Chapter 17. Personal Reflections on the Educational Potential and Future of Closed Captioning on The Web

Sean Zdenek

Texas Tech University, Lubbock, Texas USA

ABSTRACT

This chapter explores the value of closed captioning in universal design. While closed captions positively impact a wide range of our students—deaf, hard of hearing, and hearing—they also have the potential to create more robust and interactive digital learning systems. Caption technology can address the current limitations of video search and retrieval by offering students fully searchable, fully clickable interactive transcripts. The future of closed captioning on the Web will offer students a means to search the video collection of an entire course, or even across all of the videos produced in all of the courses of a department, college, or university. In this future learning environment, captions will enable students to use keywords not only to find and review course content across multiple videos but also to insert their own "margin" notes, share comments with students, and create customizable video mash-ups as study guides.

Keywords: *Closed captions, interactive transcripts, deafness, video, web accessibility*

INTRODUCTION

As closed captions on the Web become more common, and online captioning technology becomes more sophisticated, we are beginning to see captions deliver on the promises of universal design (Chisholm & May, 2009) to make digital video more accessible to every student, regardless of hearing ability. This chapter describes how captions are enabling more robust data mining techniques by facilitating digital video search, retrieval, analysis, and synthesis.

BACKGROUND

On television, nearly all English language content is required by U.S. law to be transmitted with closed captions (FCC, 2010). On the Web, Section 508 of the Rehabilitation Act of 1973, as amended in 1998, requires federal agencies that "develop, procure, maintain, or use electronic and information technology" to make their products and services, including their websites, accessible (Section 508). §1194.22b of Section 508 mandates the use of synchronized alternatives (e.g. open or closed captions) for video content: "Equivalent alternatives for any multimedia presentation shall be synchronized with the presentation" (Section 508). In the private sector, businesses that have contracts, or hope to have contracts, with the federal government must ensure that the products they deliver to the government comply with Section 508. The Americans with Disabilities Act (ADA) *may* also require closed captioning

for the Web's private sector. For example, the judge presiding over the landmark *National Federation of the Blind v. Target* case ruled in 2006 that the Americans with Disabilities Act (ADA), which was signed into law before the advent of the Web, applies to private businesses regardless of whether goods and services are offered in brick-and-mortar stores or online. "Judge Marilyn Patel rejected Target's position that their site couldn't be sued under the Americans with Disabilities Act (ADA) because the services of *Target.com* were separate from Target's brick-and-mortar stores" (Chisholm & May, 2009, p. 16). But because Target settled the case in 2008 "without admitting any wrongdoing," "the question of the ADA's applicability to the Web [is] somewhat unresolved" (p. 16). Regardless, the Department of Justice has declared that the ADA does indeed apply to the Internet. According to Thomas E. Perez, Assistant Attorney General in the DOJ's civil rights division,

> It is and has been the position of the Department of Justice since the late 1990s that Title III of the Americans with Disabilities Act (ADA) applies to websites. We intend to issue regulations under our Title III authority in this regard to help companies comply with their obligations to provide equal access. (quoted in Evans, 2010)

Finally, the "21st Century Communications and Video Accessibility Act," signed into law by President Obama in September 2010, requires all TV and comparable programming, when (re)transmitted on the Web, to be available with closed captions (see Valentino-DeVries, 2010). This landmark legislation will ensure that when TV shows and movies are redistributed on the Web by the original TV networks and authorized retransmitters like Hulu™ and Netflix™, they are accompanied with closed captions (e.g. see Zdenek, 2009).

In an educational setting, captions are designed to provide a synchronized text transcription of speech and other significant sounds for students who are deaf and hard of hearing. But captions also benefit our non-disabled students, particularly when these students are temporarily or *situationally disabled* due to "changes in one's abilities based on environment, device, or other temporary conditions" (Chisholm & May, 2009, p. 12). For example, a hearing student who tries to study with her laptop in a noisy student union building (assuming she left her headphones in her dorm room) may have difficulty fully hearing and thus fully understanding an instructor's uncaptioned video lecture. If this student is also a visual learner who retains information better through written language, then closed captioning may be necessary to help this student reach her potential, regardless of whether she studies in a quiet or noisy area. A wide range of our student population stands to benefit from captions: students who are deaf or hard of hearing, very young children learning to read, children and adults learning a second language, military veterans with hearing loss who are returning to college,[1] college students reviewing a professor's video lectures in preparation for an exam, late-career adults and seniors returning to school to pursue a second or third career,[2] and so on. Captions have an arguably much wider appeal than we have hitherto assumed. Instructors should never settle for showing uncaptioned videos in the classroom. Whenever possible, captions should also be available to students reviewing course content outside of the classroom, regardless of the student's age, class level, or presumed hearing ability. Captions are key to universal design and achieving the goal of an optimally accessible learning environment.

