Accessibility through Captioning

Cindy Camp
cindy.camp@pepnet.org
Pepnet 2 Mission Statement

Our mission is to improve postsecondary outcomes for individuals who are deaf or hard of hearing, including those with co-occurring disabilities, through provision of evidence-based resources and services.

www.pepnet.org
Probably the main reason you are attending this webinar is because you are concerned about providing legally mandated access to D/HH students on your campus. However it is important to realize that captions aren’t just for those with a hearing loss. They benefit everyone. Let’s look at a few of these additional benefits.

- Makes your video (and advertising) accessible to over 28 million Americans who are deaf or hard of hearing. Helps to meet regulatory compliance measures for governmental and educational institutions (e.g., “Section 508”).

- Promotes literacy for children and adults by strengthening reading speed, comprehension, spelling, and grammar skills.

- Improves clarity and comprehension of the sounds and dialogue, particularly when poor audio, heavy accents, background noises, and other such media elements are present.

- Allows individuals to follow along with the video dialogue even when they do not have access to sound on their PC or mobile device (e.g., in a noisy area with no headphones or when using devices with faulty or missing sound cards or drivers).

- Creates a complete text transcript of the video, which can be indexed by search engines to provide far more accurate search results than keyword tagging alone.
• Enables translation into multiple languages.
I also think it is important to understand what fully accessible captions are. It may seem like a no-brainer but as I have come to realize as I provide training on the topic that not everyone has the same understanding of what captions are and should be.
Now that we have a definition of captions let’s look at what that really means.

If you are just putting “words on the screen” then that does not provide full access to the content. How you put those words on the screen is critical.

I’d also like to clarify that we are talking about post production captioning. This means captions added to a recorded video. We are not talking about real-time captioning such as you are seeing on screen today. The rules and standards for each is different. For example the real-time captions will be a bit behind what I am saying since our captioner is not psychic. There could also be mistakes since even wonderful captioners such as were are using aren’t perfect. However, captions created for a video should not contain time lag or errors since there is time to fix all mistakes.

Next we are going to watch a video clip with captions. Please pay close attention to the quality and accuracy of the captions. Imagine if you were asked to take a test on the material being presented.
A little later we are going to talk about standards for accessible captions but for now let’s just look at some of the glaring errors in this clip.

First the timing of the captions does not match the audio. This can be confusing to an individual who is hard of hearing and trying to use the captions as support for what they can hear. It is also confusing when the captions don’t match the action on the screen. This means a deaf individual would either be behind the action or ahead. Either way can be confusing and frustrating. With post production captions there is no reason why the captions should not be perfectly time synced with the audio.

Next let’s look at some of the horrific spelling errors such as: sign tifik for scientific and knee gross for Negros

There is no speaker identification so we don’t know when the speaker changes from the narrator to Carver reading a letter he wrote. To a person who can’t hear the voice change it appears as if one person is saying all of this.

Also notice that each word appeared on screen one at a time and as the last row ended the text would scroll up. This is a lot of unnecessary movement that can cause eye fatigue.
Now that we have seen an example of poor captions it will be easier to recognize the importance of quality captions. One thing I’ve learned is that because we as hearing people don’t depend on captions for access we too often overlook errors. I’ve had some administrators view the clip I just showed and say, “the captions weren’t great but they were adequate.” I would disagree. What happens is that sometimes people don’t really pay attention to the captions themselves. They simply see words appear and think that makes the video accessible.

However, the quality of captions is critical and we need to make sure that we are paying attention to what word appear and how they appear. If you don’t have high quality captions then you are throwing away your time and money. Captions should be time synced with the video. Look at the first example in the top left. Obviously the caption does not match the action.

In the second example the text is obviously wrong. The word which came up as “fertile” should have been “first.” While it is amusing to see such mistakes you probably wouldn’t be amused if you were a student in a class being tested on the material being shown.
Let’s take a look at the basic process of captioning.

What’s Involved?

