

## 3Play Webinars | WBNR 11-14-2019 QSC

---

SAMANTHA SAULD: So, thanks for joining this webinar entitled Quick Start to Captioning. I'm Samantha Sauld from 3Play Media, and I'll be presenting today.

Today, we'll cover the following topics. The first question we'll cover is, what are captions? Followed by-- how do you create captions? Where do you publish them? Why you should caption. And then finally, we'll end off with who 3Play Media is, and then the Q&A.

So, what are captions? Closed captions are time-synchronized text that can be read while watching a video and are usually noted with a CC icon.

Captions originated as an FCC mandate in the 1980s. But the use has expanded to online video and internet applications.

Captions assume the viewer can't hear. So they include relevant sound effects, speaker identifications, and other non-speech elements to make it easier for the viewer to understand who's speaking.

An example of this additional information would be if someone is opening the door and you can visually see their keys jingling, you wouldn't need to caption that. But if the keys are jingling off-screen, you would include that non-speech element.

While closed captions are used for pre-recorded video, live captioning is for events happening in real time. For example, in this webinar or a meeting, fitness classes, and even conferences.

Live captions ensure that all your live events are accessible to deaf or hard-of-hearing individuals, as well as making sure that your content is more engaging. Live captions are usually created by an automatic software or a stenographer. In this webinar, we use a live stenographer. There might be slight delays in live captioning as the computer is processing the words or the stenographer is typing.

It's important to distinguish between captions, subtitles, and transcripts, as they all mean something different. Captions assume the viewer can't hear the audio, they are time-synchronized, and they include

relevant sound effects. You can spot if a video has captions when you see a CC icon.

Subtitles, on the other hand, assume the viewer can hear but can't understand the audio. Their purpose is to translate the audio. So like captions, they are also time-synchronized. And then transcripts are a plain-text version of the audio. It's not time-synchronized. And it's good for audio-only content.

In the US, the distinction between captions and subtitles is important. But in other countries, like countries in Europe, these terms are used synonymously.

So, how do you create captions? There are a few ways to create captions. You can either do it yourself, use automatic speech recognition, or ASR, or you can use a captioning vendor.

So one way, if you have the time, is to manually transcribe the video yourself. You'll need plenty of time, as this method can take five to six times longer than the length of the video. And this doesn't include the time it takes to add non-speech elements. So this method could be a very costly method at scale.

The second way to caption your video is starting with ASR. You can use YouTube's automatic caption generator to create captions. YouTube uses speech recognition technology that aligns your transcript with the audio and breaks it up into time caption frames.

It does a pretty decent job, especially with high-quality audio and clearly-spoken English. And YouTube even lets you export the caption file for use in other applications. So although accuracy is a big issue when using automatic captions, YouTube does allow you to go back and edit the script for accuracy.

And like I said, also you can outsource to a vendor like 3Play Media. Our process combines technology with human labor in a really careful and strategic way so that we have the best of both worlds.

So for our caption process, first the file goes through ASR, which produces a rough transcript. Then one of our professional editors, our transcriptionists with all different areas of expertise, goes through and corrects the file. And then finally, we have a third round of QA review where a QA manager will conduct a final review of the transcript and captions to ensure we meet our guaranteed 99% accuracy rate.

Now let's talk a little bit about caption quality. So when it comes to captioning, it's important to follow the best practices for captioning quality. The industry standard for spelling is a 99% accuracy rate. 99% accuracy, though close to perfection, means that there's still a 1% chance of error.

In a 10 minute file with 1,500 words, this leniency allows for about 15 errors in total. Now, if your video is scripted content, then you'll want to ensure your captions are by verbatim. So for broadcast, you'll want to include the "um's" and the "uh's" because it's scripted. For lectures or live captioning, a clean read is preferable, meaning you'll want to eliminate any filler words.

Now, each caption frame should be around one to three lines with 32 characters per line. The best font to use is non-serif. You should also ensure they're time-synchronized and last a minimum of a second on the screen so that it gives your viewers enough time to read.

Another key thing to keep in mind is caption placement. Typically, captions are placed in the lower-center part of the screen but should be moved when they are in the way of other important text or elements in the video. And as for silent bits, you want to make sure your captions go away when there's a pause or silence so they don't confuse the viewer or hang on for too long.