Because closed captions on the Web are saved as text files, they can be fed to search engines and retrieved by keyword searches. Search engines are not very good at indexing the content of audio or video files. Indeed, Google's™ search engine has been metaphorically compared to both a blind and a deaf user (see Chisholm & May, 2009, p. 14). But search engines thrive on plain text: Tags, keywords, text descriptions, text transcripts, and, of course, closed captions (Ballek, 2010; Stelter, 2010; Sizemore, 2010). Without the benefit of searchable text captions, students will often find it difficult and frustrating, to say the least, to manually scan lecture videos looking for that one example, anecdote, or solution that they vaguely remember from class but can not locate quickly or easily in the recorded video lectures.

Interactive transcripts raise the value of captions further by allowing users to click on a single word in a video transcript and be transported to that moment in the accompanying video where that word is spoken or appears. I first became aware of, and then immediately recognized the immense game-changing power of, interactive transcripts on TED.com. Because captions on TED.com are crowdsourced out to regular users ("TED Open Translation Project"), many of the videos on the site are available in an impressive number of languages. One could, for example, listen to Aimee Mullins (2009) speaking in English, read the captions in a second language such as French, and browse the interactive transcript in a third language such as Japanese. (Or one could simply load captions and interactive transcript in English, which is what I do.) In the case of Mullins' (2009) TED talk, users can choose from thirty-two languages. YouTube™ has also started offering interactive transcripts for the captioned videos in its collection (Chitu, 2010). Companies such as 3Play Media™ and ProTranscript™ also provide, as part of their regular video transcription service, a video player plugin that serves up interactive, clickable transcripts alongside closed captions. 3Play Media™ also supports "archive searching" across a website's video collection, and has recently introduced a "clipping plugin" that allows users to "[c]lip video segments simply by highlighting the text. Rearrange clips from multiple sources and create your own video montages" (3Play Media). The video clipping plugin will output a Uniform Resource Locator or URL for sharing montages with other users. When these users are college students attending the same university, or enrolled in the same course, the video montage—fully accessible because it is built on closed captions—could be a powerful, accessible learning tool indeed.

