style. All discussions were videotaped. Following the clip viewings and discussion, all participants were asked to complete a post-study survey that asked participants to summarize their opinions of all caption varieties presented to them.

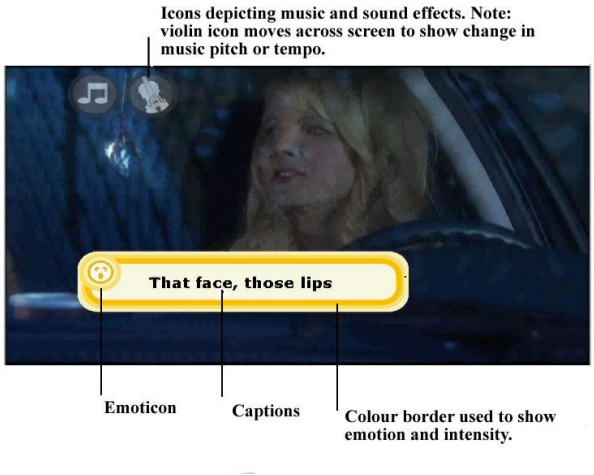

Figure 1: Dynamic emotive captions for Traffic Jammed

Figure 2: Static emotive captions for Forever and Ever

Figure 3: Conventional captions

Data analysis

Data were collected using a questionnaire/interview format for the pre and post-study survey. Discussion and dialogue that occurred during the viewing of the clips was recorded using a note taker and videotape. Facial expressions and sign language utterances as well the verbal interpretation by the ASL interpreter were also recorded. In this paper, only qualitative analyses based on the questionnaires, and observer notes are presented. A detailed video analysis has not yet been completed.

RESULTS

Pre-study questionnaire – Deaf participants (D)

People were asked to identify and rank five of the most important characteristics of closed captioning that should be modified or added (out of a list of 14 items as seen in Table 1). Five participants suggested that using text to describe background noise, and using graphics or text identify speakers were the most important additions that could be made to improve captioning. Four people suggested that using overlay captions, and using describing background sounds was important. Three participants suggested that it was important to use colour for emphasis or emotion in captions.

Ranked Item	Rank		Ranked Item	Rank	
	D	HOH		D	HOH
1. Use of graphics or symbols to denote background elements such as applause or musical inserts.	4	3	8. Different fonts or text size		
2. Use of graphics to represent background noise			9. Graphics to represent music		4
3. Text descriptions of background noise.	1		10.Text descriptions for music		
4. Use graphics or text to identify speaker	2	5	11.Faster speed of caption display		2
5. Use of graphics to represent emotion.		1	12.Use of overlay/floating captions	3	
6. Use of text descriptions for emotional information.			13.Flashing captions for emphasis.		
7. Use of colour in captions for emphasis, or emotion.	5		14.Add ASL window		

Table 1: Ranking of five most important characteristics to be modified or added to conventional closed captions.

Video segment viewing and post-viewing questionnaire – Deaf participants

Post viewing questionnaires were administered after all three related segments (e.g., static, dynamic and conventional versions of each segment). People were asked 18 forced-choice questions regarding their interpretation of the content, their opinion of the detailed positive and negative aspects of the caption presentations (e.g., colour, movement, size, speed of display, icons, and animation) and their willingness to discuss the content with hearing friends. All questionnaire responses consisted of five response Likert scales.

A Mann-Whitney analysis was carried out to determine whether there were any significant differences in questionnaire responses between the two content segments. There were no significant differences found in any of the questionnaire responses. The questionnaire responses for the two segments were then aggregated.

A CHI-square analysis of the post-questionnaire responses was carried out with the aggregated data. Significant differences were found for: 1) the colour of the emotive captioning [χ^2 (12,3)=11.33, $p<0.05$] where 9/12 (75%) strongly disliked or disliked the colours; 2) speed of display for emotive captions [χ^2 (12,3)=9.5, $p<0.05$] where 11/12 (92%) liked or liked the display speed very much; and 2) the size of conventional captions [χ^2 (12,3)=8.0, $p< 0.05$], where 10/12 (83%) liked or liked the size of the conventional captions.

From the observer notes, all of the deaf participants suggested that the emotive captions were too large and colourful. At the beginning of the content presentation, all participants reported that they found the colour distasteful and were "not comfortable with the colours on the boxes." At this point, most participants recommended that the text colour could be used for speaker identification, to emphasize emotion at certain points, or that it would be good for children.

