

**Requested Technical and Operational Information
Related to the Provision of Closed Captioning by
Canadian Private Broadcasters**

Prepared by:

**The Canadian Association of Broadcasters'
Working Group on Closed Captioning**



Quebec City, Quebec

July 17th, 2006

Introduction:

In November 2005, the Canadian Association of Broadcasters (CAB) met with representatives of the Canadian Association of the Deaf (CAD) and CRTC staff in Winnipeg to discuss a number of issues related to closed captioning. The CAB was asked to follow up on a handful of issues most of which were technical or operational in nature.

Specifically, the CAB was asked to provide information relating to the following four areas: status of voice-recognition technology for closed captioning; the impact of HDTV and digital transmission on closed captioning; interference with closed captioning as a result of "cut-aways" to commercial messages; and the provision of emergency messages to captioning consumers.

This report was prepared in response to the CAB's undertaking and will be presented by the Chair of the Joint Societal Issues Committee, Jon Medline, during the Deaf Canada Congress in Quebec City on July 17, 2006.

I) Commercial Messages – Interference with Captioning

Issue:

Due to 5-6 second time delay, captioning offered by some broadcasters is cut off when leading into a commercial break.

Description:

This problem is related to the technical infrastructure of how a station's signal gets to air. For some broadcasters, the closed caption encoder is on the output of the production room, while the master control that inserts the commercials is in another part of the facility, or even another part of the country. In this scenario, if the captioning is not complete when the commercial starts, the closed caption encoder is effectively "cut off".

A contributing factor to this problem is the delay in the captioning itself. The delay is caused by a couple of factors: 1) the reaction time from when a captionist hears the words, to when they type it on their keyboard; and 2) the technical means by which a broadcaster receives the captioning. Many closed captioning companies use satellite reception, typically Star Choice and/or Bell ExpressVu, to get their audio. Captionists are often not in the same city as the station they are captioning which presents an added challenge since being able to see the pictures gives the captionist context that helps them both place their captions and improve accuracy. Using satellite reception adds somewhere between a 1½ to 2 ½ second delay to the captionist's hearing the audio, on top of any reaction delay.

Options:

One possible option is to ensure the closed caption encoder is placed downstream of the master control facility; however there are many factors that do not always make this practical. Many companies like to do a test before the broadcast to ensure the captioning is working. This would be difficult to do downstream of master control, as the test would go to air thereby disrupting whatever program is on the air. This would be even more problematic for news programs which run on multiple stations. For example, a national news program may run on 15 stations. If the closed captioning encoder were downstream of the station's master control, you would need 15 captionists working on the program. This is a technical solution that may or may not be viable for all broadcasters and/or stations.

Another possible option which some broadcasters have implemented is to alter their newscasts so that there is no material to be captioned at least 5 seconds before the commercial break starts. For example, a bumper and music with on screen text promoting what is coming up next on the program which does not require captioning. This option is recommended in the CAB's Closed Captioning Standards and Protocol Manual for English-language Broadcasters.

II) Captioning Interference with Emergency Messages**Issue:**

Closed captioning interference with emergency messages

Description:

There are many different kinds of emergency messages including systems which break into the signals to messages that "crawl" across the bottom of the screen. Captioning is most likely to interfere with emergency messages when "crawls" are used.

Closed captions are actually displayed by the user's television. The television decodes the data stream with the closed caption information from the vertical interval, and displays it on the screen assuming the user has told his/her television to do this. Because the captioning data is added by the television, there is no way to place "crawls" or other information over the closed captioning. The closed captioning will always be on top of everything else.

It should also be noted that in most cases broadcasters do not control where the captions are placed. Foreign programs come with captioning, and those captions could appear anywhere on the screen.

Option:

Since captions are "usually" placed in the lower third of the screen, doing an emergency "crawl" at the top of the screen may reduce, but not eliminate the issue of crawls appearing over captioning.

It might also be possible to "stop" captions from appearing while the emergency crawl is on the air. This could be done by "keying out" line 21 on the vertical interval. It should be noted that is not something most master control systems are designed to do so it may not be technically possible. Some stations may be able to do this while others will not depending on their equipment. However, since crawls can run quite often, for example every 15 minutes in the case of Amber Alerts, it may make the program somewhat unwatchable to captioning consumers.

III) HDTV and Digital Transmission Impact on Closed Captioning

Common Questions:

How does HDTV and/or digital transmission change the way closed captioning works?

In standard definition (SD), (either NTSC analog or SD-SDI digital) captions are encoded onto line 21 of the Vertical Blanking Interval. This means that captioning data information essentially becomes part of the video signal and as such follows the same path to the viewer. With HDTV a separate caption data packet is used for the transport of closed captioning data that must be carried intact all the way to the viewer.

Generally speaking, it requires the same amount of effort to caption an HDTV show as it does for an SD show. However, HDTV captioning provides greater flexibility by enabling the viewer to control the caption display including font style, text size and color, and background color.

Also, broadcasters only have to deal with one standard when handling SD. Alternatively, HDTV comes in many standards (e.g. 1080i, 720p etc.) and within the broadcast plant it is important to ensure that captioning data is maintained during the various conversion processes entailed in handling multiple standards.

