Annie Schweikert
Video Preservation I, CINE-GT 3403
Kelly Haydon and Ben Turkus
15 October 2018

Closed Captioning: Technical Outlines and Archival Access

I. A brief history of closed captioning

Closed captioning is the practice of displaying a transcription with dialogue and other audio cues over a broadcast or streaming audiovisual program. Captioning allows people with hearing impairments to understand the words spoken onscreen and identify who spoke them, and to recognize the musical cues and sound effects that accompany the dialogue. In this way, captioning differs from subtitling, which is intended for those who can hear those non-spoken elements but cannot understand the speech itself.[1] While captions are an integral part of providing access for people who cannot hear, they also benefit anyone who cannot hear a program in a noisy place, as well as serving as an educational tool for those trying to learn a second language (for example, English learners in the United States).[2] In addition, the ability to extract captions gives archives of broadcast materials an excellent means of searching for and describing content, both for hearing and hearing-impaired patrons.

The history of American captioning is relatively recent. The first captions were broadcast on television by PBS in 1972 as "open" captions—that is, visible to everyone and could not be turned off). However, these open-captioned programs ended up competing with similar uncaptioned programs for broadcast times. Closed captioning was previewed in 1971 and tested over the next several years until 1976, when the Federal Communications Commission (FCC) approved and reserved the use of line 21 to transmit the captions. Despite this approval, the first

[1] University of Washington Disabilities, Opportunities, Internetworking, and Technology Center, "What is the difference between open and closed captioning?" (factsheet), n.d., accessed 14 Oct. 2018, https://www.washington.edu/accessit/print.html?ID=1050

[2] George Spanos and Jennifer J. Smith, "ED321623 1990-08-00 Closed Captioned Television for Adult LEP Literacy Learners. ERIC Digest," ERIC Institute of Education Sciences, U.S. Department of Education, Aug. 1990, pg. 3. https://files.eric.ed.gov/fulltext/ED321623.pdf

closed-captioned shows on American television were not broadcast until 1980. Real-time captioning followed in 1982, in which court reporters transcribed programs live.[3]

Closed captioning was strengthened through legislation in the 1990s and 2000s. The Television Decoder Circuitry Act of 1990 required new televisions 13 inches or larger to have the built-in ability to decode captions, decreasing the cost barrier to captions; prior to this act, captions could not be displayed without a separate decoder.[4] 1990 also brought the Americans with Disabilities Act, mandating access to verbal information in public places (not including movie theaters)—a requirement that captioning can satisfy.[5] Captioning availability increased again in response to the Telecommunications Act of 1996, when the FCC began to require closed captioning for almost all English and Spanish television programming.[6] 2010's 21st Century Communications and Video Accessibility Act expanded captions' reach to online and streaming media by requiring captions on Internet distributions of captioned broadcasts.[7] Most recently, the World Wide Web Consortium's Web Content Accessibility Guidelines (version 2.0), which requires captions for all prerecorded and live audiovisual content, has become a benchmark for measuring compliance with the ADA and has been adopted in phases by the U.S. federal government. [8] [9] Captions have become mandatory in every arena, making them a reliable way to provide access both to viewers and to researchers—assuming, of course, that the quality of the captions is good.

II. Standards and mechanics of closed captioning

a. *NTSC: EIA-608, CEA-608, or line 21 closed captions*

---

[3] The previous paragraph is a summary of the excellent narrative at "A Brief History of Captioned Television," National Captioning Institute, n.d., accessed 14 Oct. 2018, http://www.ncicap.org/about-us/history-of-closed-captioning/

[4] Television Decoder Circuitry Act of 1990, Pub. L. No 101-431, § 3, 104 Stat. 960, 15 Oct. 1990, https://www.gpo.gov/fdsys/pkg/STATUTE-104/pdf/STATUTE-104-Pg960.pdf

[5] Americans with Disabilities Act of 1990, Pub. L. No. 101-336, Appendix A, 104 Stat. 328, 26 July 1991, https://www.ada.gov/reg3a.html

[6] "A Brief History of Captioned Television," National Captioning Institute.