While speaker identification was identified in the pre-study questionnaire as generally lacking from closed captions, the reaction to the caption boxes moving around the screen was confusion. Participants were confused on where to focus their attention, as they were unable to anticipate where a caption would appear. One participant also mentioned that "it forces us to lip read and I don't appreciate that". It is important that if caption position is to be used for speaker identification, there is no relationship to lip reading.

Participants also objected to having the icons represent emotions as some believed that these were "leading them to water" rather than "being able to use my own imagination." People stated that it was important to watch and interpret facial grammar to be able to "get it on my own."

The sound effects icons were better received. In Traffic Jammed, sound effects icons were used for traffic noises and some of the important musical events. For example, the conventional captions used "[soaring music]" to explain what was happening with the music. For the emotive captions, a violin icon travelling from the right edge of the screen to the left was intended to show this effect visually. Three of the six people did not notice these icons and three reported liking them, particularly the musical note icon. In the discussion of these icons, people were trying to speculate an acceptable approach to representing sound effects. Three suggested that text descriptions were sufficient, and the other two suggested that there may be other graphical alternatives but could not decide on specific examples, and one liked the icons in general.

The quality of the text was reported as good. Participants liked the mixed case and thought that using upper case letters helped convey or emphasize strong emotions (e.g., "we know it is screaming"). An example that was discussed is the use of the capital letters for "FAIR" in Traffic Jammed to represent a stronger emotion.

Some participants suggested that the emotive captions boxes were too large while others liked the size. All participants rated the speed and quality of the text as good. It was suggested that a transparent background for the caption boxes and for the sound effects icons would be a beneficial change. All participants liked the conventional captions.

One interesting observation was that as the study progressed, some individuals suggested that having colours and icons that are not part of the text might be useful for specific parts of a show (e.g., "once or twice would be okay during a show").

Pre-study questionnaire – Hard of hearing participants

People were asked to identify and rank five of most important characteristics of closed captioning that should be modified or added (out of a list of 14 items as seen in Table 1). Four participants suggested that using graphics to represent emotions was most important. Three participants suggested that displaying captions at a faster speed, using graphics or symbols to identify background noise, using graphics to represent music, and using text or graphics to identify speakers were important.

Video segment viewing and post-study questionnaire – Hard of hearing participants

Similar to the deaf participants, there was no significant difference found in the questionnaire responses between the two video segments. The questionnaire responses for the two segments were also then aggregated.

A CHI-square analysis of the post-questionnaire response was carried out with the combined data for both content pieces. Significant differences ($p<.05$) were found for: 1) icons [χ^2 (10,3)=8.33, $p<0.05$] where 77% (7/9) of responses liked or liked very much the icons representing the emotions; 2) size of conventional captions [χ^2 (10,2)=6.2, $p<0.05$] where 8/10 (80%) liked or liked very much the size of the conventional captions; and 3) willingness to engage in conversation with hearing friends about content [χ^2 (10,1)=6.4, $p<0.05$], where 90% (9/10) were very willing to discuss content with hearing friends.

From the observer notes, all of the hard of hearing participants suggested that the emotive captions were too large but that the size of the text within the caption boxes was good. However, by the end of the study all participants reported that they liked the colour and thought that it added value and made the video segments more "interesting." One person stated that he was "more laid back when watching" and two people reported that compared to conventional captions, watching emotive captions was like the "difference of using a black and white television versus colour." Some of the participants mentioned that they did not initially enjoy the colours but by the second set of captions they actually preferred it over the conventional captions (e.g., "I need time to get use to it" and "I am getting comfortable with the colours"). Only one person reported disliking the colours and had no change of opinion between the first and second viewing.

Most participants rated the quality of the text descriptions for the conventional captions and the emotive captions as good. One person was neutral about the text descriptions for the emotive captions. Two people reported disliking the closed caption text for the conventional captions and two people were neutral about it. All participants reported liking mixed case text.

It was unanimous that the moving icon to show music and the sound effects icons were very likable and aided people's awareness of music and background noise respectively. One person suggested that "people who are hard of hearing miss music and information like punch lines in jokes." In addition, most people reported liking the faces in the caption graphics making the vignettes more interesting and entertaining (e.g., "see what the actors are feeling").

As with the deaf participants, the HOH participants found that the placement of the emotive captions on the screen, particularly for the dynamic captions, was problematic. Some people noted that it helped them identify the speaker while others suggested that the moving captions were difficult to follow. Some of the captions were too close to an actors face, almost covering it. All participants also mentioned that facial expressions and grammar were critical to them and captions should never interfere with the actor's facial expressions, especially their lips. One person suggested that male/female icon faces could be used with static captions to identify who is speaking in the video clips shown.