CASE STUDY

In my graduate-level course on Web Accessibility and Disability Studies, I have begun experimenting with interactive (clickable and searchable) transcripts for my video screencasts. Interactive transcripts exemplify inclusive multimedia design as well as provide students with a more accessible way to mine the video content of my course. I ask students to imagine an open, fully searchable university populated by hundreds or thousands of lectures and other videos. Because I teach in a technical communication program, I also ask students to consider the value and limits of searchable, captioned media as a form of user documentation. In Figure 1, the video is closed captioned. Each word in the transcript below the video is time-stamped and clickable. The transcript is fully searchable and automatically scrolls in time with the video. Individual words are highlighted as they are spoken.
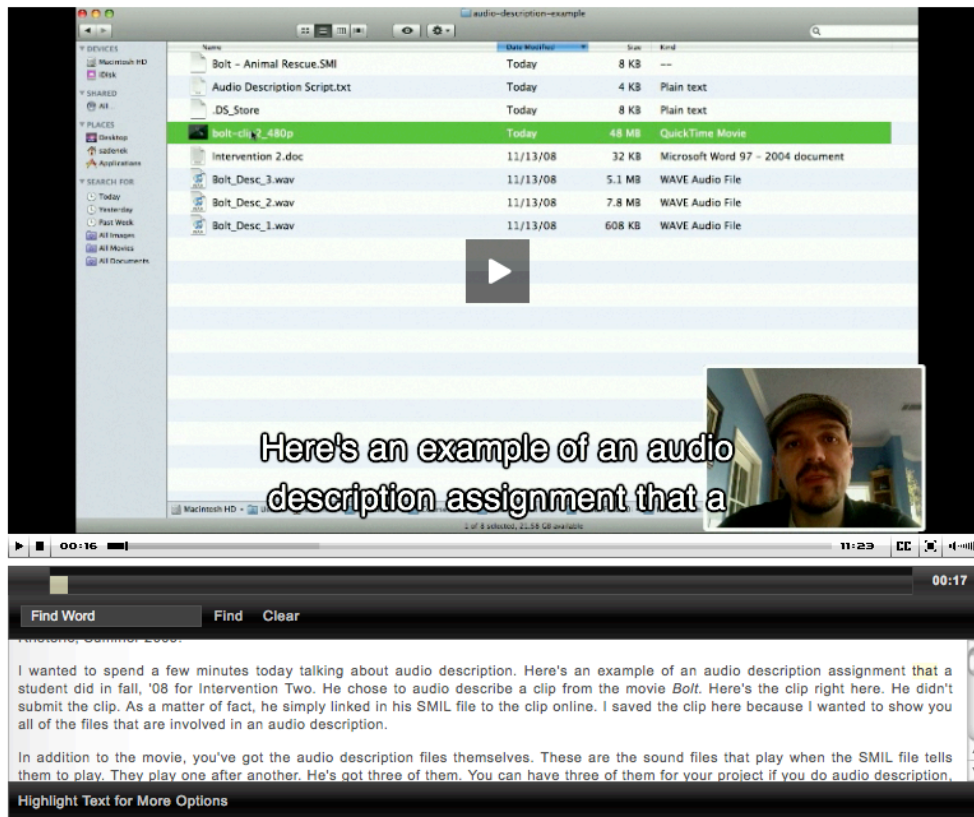
Figure 1. A Screen Grab from Course Screencast
Interactive Transcript Plugin ©2010, 3Play Media.com. Used with permission.

Interactive transcripts already provide users with an excellent way to search for and find information within a single video. As video becomes more popular and captioning technology provides a way to index large databases of video context, students will be able to search the video collection of an entire course, or even across all of the videos produced in all of the courses of a department, college, or university. In this future learning environment, captions will enable students to use keywords not only to find and review course content across multiple videos but also to insert their own "margin" notes, which could take the form of time-stamped text comments or pop-up idea bubbles (e.g. see BubblyPly.com), their own video responses or notes produced on the fly with their web cams, links to other related video moments in the course's video collection, links to external Web resources, and comments from other students that have been made public. This added content may or may not be searchable/captioned, but it would at least be tagged and easier to find as visible nodes in the student's personalized video stream. The instructor's lecture video would thus be transformed into the student's personalized study guide and an opportunity for collaborative learning. In addition, keyword searches would not simply return a list of matching video clips but also, perhaps, a single mash-up comprised of all the clips that satisfied the search query, plus any accompanying student commentary. The inherent limitations of uncaptioned video would thus be addressed by a robust video captioning and search system that allows students to personalize and reconfigure the content of a course according to their needs. The promise of universal design could be achieved, in other words, by an accessible system that levels the playing field for all students—deaf, hard of hearing, and hearing.

CONCLUSION

We need to continue to push for and applaud advances in caption technology that will leverage the power of searchable text to provide a more inclusive, more accessible learning environment for our students. While it is naïve to think that a fully accessible video library is cheap or easy to achieve—particularly at a time when some academic librarians are opting for cheap and inaccessible solutions like Netflix™ (Kaya 2010),[3] users are uploading uncaptioned, so-called "disposable" videos (Reid, 2008) by the millions each month to Facebook™ and YouTube™ (Bilton, 2010), and many others are simply unaware of either the need for or the benefits of captioning—it is nevertheless important for Web accessibility advocates to continue to publicize all of the reasons (ethical, legal, business, user-centered, etc.) that accessibility makes sense for our students and our pedagogies. I am optimistic about the ongoing efforts of Google™, Hulu™,[4] and others to address our pressing need for better solutions to the problem of video search, retrieval, analysis, and synthesis. As the number of distance learning, video-enriched courses grows on our college campuses, educators and students will require solutions that combine the richness of video with the data mining benefits of text-based captions and transcripts.