When you use ASR technology, the accuracy rates are pretty abysmal. A lot of ASR errors make sense acoustically but not linguistically. Here, I'll show you an example of a transcript captured by ASR. Just listen closely to the audio and compare with the words on screen, and see if

you catch the errors. You can type any errors you notice in the chat window.

[AUDIO PLAYBACK]

- One of the most challenging aspects of choosing a career is simply determining where our interests lie. Now, one common characteristic we saw in the majority of people we interviewed was a powerful connection with a childhood interest.

- For me, part of the reason why I work here is when I was five years old growing up in Boston, I went to the New England Aquarium. And I picked up a horseshoe crab. And I touched a horseshoe crab. And I still remember that. And I've still-- you know, I love those types of engaging experiences that really register with you and stick with you.

- As a child, my grandfather was a forester. And my childhood playground was 3,600 acres of trees and wildlife that he had introduced there.

[END PLAYBACK]

**SAMANTHA SAULD:**

So one of the issues that I noticed right away is a lack of punctuation. In this transcript, there are few periods and incorrect capitalizations. Another issue is that hesitation words are not removed, which spill over into other words and make it inaccurate. A lot of things like speaker IDs are also not captioned.

So some of you guys said in the comment section that New England Aquarium was off. Yeah, it was off. So yeah, nice catch.

I also noticed, too, that it said new wing of the Koran, and then forester-- the speaker said forester. But in the transcript, it read four stories. So a human wouldn't make any of these errors.

So, most live captioning accuracy is over 90% accurate. But noise level, accents, or connectivity issues can affect the accuracy.

Much of the same quality standards for closed captioning apply to web captioning as well. For the best accuracy in live captioning, you'll want to have a strong network connection, a good audio quality, little to no background noise, a single speaker, and clear speech and pronunciation.

Now let's dive a little bit into how to publish captions. So there are a number of ways to publish your captions. The most common is through a side car file, which is basically a file that stores the captions so they can be associated with the corresponding video.

So when you upload your caption file on YouTube, you're uploading it through a side car file. And these types of captions give the user control to turn the captions on or off.

Another way is to encode captions onto the video. For example, these are found on kiosks or offline videos, and could also be turned off or on. Open captions are burned into a video and cannot be turned off or on. For social media videos on Instagram or Twitter, adding captions as a side car file is not possible. Open caption encoding is one way to overcome this barrier and keep your videos engaging and accessible.

And then lastly, you have integration. So integrations are simply a publishing process for captions. It's a preset workflow between your captioning process and video publishing process to really just make everything more streamlined.

So, why should you caption? The number one reason why you should caption is for accessibility. There are 48 million Americans with hearing

loss, which is about 20% of the US population. And then there are 360 million deaf or hard-of-hearing individuals around the world. And captions just help make your content accessible to them.

So you should also caption for engagement. Captions are important in order to make your video comprehensible without sound. So we found this stat that's 41% of videos are incomprehensible without sound or caption, which means that if someone doesn't have headphones, they won't watch your video.

And you may be wondering, are people even watching videos with the sound off? Well, they actually do. Facebook uncovered that 85% of Facebook videos are watched with the sound off. So if your video relies heavily on sound, a lot of people are probably scrolling past them.

Video accessibility has tremendous benefits for improving SEO, the user experience, your reach, and your brand overall. A study by Liveclicker found that pages with transcripts earned an average of 16% more revenue than they did before transcripts were added. And according to Facebook, videos with captions have 135% greater organic search traffic. And a research study from the *Journal of the Academy of Marketing Science* found captions helped improve brand recall, verbal memory, and behavioral intent.

So we conducted a nationwide study with Oregon State University. And we surveyed students to see how and why they use captions. And the results prove that captions truly help students learn.

98.6% of students found captions helpful. In addition, 65% of students use captions to help them focus. And 75% of all students who use captions-- not just those who are deaf or hard-of-hearing-- use captions as a learning aid.

So, what are the laws that require captions? There are a number of accessibility laws in this country. The first one is the Rehabilitation Act of 1973. And that was the first major accessibility law in the US.

It has two sections which specifically impact video accessibility. Section 504 is a broad anti-discrimination law that requires equal access for individuals with disabilities, which applies to federal and federally-funded programs.

Section 508 requires federal communications and information technology to be made accessible. And Section 508 references web content accessibility guidelines, otherwise known as WCAG. So what's unique about the Rehabilitation Act is that closed captioning and audio description requirements are written directly into Section 508.