Comparison of groups

The data for all subjects were aggregated together and a Mann-Whitney analysis was carried out to determine differences in responses between the two video segments. There were no significant differences in responses between the two video segments. The responses to the post-questionnaire data were then grouped together by participant for the two content pieces. There was thus a possible two responses per person per question for a total of 22 data points for each question. Note some participants did not answer all questions on both questionnaires.

A cross-tabs analysis of the questionnaire data showed that there were significant differences between deaf and HOH participants for willingness to engage in conversation with hearing friends about content that subject had just watched [χ^2 (22,2)=11.43]. For this question, 9/10 (90%) of the HOH responses indicated that respondents were very willing to discuss the content with friends whereas a majority of responses by deaf participants, 6/12 (50%), were neutral. There was a significant difference between the HOH and deaf groups in response to the question about appearance of the face icons shown in the graphics of the emotive captions [χ^2 (22,4)=12.74]. A majority, 80%, of HOH responses liked the face icons while a majority (75%) of deaf participants either disliked or strongly disliked them. A significant difference was found between HOH and deaf responses for the graphical representation of the emotions [χ^2 (22,3)=12.45] where 90% (9/10) HOH responses liked them very much or liked them and 67% (8/12) of deaf responses either disliked them or strongly disliked them. Finally, there was a significant difference between HOH and deaf participants for the question regarding the colour of the emotive captions [χ^2 (22,4)=11.54] where 70% (7/10) of the HOH participants liked them or liked them very much, and 9/12 (75%) of deaf participants disliked them or strongly disliked them.

DISCUSSION

The most interesting results from this study were the differences between the deaf and HOH groups. While most participants in both groups reported liking verbatim conventional captions in their pre-study questionnaire, they also identified that conventional closed captioning lacked important information such as background music, information about the emotive characteristics of the dialogue or music, and speaker identification. In this study the deaf and HOH groups seemed to diverge considerably on how that information should be expressed, and what is acceptable and what is not.

The HOH participants seemed to like, and in some cases prefer, using graphics, icons and colour to represent some of the sound information. The deaf participants generally did not accept this style of captioning as a suitable captioning alternative. In addition, both groups initially did not respond positively to the colours. However, after the second viewing the HOH group thought that colours added value and made the video segments more "interesting." The deaf viewers did not consider colour acceptable unless it could be part of text. The strong negative response of the deaf sign language users could be related to how deaf culture interacts with hearing environments in general. Many individuals in the deaf community are trying to maintain a distinct and unique culture (Israelite et al. 2002). Having access to cultural artifacts such as television from the hearing world is a matter of interpretation rather than translation because text is seen as an inadequate method of conveying information to sign language users. This may also be reflected in the fact that the deaf participants were not interested in engaging in a conversation with hearing users about the content of either piece content. It may also be a matter of comfort levels and experience as this particular group had many years of experience with closed captioning.

HOH viewers may want more translation of the sound information in a way that they find interesting and that uses less text. Many HOH viewers experienced their hearing loss over time (often having been hearing users at one point in time) and may relate better to translations of hearing information than people who are congenitally deaf and have less experience with hearing sound. The willingness of the HOH viewers to interact with hearing viewers in a discussion about the content may be an indicator of their interest in accessing hearing cultural artifacts.

Another potential cause of differences between the HOH and deaf groups found in the study was the difference in the ages (age range for HOH group was 25-34 years and for deaf group it was 45-54 years). A younger age group may be more accustom to having graphics and animation on screens due to their experience with computers and computer games and it may be a matter of gaining a comfort level with the different style. For example, some of the HOH participants reported that they liked the captions after they were "used" to them. Hence this age group, whether deaf or HOH, may be more willing to accept colours and graphics as captions. The older age group are more experienced and habituated to the conventional style of captions making it difficult for them to envision new appealing alternatives. In future studies attempts will be made to balance the age range for the two groups.

Deaf individuals and HOH individuals rely heavily on para-linguistic information that can be seen through facial expressions, and gestures. Overlays such as captions and graphics should never interfere with access to this information. In an attempt to facilitate speaker identification, captions were placed close to the person singing/speaking (free-form). Due to the size of the captions and the screen dimensions, some of the captions were placed too close to the speaker's mouth. Participants believed that this was forcing them to lip read when they did not want to, or that it was slightly covering up the mouth so they were unable to see the speaker's lips moving. CEA-708B (Association 1999) recommends standardizing on eight screen locations for captions to be placed in an attempt to solve the difficulty with speaker identification. Further studies are required to determine whether these locations would be more useful than the free-form locations used in this study and whether the difficulties encountered with keeping track of the captions and covering over the paralinguistic information are still found.