[7] 21st Century Communications and Video Accessibility Act (CVAA), Pub. L. No 111-260, § 718, S. 3304, 5 Jan. 2010, https://www.gpo.gov/fdsys/pkg/BILLS-111s3304enr/pdf/BILLS-111s3304enr.pdf

[8] "Guideline 1.2 Time-based media," World Wide Web Consortium, "Web Content Accessibility Guidelines (WCAG) 2.0," eds. Ben Caldwell; Michael Cooper; Loretta Guarino Reid; Gregg Vanderheiden et al., 11 December 2008, https://www.w3.org/TR/WCAG20/

[9] Architectural and Transportation Barriers Compliance Board, "Information and Communication Technology (ICT) Final Standards and Guidelines," U.S. Federal Register 82 (11): 5790-5841. https://www.access-board.gov/attachments/article/1877/ict-final-rule.pdf

Analog NTSC broadcasts are captioned in the 608 or line 21 format. The prefixes EIA and CEA refer to the Electronic Industries Alliance and their subdivision Consumer Electronics Association, and are interchangeable.[10] In this format, the captioning data is transmitted in the vertical blanking interval. Data is carried on line 21 on both fields during this interval; two bytes are transmitted per field, which allows a maximum of 120 characters per second.[11] These bytes are equivalent to two characters, or to a two-byte command controlling display styles, though there are few options beyond white text (often all capitals by convention) on black background.[12] [13] Each caption field contains two closed caption channels and a text channel, with a total of four CC channels and two text channels available per frame. However, due to bandwidth limitations, one two CC channels are used; a typical layout reserves CC1 for English captions and CC3 for Spanish captions.[14]

b. *ATSC/DTV: EIA-708 or CEA-708*

Digital ATSC broadcasts are captioned in the 708 format, in which the closed captions are embedded within the transmission bitstream. Specifically, the caption text is carried in the Picture User Data of the MPEG-2 Transport Stream,[15] which is the standard ATSC container format for audiovisual materials.[16] This means that 708 captions are processed like image content, instead of transporting captions in inactive video (as in 608). 708 streams can also carry 608 captions in a format known as 608 over 708; in this format, the viewer's digital to analog converter inserts the 608 captions into the vertical blanking interval.[17] The FCC requires both 708 and 608 over 708 captions in digital broadcast, and for this reason many broadcasters still begin with 608 captions and "upconvert" to 708.[18] 708 features include widely expanded font,

---

[10] "CEA-708," The Closed Captioning Project LLC, 2018, accessed 14 Oct. 2018, http://www.theclosedcaptioningproject.com/?p=389
[11] Doug Keltz, "Understanding & Troubleshooting Closed Captions," July 2014, accessed 14 Oct. 2018, https://www.smpte.org/sites/default/files/section-files/2014_July_Closed_Captioning.pptx
[12] Glenn Eguchi, "Introduction to Closed Captions" (technical paper), Adobe, April 2015, pg. 13, accessed 14 Oct. 2018, https://www.adobe.com/content/dam/acom/en/devnet/video/pdfs/introduction_to_closed_captions.pdf
[13] "The Basics of 608 vs. 708 Captions," Aberdeen Broadcast Services, 18 June 2009, accessed 14 Oct. 2018, https://www.abercap.com/blog/2009/06/18/the-basics-of-608-vs-708-captions/
[14] Ibid.
[15] Eguchi, "Introduction to Closed Captions," pg. 13.
[16] Advanced Television Systems Committee, "Recommended Practice: Guide to the Use of the ATSC Digital Television Standard, including Corrigendum No. 1," 4 Dec. 2003, accessed 14 Oct. 2018, https://www.atsc.org/wp-content/uploads/2015/03/a_54a_with_corr_1.pdf
[17] Eguchi, "Introduction to Closed Captions," pg. 13
[18] Eguchi, "Introduction to Closed Captions," pg. 14.