And then we have the Americans with Disabilities Act, which was the second major accessibility law in the US. It has two sections that impact video accessibility. Title II applies to public entities. And Title III applies to places of public accommodation, including private organizations that provide public accommodation, basically-- so, a doctor's office, a library, a hotel, a restaurant, and many other places.

The context of a place of a public accommodation has been tried in many lawsuits in regards to how it impacts internet-only businesses. And in several cases, Title III has been extended to the online spaces as well. For example, there were suits against Netflix both in regards of closed captioning and audio description. And in both cases, the outcome was that Netflix had to provide accurate captions for their streaming shows and audio description for their Netflix originals.

The third major accessibility law in the US is the 21st Century Communications and Video Accessibility Act, or the CVAA. For caption requirements, the CVAA applies specifically to online video, but has previously been aired on television. And any online video that appeared on TV with captions has to be captioned when it goes online, including video clips and trailers. And as for audio description, the CVAA is phasing in audio description requirements by 2020.

And one other thing to know are the standards that should be met to mitigate the risk of legal action. WCAG is the international standard

and best practice for accessibility. It's important to note that there is a WCAG 2.1. But for the time being, WCAG 2.0 is currently what's referred in lawsuits and legal recommendations.

So, WCAG basically is a standard to make digital content accessible for all users, specifically users with disabilities. It outlines best practices for making web content universally perceivable, operable, understandable, and robust.

So WCAG has three levels of compliance. Level A is the easiest level to maintain. Then we have level AA, which is what most people are aiming for. And this is the mid-level of standards. And then we have level AAA, which is the most comprehensive and the highest accessibility standard.

Most laws and lawsuits mentioned WCAG 2.0 compliance, which I mentioned. So for now, that is what's legally required. Only if a law explicitly states that web developers have to adapt to the newest WCAG version do you need to make your content WCAG 2.1-compliant.

The W3C does suggest that any new website should be created following WCAG 2.0 guidelines since they are more inclusive and more mobile-friendly. And to be compliant with WCAG, you are required to caption pre-recorded video for level A compliance and caption live video for level AA compliance.

So in the lawsuit, *The National Association of the Deaf versus Harvard and MIT*, the universities were sued for failing to caption and for having unintelligible captions on a lot of their online courses. This was the first time that accuracy had been considered in legal ramifications for closed captioning. And by the way, these were YouTube captions. So this lawsuit represents a violation of Title III of the Americans with Disabilities Act and has extended the requirements to the internet. The outcome will have huge implications for anyone that's in higher education.

Also, Section 504 of the Rehabilitation Act was violated for failing to provide adequate closed captioning for online video. So UC Berkeley, Penn State, and Miami University are all examples of schools that have been sued or have entered into consent decrees in regard to inaccessible video.

*NAD versus Netflix* was the first web accessibility lawsuit. So under Title III, the court ruled that Netflix is considered a place of public accommodation, and therefore it needed to make their content accessible.

Now we'll talk a little bit about who 3Play Media is. We are video accessibility company and spun out of MIT in 2007, and are currently based in Boston.

We started out offering captioning, transcription, and subtitling services. We also offer audio description now, which is a service for blind and low-vision individuals. And we recently released a live automatic captioning solution as well.

We have over 2,500 customers spanning higher education, media, government, e-commerce, fitness associations, and enterprise. And our goal is really to just make the whole captioning and video accessibility process much easier. We also have a number of turnaround and workflow options to choose from.

Our number one goal is to make video accessible and make that whole process easy. And we do that in a number of ways. We have an easy-to-use online account system where you can manage everything easily from one place. We have a number of different options for turnaround, anywhere from a couple of hours to over a week, whatever fits your needs.

We have different video search plugins and integrations for captioning and description that help simplify the process of creating accessible video. And then, of course, we have the 3Play plugin. And then what we're really working toward is being a one-stop shop for captioning,

description, transcription, and subtitling, and video accessibility as a whole.

And besides our services, we also offer tons of free resources all centered around online video and video accessibility. We post weekly blogs. We have free white papers, how-tos, and checklists. We have research studies on the impact of video accessibility in certain environments like education.

We offer monthly webinars like this one, sometimes several webinars a month where we bring in an accessibility expert to share their knowledge. And we will soon be offering a free video accessibility certification. That will be available online this winter.