REFERENCES

3Play Media. Interactive transcript plugins. Retrieved from http://www.3playmedia.com/services/interactive-transcript-plugins/

Ballek, M. (2010, April 15). FACT: YouTube indexes captions. Retrieved from *vidiSEO* blog: http://vidiseo.com/youtube-indexes-captions/

Bilton, N. (2010, June 11). Mobile uploads spur Facebook video growth. *The New York Times*. Retrieved from http://bits.blogs.nytimes.com/2010/06/11/mobile-uploads-spurs-facebook-video-growth/

Chisholm, W., & May, M. (2009). *Universal design for Web applications*. Beijing: O'Reilly.

Chitu, A. (2010, June 4). YouTube's interactive transcripts. Retrieved from *Google operating system* blog: http://googlesystem.blogspot.com/2010/06/youtubes-interactive-transcripts.html

CaptionsOn (n.d.). http://www.captionson.com/watchVid.asp.

Evans, J. (2010, June 6). Website accommodation rules on the horizon. *Zanesville Times Recorder*. Retrieved from http://www.zanesvilletimesrecorder.com/article/20100606/OPINION02/6060308

Federal Communications Commission (2010). Part 79—Closed captioning of video programming: Current through February 19, 2010. Retrieved from http://www.fcc.gov/cgb/dro/captioning_regs.html

Hemstreet, T. (2010, April 6). Hearing loss number one diagnosis for military. Retrieved from http://www.northwestmilitary.com/news/veterans/2010/04/northwest-military-fort-lewis-ranger-newspaper-mcchord-airlifter-tyler-hemstreet-hearing-loss-vetera/

Hulu. (n.d.) Captions search. Retrieved from http://www.hulu.com/labs/captions-search

H.R. 3101. Twenty-first Century Communications and Video Accessibility Act of 2009. Retrieved from http://www.govtrack.us/congress/bill.xpd?bill=h111-3101

Kaya, T. (2010, September 18). Academic libraries add Netflix subscriptions. *The Chronicle of Higher Education*. Retrieved from http://chronicle.com/blogPost/Academic-Libraries-Add-Netflix/27018/

Mullins, A. (2009). Aimee Mullins and her 12 pairs of legs. Retrieved from http://www.ted.com/talks/lang/eng/aimee_mullins_prosthetic_aesthetics.html.

Netflix (2010, April 15). Subtitles now available for some titles for PC/Mac viewing. Retrieved from *The Netflix Blog* http://blog.netflix.com/2010/04/subtitles-now-available-for-some-titles.html

ProTranscript (n.d.). http://www.protranscript.com/.

Reid, A. (2008, May 15). Disposable video comments. Retrieved from the Digital digs blog: http://alexreid.typepad.com/digital_digs/2008/05/disposable-vide.html.

S. 3304. Equal Access to 21st Century Communications Act. Retrieved from http://www.govtrack.us/congress/bill.xpd?bill=s111-3304

Section 508 (1998). Amendment to the Rehabilitation Act of 1973. Retrieved from http://section508.gov/index.cfm?FuseAction=Content&ID=12

Sizemore, E. (2010). Google video test results: Captions vs description vs speech recognition. Retrieved from *Everett Sizemore: SEO Consultant* blog: http://www.esizemore.com/google-video-test-results-captions-vs-description-vs-speech-recognition/.

Stelter, B. (2010, June 21). On Web video, captions are coming slowly. *The New York Times*. Retrieved from http://www.nytimes.com/2010/06/21/business/media/21captions.html

TED Open Translation Project. Retrieved from http://www.ted.com/translate/about

U.S. Census Bureau (2008). National population projections. Retrieved from http://www.census.gov/population/www/projections/summarytables.html

Valentino-DeVries, J. (2010, September 29). New rules to make technology more accessible for blind, deaf. *The Wall Street Journal*. Retrieved from http://blogs.wsj.com/digits/2010/09/29/new-rules-to-make-technology-more-accessible-for-blind-deaf/

Zdenek, S. (2009, August 17). Caption watch: Hulu.com. Retrieved from Accessible Rhetoric blog: http://seanzdenek.com/?p=36

ADDITIONAL READING

AccessibleRhetoric.com

Captioning Tools. https://billcreswell.wordpress.com/other-caption-resources/captioning-tools/

Captioning Key. http://www.dcmp.org/captioningkey/index.html

How People with Disabilities Use the Web. http://www.w3.org/WAI/EO/Drafts/PWD-Use-Web/