color, placement, and background options, as well as an increase in transmission rate—9600 bits per second, as opposed to 608's 960 bits per second.[19]

c. *Streaming and online media*

Web captions exist in several formats, the most prominent of which are WebVTT (Web Video Text Tracks, formerly known as WebSRT for Subtitle Resource Tracks) and TTML (Timed Text Markup Language). WebVTT, originally developed by Google, is a text format with a header and cues, or text sequences mapped to start and end times.[20] Advantages of WebVTT include its consistency (all instances of WebVTT must conform to a single specification), as well as its interoperability: both 608 and 708 captions can be translated to WebVTT, integrated with HTML5, and played across web browsers, iOS, and Android.[21] TTML, another text format, delivers caption text in an XML-based format. Though it is less web-accessible, it is favored by the broadcast industry for its integration with standards bodies. Unlike WebVTT, it is a recommendation of the World Wide Web Consortium, and a SMPTE-developed profile of the TTML format called SMPTE-TT has been endorsed by the FCC.[22]

d. Caption extraction

Technical methods of capturing and extracting captions vary based on format. When digitizing analog videotapes with 608 captions, the line 21 signal will be captured as a data packet within the vertical blanking interval. Since analog decks and cables transmit all active and inactive video lines, the captions will be retained as part of the video signal.[23] It is important to note that the video signal must be captured uncompressed, as compression may scramble the caption information.[24]

Digital captions can be captured through SDI or by simple extraction. ATSC captions (708 and 608 over 708) may be captured through an HD-SDI connection. These digital captions

---

[19] Gerry Field, "DTV Captioning: Transitioning from 608 to 708" (presentation), DTV Caption Summit, WETA, Arlington, VA, 14 Mar. 2001, accessed 14 Oct. 2018, http://main.wgbh.org/wgbh/pages/ncam_old/dtv/dtvccsummit/
[20] Eguchi, "Introduction to Closed Captions," pg. 16.
[21] World Wide Web Consortium, "TTML and derivative Captions Formats," 7 Jan. 2016, accessed 14 Oct. 2018, https://www.w3.org/AudioVideo/TT/docs/TTML-Profiles.html
[22] Ibid.
[23] Keltz, "Understanding & Troubleshooting Closed Captions."
[24] Kimberly Tarr, Michael Grant, Lauren Alberque, and Carleton L. Jackson, "Capturing Captioning: Problems in Preservation and Presentation of Timed Text" (presentation), 11 Nov. 2016, description accessed 14 Oct. 2018, http://www.amiaconference.net/amia-2016-presentations/

are embedded within the SDI bitstream in the picture user data, in accordance with SMPTE's standards for embedding vertical ancillary data information.[25] For web streaming videos, several tools exist to extract the closed caption text files as .vtt (WebVTT), .srt (WebSRT), or .xml (TTML) files, among other extensions.

In all cases, the digitized captions may be extracted with a number of proprietary and open-source tools, notably FFmpeg and CCExtractor. Please see Appendix II for sample extraction commands. A number of file formats and codecs have the ability to store captions, including MXF with JPEG 2000, MPEG-2 with MPEG-2, MPEG-4 with H.264, DV (608 only), and QuickTime with Apple Pro Res.[26] The existence of captions in digital video may be detected with audiovisual identification programs such as MediaInfo, which will report captions as text fields.[27] Not all players support closed caption display; as of 2015, Windows Media Player only supported captions in .wmv video files, QuickTime Player only supported .mov files, and VLC supported a partial list of file and caption formats, while Chrome, Firefox, Safari, and Internet Explorer only supported WebVTT consistently.[28]

III. Significance in digitizing videotapes

Capturing the captions in an analog or digital videotape, broadcast, or stream is crucial for a number of reasons. Captions are not just metadata—they are carried in the video signal itself, and as such represent an integral part of the video's content that must be captured. Captions do differ from the video portion of the content in an important way: they can be searched for keywords. Once extracted and turned into a transcript—with little extra effort on the part of the archivist—captions can greatly expand access to undescribed audiovisual materials for all users.