We do have a page where you can sign up for updates. And if you're interested, we'd be more than happy to share the link with you. You can visit [3PlayMedia.com/Certification](https://3PlayMedia.com/Certification).

Now we'll begin Q&A. Let me look at the first question that we have so far. So, the first question that we have is, do we need both captions and the transcript?

So, captions and transcripts are two different things. Transcripts are not time-synchronized. So they're preferable for audio-only content. But if you provided the transcript, it doesn't necessarily take the place of captions. So it's sometimes possible that you will need both. And you can see from the research that I mentioned that there are definitely benefits for each of these different types of tools.

The next question is, how can we sign up for the certification? You can go to the link [www.3PlayMedia.com/certification](https://www.3PlayMedia.com/certification). And you can just sign up there. And you'll get updates on when it's going to roll out. Thanks for the question.

The next question is, how do you ensure accurate captions on more complex or challenging content? Good question. Our process includes two rounds of human editing. So no matter what the content is, we're

getting back at least a 99% accurate file.

Our editors choose a file to work on from the marketplace, which allows them to choose content they know really well. Our editors come from very diverse backgrounds, including STEM fields and medical fields, things like that. And we also allow you to upload cheat sheets, which are basically a glossary where you can upload specific terminology or any names you think the editor may not be familiar with.

The next question is, can you tell us more about your live captioning solution? Absolutely. So last week, we released an automatic live captioning solution that integrates with top live streaming video and meeting platforms like YouTube, Zoom, Brightcove and JW Player. The process varies depending on the platform. But essentially, you schedule your live captions via the 3Play platform. And then once you begin streaming, your captions will appear right in the player.

And right now, we're also giving 10 free hours of live automatic captioning. And you can get more information by emailing [livecaptioning@3playmedia.com](mailto:livecaptioning@3playmedia.com).

And the next question asked, how does an integration work?

Integrations link disparate systems or platforms to make it easier to share information and build workflows between the two.

Our integrations are engineered to make the captioning process a lot easier. So we integrate with most leading video platforms. And our integrations allow you to select the files you want captioned directly from your video platform or cloud storage. And basically, integration saves you a ton of time by streamlining the whole captioning process.

The next question is, how do you handle the fact that there is a large number of captioning formats? So we offer over 50 formats. And you're able to download them directly from the account system.

The next question is, are there quality standards for how captions should look? Yes, there are. And with accuracy, the FCC states that

captions must match the spoken words and the audio to the fullest extent. And this includes preserving any slang or accents in the content, and adding the non-speech elements.

For live captioning, some leniency does apply. But captions must be synchronized, they must align with the audio track, and each caption frame should be presented at a readable speed of about three to seven seconds on the screen.

Completeness is also important. Captions must run from the beginning to the end of the program and not drop off at any time. And captions should be placed so that they do not block any important visual content.

The next question is, how do you get access to the 3Play plugin? It's free for any 3Play Pro or Enterprise customers. So it just comes right with it. So as soon as you sign up to be a player customer, you have access to the 3Play plugin.

The next question is, how do you get started with 3Play? You can feel free to contact us at [sales@3playmedia.com](mailto:sales@3playmedia.com), and we'd be happy to get you started.

The next question is, are audio descriptions now mandatory in addition to video captioning to present 508 compliance for federal organizations? According to Section 508, you need to be compliant with WCAG 2.0 level AA. And audio description for recorded video is a level A requirement.

OK. So the next question we have is, can we see a demo on how to create closed captions or external videos such as YouTube videos by your tool? Yes, you can. So the best way to schedule a demo is to contact sales. And you can reach our sales team at [sales@3playmedia.com](mailto:sales@3playmedia.com)

So we have time for a few more questions. So, someone asks, relating to a previous question, we are actually supposed to include the ums

and the-- so basically, any of the filler words. Yes. But this is only for scripted content. So you want to make sure that it's by verbatim. But if you have a live broadcast where someone is saying "um" or "uh," then you don't want to include that. A clean read is preferable in those instances.

So the next question is, can automatic captions be edited later for accuracy? Yes, it can be edited for later. You can edit it yourself. Or you can use a third-party vendor like 3Play where you can submit your file. And then we can edit later for accuracy as well.

Thanks, everyone, for joining.