1. Obtain permission from the copyright holder to add captions
2. Create a verbatim transcript
3. Import transcript into captioning software
4. Create grammatically correct line breaks
5. Time code the captions
6. Export the captioned file(s) to the final destination
Copyright vs. Fair Use

- Copyrighted media should not be captioned without permission.
- “Fair use” does allow reproduction of copyrighted works for nonprofit educational use within limits.
- Report of the Advisory Commission on Accessible Instructional Materials In Postsecondary Education of Students with Disabilities (Full Report)
  – http://www2.ed.gov/about/bdscomm/list/aim/meeting/aim-report.pdf

However, “fair use” limits include:
- How much of a work can be reproduced.
- How long the copies can be kept.
- Alteration of the work.
Transcript Standards

- Captions are only as good as the transcript.
- A transcript should be verbatim.
- It should include speaker identification.
- Spelling, capitalization, and grammar must be correct.
- Punctuation follows standard rules, but also special rules unique to captioning.
- All essential sound effects are included, either in words or symbols (e.g., “buzz” or 🎵).
Now let’s look at guidelines for creating high quality captions. These guidelines come from the Described and Captioned Media Program or DCMP. They are a federally funded project and the only group to have researched and developed captioning standards. You can access a copy of the Captioning Key via the URL listed on your screen: http://www.dcmp.org/captioningkey
Notice that these captions use both upper and lower case letters. They also use proportional spacing of letters. It is easier to read, because it is what we are used to seeing in books and newspapers. Using all upper case letters is very difficult to read because we have to look at each individual letter to figure out the word. When upper and lower case letters are used the brain more easily recognizes the shape of the word.

Also notice that when each person is speaking the captions move to their side of the screen. This is one way to clearly show who is speaking.

Sound effects are indicated such as “echoing” in brackets, and “upbeat hip-hop music” and the music note graphic

During the song portion the speaker is identified in parenthesis as man or woman.
Now that you understand the benefits of captioning and the basics involved it is time to decide how you will deal with captions at your institution. Your two basic choices are to pay a professional company or to purchase equipment to create captions yourself. The first consideration most administrator look at is cost. Of course expense is important to us all. However, I challenge you to just as strongly consider quality. If captions are not high quality, they are not providing truly equal access, thus they are a waste of money if done “cheaply.”

When considering cost it is important to look at all factors. Purchasing high end captioning software and equipment can run from $5,000 to $10,000 or more. In theory you could get a one time grant and have the equipment you need. However, this does not take into account the staffing needs. Some institutions try using student workers but the reality is that once they are trained to do the job efficiently they move on. Captioning is not a simple or easy process. The more you do it the more comfortable you are with the process and the efficient you’ll be. So if you are willing to hire full time or part time staff who can be devoted to captioning then you can establish good in house captions. However, if you have inconsistent staff and they only work with the equipment infrequently they are going to waste a lot of time relearning the process each time. It can take 7 to 10 hours to caption a 30 minute video start to finish.

Outsourcing captioning may seem expensive in the beginning. Companies charge between $8 to $28 per minute to caption a video. But it could be the best way for you to get high quality captions without a lot of personal time and hassle involved.
What type of media do you need captioned?
   Digital: online courses, online promo videos
   DVD or CD
   VHS

How much media do you need captioned?
How often will you need to caption in the future?
What is your budget?
   Money and Time

Ideally setting up for in house captioning would involve creating a department with full time staff to work on captions. Trying to use existing staff and part time student workers will probably mean that workers are soon overwhelmed and the quality of the captions is not at an appropriate level.
Do your research. Not all companies are equal. Also remember your decision should be based on more than just the cost involved. In fact some companies will offer you lower prices for lower quality captions.

What formats do they accept?  
What formats can you get back?  
What is the turn around time?  
How is a transcript created?  
How is the time coding done?  
How are line breaks decided?  
Is there quality control?  
Can corrections be made?
If you pay for low quality captions which don’t provide true access then you have wasted your money.
Several years ago I did a comparison of about 25 different captioning agencies.

These are two quotes I got during those phone calls.

I was shocked that the sales person was trying to get my business by telling me that they cut corners and that was how I was going to get a good price. But it also made me realize that many institution may think that way. If they don’t truly understand captions then cutting corners may not seem so bad. I would hope that after today’s presentation you’ll be able to educate your colleagues about just how important quality captions are.

So what’s the difference?

• “Educational videos don’t need to be works of art. You just need words on the screen.”
• “We don’t check spellings for anything. For proper names, we spell phonetically as best we can. After all, if a hearing kid hears the name of a foreign river, they don’t know how it is spelled either.”
Another popular solution to captioning is voice recognition software. Some captioning agencies use voice recognition software to create transcripts and some institutions are trying to use voice recognition software to create transcripts in house. Before going either route you need to understand what the software is capable of doing and what it isn’t.