It is also important not to lose sight of the users for whom captions were developed: people who are deaf or hearing impaired. For these researchers, who experience many barriers to the use of audiovisual materials, captions are not just helpful but essential. Institutions uploading

---

[25] SMPTE 334M, as described by Sarkis Abrahamian in "EIA-608 and EIA-708 Closed Captioning," Evertz, n.d., accessed 14 Oct. 2018, https://evertz.com/resources/eia_608_708_cc.pdf

[26] https://sourceforge.net/p/mediainfo/support-requests/31/

[27] jimet65, "VLC support of Closed Captions" (forum post), 21 Aug. 2016, accessed 15 Oct. 2018, https://forum.videolan.org/viewtopic.php?t=135163

[28] "Video Player support for closed captioned videos," Telestream Knowledge Base, Aug. 2015, accessed 15 Oct. 2018, http://telestream.force.com/kb/articles/Knowledge_Article/Video-Player-support-for-closed-captioned-videos

their captioned content also have the opportunity to expand access to these users from afar. Finally, captioning web content is not just a bonus way to expand access, but a requirement; institutions trying to meet the Web Content Accessibility Guidelines 2.0 must provide captions for all prerecorded audio and video (for level A compliance) and live audio in audio and video streams (for level AA compliance).[29] Archives that upload already-captioned content have a leg up in meeting these requirements. In sum, closed captioning is both an integral part of accessibility standards and a way to increase discoverability for all users of audiovisual content.

[29] W3C, "Web Content Accessibility Guidelines (WCAG) 2.0."

Appendix I. Further reading

Captioning overviews

For captioning history:

> "A Brief History of Captioned Television," National Captioning Institute, n.d., http://www.ncicap.org/about-us/history-of-closed-captioning/

For a technical overview of captions for analog, digital, and web-based captions:

> Glenn Eguchi, "Introduction to Closed Captions" (technical paper), Adobe, April 2015, pg. 13, accessed 14 Oct. 2018, https://www.adobe.com/content/dam/acom/en/devnet/video/pdfs/introduction_to_closed_captions.pdf

For more detail on how 608 and 708 captions are transported and extracted:

> Doug Keltz, "Understanding & Troubleshooting Closed Captions," July 2014, accessed 14 Oct. 2018, https://www.smpte.org/sites/default/files/section-files/2014_July_Closed_Captioning.pptx

Web captioning

For an example of how analog and digital captions are translated to file-based formats, see the map between 608/708 captions and WebVTT:

> World Wide Web Consortium. "Conversion of 608/708 captions to WebVTT: Draft Community Group Report." ed. Silvia Pfeiffer. 15 Oct. 2018. https://dvcs.w3.org/hg/text-tracks/raw-file/default/608toVTT/608toVTT.html

For the text of the Web Content Accessibility Guidelines (WCAG) 2.0, adopted by the U.S. federal government and others:

> World Wide Web Consortium. "Web Content Accessibility Guidelines (WCAG) 2.0." eds. Ben Caldwell; Michael Cooper; Loretta Guarino Reid; Gregg Vanderheiden et al. 11 December 2008. W3C Recommendation. https://www.w3.org/TR/WCAG20/

Appendix II. Caption extraction with open-source tools

608

FFmpeg command to extract 608 captions (from ffmprovisr: https://amiaopensource.github.io/ffmprovisr/index.html#readeia608)

```
ffprobe -f lavfi -i movie=input_file,readeia608 -show_entries
frame=pkt_pts_time:frame_tags=lavfi.readeia608.0.line,lavfi.readeia608.0.cc,l
avfi.readeia608.1.line,lavfi.readeia608.1.cc -of csv > input_file.csv
```

708

CCExtractor command to extract 708 captions (from CCExtractor documentation: https://www.ccextractor.org/public:gsoc:olegkisselef_cea_708)

```
ccextractor input_file -svc all[CHARSET]
```

WebVTT

FFmpeg command to extract or convert embedded WebVTT captions (from FFmpeg documentation: https://trac.ffmpeg.org/wiki/ExtractSubtitles):

```
ffmpeg -webvtt text -i input_file out.vtt
```