EIA-708 is a new closed caption standard developed for high definition television that is recognized internationally. However, this standard is not currently supported by a large number of high definition equipment manufacturers so additional support has been developed for the original NTSC closed caption standard EIA-608 to be translated through the television sets.

Should we re-think closed captioning because of HDTV (e.g., color-coding the dialogue of different characters)?

It should be noted that while color-coding of dialogue is possible in SD, HDTV offers even more capability for more sophisticated captioning displays to be used e.g. different fonts for different genres. Despite this, it must be kept in mind that for some time to come broadcasters will have to service both SD and HDTV viewers and that intelligibility of captioning is the most crucial factor for both services

Fortunately, the EIA-708 standard was developed with this in mind but because it is not currently supported by all television manufacturers some of these advanced features cannot be exploited by either the consumer or the broadcaster. Again the EIA-708 standard was designed to improve opportunities for captioning consumers but unless support of EIA-708 is mandated for equipment manufacturers it is not likely to become fully supported.

Finally, the CAB has undertaken to review its Closed Captioning Standards and Protocol Manual on a regular basis and could consider establishing certain protocols for captioning in HDTV where possible and if desired.

Does the closed captioning cost equation change because of HDTV/digital transmission?

Broadcasters must make a large initial investment in HD captioning equipment. However, once broadcasters have made this investment the cost of producing captioning is the same as SD.

Do we need parallel technological solutions (i.e., one for continued analogue transmission and one for digital transmission)?

This is very important. The recognized need for EIA-608 translation through high definition devices addresses this.

Can distributors handle HDTV closed captioning?

The closed caption information is contained within the Metadata of the high definition transport signal. Some equipment manufacturers do not “pass through” this information so testing is required within the particular transport stream to ensure that this information is indeed passed through. Distributors encountered some problems during the very early days of HDTV in Canada. Fortunately, the situation appears to have been resolved.

IV) Voice Recognition Technology (VRT) for Closed Captioning

Status of VRT:

Most of the voice recognition software that currently exists is for commercial or personal use and not intended for use by broadcasters or for closed-captioning. However, TVA, in partnership with le Centre de recherche informatique de Montréal (CRIM) and le Regroupement québécois pour le sous-titrage (RQST), have been developing a voice recognition software since 2002 to produce live closed-captioning for French-language news programming. Since September 2005, TVA uses this technology to caption all of its live segments from morning shows to nightly news. Other French-language services have expressed interest in using the VRT developed by TVA/CRIM/RQST. In addition, VRT is currently being researched for applications in the English market where very few broadcasters use this technology.

Over the past year, they have received three prestigious awards for this new technology:

- OCTAS Award given by the Fédération informatique du Québec in the “Strategic Partnership” category;
- An award from l’Association de développement et de recherche industrielle au Québec in the "Partnership" category; and
- A silver award of excellence given by the Canadian Information Productivity Award in the “Customer Care” category

While VRT for closed captioning is still not developed enough to replace the stenographer approach, there are clearly benefits to using this technology which include:

- Equipment cost is more affordable;
- It takes about 3 hours to train a new re-speaker compared to 3-5 years for training a stenographer; and
- The voice recognition “dictionary” is unique and can be customized.

It should be noted that the application of VRT to closed captioning does not result in an automated process but still requires oversight by a trained and experienced captionist.

Although the quality for any technology used for live closed-captioning will always be lower than for off-line closed-captioning – since the software is calculating the probabilities for spelling a “sound” correctly and there is not enough time to manually correct spelling mistakes - the voice recognition technology developed by the CRIM is very promising. The results are more likely to increase with wider use and adoption. Some challenges remain for proper nouns and homonyms.

Recently, Réseau des Sports (RDS) also used CRIM’s resources during the hockey playoffs to produce live closed-captioning for hockey games. In addition, TVA, in partnership with CRIM, is now working on the possibility of providing live closed-captioning for French-language entertainment programs.

Common Questions:

Is voice recognition software capable of handling all closed captioning requirements (i.e. could/should we eliminate the stenographer approach)?

Voice recognition software exists and is used primarily in live news or sports applications. The current level of performance on the available products is approximately 70-80% accuracy when used with a “trained” voice. While it is possible to “train” a newscaster’s voice using voice-recognition technology for other program categories would be impractical. For this reason, voice-recognition technology is a long way from ever replacing the traditional approach to captioning.

What is the impact of voice recognition technology on closed captioning quality (e.g., more typos; mangled names etc)?

The current level of accuracy for voice recognition technology, without a re-speaker, is unacceptable for today's standards.

How readily available is this technology -- and how costly would it be to implement for a broadcaster - both conventional and specialty (hardware; software)?

Pricing for VRT is comparable to the price of software currently being used. There are cost implications however since additional work is involved in getting transcripts of the programming to voice the dialog.

Who makes voice recognition software?

The following are known voice recognition software manufacturers: Softel, Starfish and Mitel all of which have Canadian versions and use a Canadian vocabulary.

Would voice recognition solve the time delay issues (e.g., could live sports be done better with voice recognition)?

Live sports and news are the dominant drivers of VRT. However, since the "captioner" has to re-voice all words announced in the original dialog any real time savings are limited. Both technologies rely on real time execution.

Would adoption of voice recognition for closed captioning be affected by HDTV/digital transmission?

Voice recognition is just one of the tools that may be used to generate captioning. Its use is the same regardless of whether the program is in HDTV or SD.