According to:  
• Speech Recognition results are inaccurate, unreliable, and inconsistent – with error rates often exceeding 20%.  
• 80 – 90% accuracy is simply inadequate for most applications.  
• Accuracy below 97% is insufficient when comprehension is crucial.  
• Transcripts with high error rates do not satisfy compliance with legal mandates.  
• It costs more to edit speech recognition scripts than it does to use a professional transcriber to start with.

This does not mean speech recognition can’t be helpful but it cannot create an accurate transcript without editing. Currently there is not cheap and easy way to create fully accessible and accurate captions. Of course technology is always advancing. At some point in the future captions may be as easy as pushing a CC button. But for now we cannot assume that having words on the screen means equal access.
Another use of automatic speech recognition software is with YouTube. I have heard of some institutions trying to upload all their video content to YouTube and let YouTube provide the captions automatically. Let’s talk about why that is not a good idea.

First let me say that I applaud YouTube for recognizing that there is a need for captions and for trying to be part of the solution. However, their automated captions have become a bit of a joke as demonstrated by this video by Rhett and Link. These two men created a series of videos called “Caption Fail.” In these videos they upload a video conversation and allow YouTube to automatically caption it. Then they film the same clip again using the transcript produced by YouTube. In this example they uploaded popular Christmas carols and we can see the YouTube version.

This carol is “Let it Snow.” The text should be “although the weather outside is frightful.” However, YouTube created the caption “although they’re outside is frightful.” This example may not seem so bad since only 2 words were wrong. But that is 1/3 of the total content.
In this example the errors are much more clear. “Chestnuts roasting on an open fire,” became “chess match rusting envelope inside.” This time 0% of the content is correct.

The problem is not necessarily that the caption are inaccurate but that not everyone recognizes that automatic speech recognition is not perfect. I have heard of instructors showing YouTube videos in class without previewing the automatic captions. They simply assumed that because there was a CC button that the video would be accessible. While the entire class got a good laugh, the desired learning experience did not take place.

I also know of institutions who have tried to use YouTube captions as a shortcut to creating their own captions. They would upload class lecture and then use the automatic captions to create a transcript which they could then correct. Almost unanimously these institutions found it more time consuming to try and “fix” the errors than if they had created the transcript from scratch.
The most important thing to remember about captions is that you need a policy and procedure in place.

Educational institutions are mandated by the ADA and Section 504 to provide access to educational content. This means that students who are deaf or hard of hearing must have access to video/audio content for face to face and online classes. The only way to provide equal access to this content is time synced captions. In the past many institutions adopted a wait and see approach, meaning they would wait until there was a request for an accommodations and then try to put a plan in place to provide at least minimal access. This approach usually leads to delayed and/or subpar accessibility.

Recent OCR and court decisions encourage institutions to be proactive and to develop policies that are practical. For example stating that each instructor is responsible for captioning media shown in their classroom, without providing institutional support or training, is not an effective policy. Some institutions have stated such policies as a way to pass blame or deny institutional responsibility. However, the courts have upheld that accessibility is not the sole responsibility of an instructor or even the disability services department. Thus the institution as a whole is held accountable for providing student access.

But let’s not view captions as something we have to do; which are expensive and only benefit a small group of students. Instead view captions as part of Universal Design. They enhance learning of all students because the brain retains information better when it is
received through multiple senses, thus hearing the audio and seeing the captions is better than just hearing. Captions provide access when there is poor audio quality or strong accents. Captions provide a transcript which makes a video searchable and can increase hits on your website. Captions benefit everyone.
Resources

- Caption it Yourself
  - http://dcmp.org/public_content/ai/ciy/
- Captioning Service Vendors
- Collaborative for Communication Access via Captioning
- "YouTube Ready" Qualified Captioning Vendors
  - http://youtubeready.dc.mp.org/
- DCMP Labs
  - http://labs.dc.mp.org/
Questions
pepnet 2 receives support from:

pepnet 2 is funded by the Research to Practice Division, Office of Special Education Programs and the US Department of Education via Cooperative Agreement #H326D110003

www.pepnet.org