One of the other contentious issues in captioning is the rate at which captions are displayed (Jensema 1998). A majority of the deaf participants (92%) liked the speed of the display of the emotive captions whereas 80% liked the speed of the conventional caption display. All of the HOH participants liked the speed of display of the emotive captions while only 70% liked the speed of display of the conventional captions. In fact, both caption styles were appearing and disappearing at the same rate. However, the text for the emotive captions was larger and perhaps easier to read at the display rate.

Study limitations

Finally, we would like to suggest some recommendations for closed captioning based on the results of this study:

1) Closed captions should never overlap paralinguistic information such as facial expressions or gestures.
2) Alternatives to monochromatic text such as colour, graphics, icons and animation seem to have potential for use in captioning emotive sound information, music and sound effects. However, additional studies are required to create an acceptable model for applying these techniques.
3) Differences and preferences between HOH and deaf viewers could be managed through preference settings or multiple caption tracks to allow viewers to choose between graphical captions, or text only captions.

CONCLUSION

The potential for closed captioning to be much more than its current text-only form is significant particularly considering the growth and implementation of digital television and computing technologies. In addition, viewers are beginning to demand more and better quality captioning since a text version of only the dialogue is insufficient to convey all of the sound information. We have begun to explore the possibilities of using graphics, icons, animation, and colour to express sound information such as emotion contained in music and dialogue, the music itself, and sound effects. The results of our small study indicate that this approach has potential to be more expressive and interesting particularly for hard of hearing viewers. We also found that there are considerable differences in style preferences between deaf viewers and hard of hearing viewers. Finally, considerably more research is required to derive usable models for expressing sound information in alternative modalities and that these models must be evaluated with deaf and hard of hearing populations.

ACKNOWLEDGMENTS

We thank Ontario Media Development Corporation for provided funding for this project through the Pioneering Content program. We gratefully acknowledge the many contributors to this project including: Marblemedia, Ryerson University, and Rhombus Media. Finally, we thank all of the deaf and hard of hearing participants.

REFERENCES

1. Abrahamian, S. (2003) EIA-608 and EIA-708 closed captioning, retrieved May 1, 2005 from http://www.evertz.com/resources/eia_608_708_cc.pdf.
2. Acton, G.S. (1998) Basic Emotions, retrieved Feb. 15, 2005 from http://www.personalityresearch.org/basicemotions.html.
3. Consumer Electronics Association, (1999) CEA-708-B: Digital television (DTV) closed captioning, Consumer Electronics Association, retrieved Feb. 25, 2005 from http://www.ce.org/standards/StandardDetails.aspx?Id=1782.
4. Fels, D.I., Polano, L., Harvey, T., Degan, S.S., and Silverman, C.H. (2001) Towards emotive captioning for interactive television, Universal Access and Human-computer Interaction, New Orleans.
5. Field, G. (2000) Implementing DTV Closed Captions and Video Description, retrieved Feb. 28, 2005 from http://ncam.wgbh.org/dtv/overview/nab2000paper.html.
6. Israelite, N., Ower, J., and Goldstein, G. (2002) Hard-of-hearing adolescents and identity construction: Influences of school experiences, peers, and teachers," *Journal of Deaf Studies and Deaf ldentification*, 7, 134-148.
7. Jensema, C., McCann, R., and Ramsey, S. (1996) Closed-captioned television presentation speed and vocabulary," *American Annals of the deaf*, 141, 4, 284-292.
8. Jensema, C.J. (1998) Viewer reaction to different television captioning speeds, *American Annals of the Deaf*, 143,4, 318-324.
9. Jordan, A.B., Albright, A., Branner, A., and Sullivan, J. (2003) The state of closed captioning services in the United States, National Captioning Institute Foundation, 1-47.
10. Mehrabian, A. (1968) Communication without words, *Psychology Today*, 2, 4, 53-56.
11. Oviatt, S. (1999) Ten myths of multimodal interaction, *Communications of the ACM*, 42, 11, 74-81.
12. Robson, G.D. (2000) FCC adopts caption standards for digital television, retrieved April 15, 2002 from www.robson.org/gary/captioning/dtv-standards-july2000.html.
13. VITAC (2003) History of DTV Captioning, retrieved Feb. 28, 2005 from http://www.vitac.com/technical/dtvhistory.htm.