Online Captioning Bibliography. http://joeclark.org/access/captioning/bpoc/bibliography.html

WebAIM.org

Web Accessibility Initiative. http://www.w3.org/WAI/

Web Content Accessibility Guidelines (WCAG 2.0). http://www.w3.org/TR/WCAG20/

KEY TERMS

Closed captioning: An accommodation that provides access to the audio portion of a video file for viewers who are deaf or hard of hearing. Closed captions are written transcripts of dialogue and other significant sounds that are synchronized to display on the screen as the sounds occur. Captions are either closed or open. Closed captions can be turned on or off. On TV in the U.S., closed captions are delivered through line 21 of the video data area. On the Web, closed captions are delivered in a separate text file and activated when the user selects the caption function/button on the video interface. Open captions can not be turned off. They are burned into the video itself. Captions are live or pre-recorded. Live captions are transcribed by a human stenocaptioner and delivered to viewers with a slight delay of 2-3 seconds usually. Captions are usually either one of two types: scroll-up or pop-on.

Interactive transcripts: Written transcripts of the audio portion of a video file. The transcripts are searchable and clickable, so users can interact with the transcript. By clicking on a word or sentence in the transcript, the user is transported to that moment in the video where that word is spoken. Vendors include 3Play Media and ProTranscript. YouTube also supports interactive transcripts.

Web accessibility: The practice of making Web pages and applications accessible to the largest number of users possible, especially users with disabilities. Web accessibility focuses on the major categories of disability: visual, auditory, motor, and cognitive. U.S. laws such as Section 508 and international guidelines such as Web Content Accessibility Guidelines (WCAG 2.0) allow web developers to adhere to a set of accessibility guidelines and best practices.

DISCUSSION QUESTIONS

1. Ask participants to share their experiences with video used for educational purposes (e.g. video lectures, screencasts, YouTube videos in the classroom). Is video becoming a more popular means of delivering course content, especially in distance learning contexts? What are some of the challenges of using video in the classroom? How do video and audio differ from writing?

These questions are intended to encourage participants to reflect on some of the challenges involved in searching and annotating uncaptioned video. For students who are deaf and hard of hearing, uncaptioned video is inaccessible. But uncaptioned video can also be difficult for hearing students to use. For example, consider how difficult it can be for students to search and retrieve a specific piece of information from a long, uncaptioned lecture video, or to use a professor's uncaptioned video as a study guide prior to an exam.

2. If participants have access to a computer lab, ask them to visit TED.com and view one of the TED videos that is accompanied by an interactive transcript. The talk by Aimee Mullins (see the References to this chapter) is an excellent example. How might interactive transcripts be used with video in the classroom? What are some of the advantages of interactive transcripts over uncaptioned videos? What are some of the drawbacks and challenges for instructors? How might interactive transcripts be coupled with search technology to give students integrated access to entire collections of videos?

7

These questions are intended to encourage participants to consider closed captioning in the wider context of universal design.

ENDNOTES

1. "According to the Deafness Research Foundation, hearing loss is the No. 1 diagnosis for U.S. soldiers in Afghanistan and more than 65 percent of Afghan war veterans are suffering from hearing damage" (Hemstreet, 2010).

2. "One third of all senior citizens have hearing problems" (CaptionsOn 2010). Moreover, the number of Americans 65 years of age and older—a population group more likely to benefit from accommodations such as closed captioning—is projected to rise from 13% in 2010 to 20% by 2050 (U.S. Census, 2008).

3. Netflix™ has only recently started offering streaming movies and TV shows with closed captions (see Netflix, 2010). But their closed captioned streaming library is small (only about 100 titles out of thousands). In addition, because Netflix™ provides no way to search for closed captioned content, Netflix's™ streaming service is at this time virtually inaccessible to users who require or prefer captions.

4. While the number of closed captioned full episodes and movies on Hulu™ remains small at approximately 4 to 6% (see Zdenek, 2009), Hulu™ has taken steps to exploit the power of captioned media to provide more fine-grained search experiences. Search results match specific time-stamped moments in captioned episodes. In addition, users are able to see visually on a "Heat Map" graph "the parts of the video that have been viewed the most; you can also click on the chart to navigate to any point within the captions" (Hulu). According to Eric Feng, the chief technical officer at Hulu™, captions have "turned into a very important part of our user experience" (quoted in Stelter, 2010).