

3Play Webinars | Live Captioning with 3Play Media

SOFIA LEIVA: Well, thank you everyone for joining us today for the webinar entitled "Live Captioning with 3Play Media." My name is Sofia Leiva, and I'm part of our marketing team. I'll be presenting today. But I'm also joined by my colleague Ryan Martinez, who's part of our implementation team. And he'll be here to help answer Q&A questions and any that you have around 3Play Media or live captioning.

On today's agenda, we are going to talk about the basics of live captioning, from live automatic captioning to a CART solution. Then we'll briefly talk about who is 3Play Media and upcoming events and content that we have for you and then [INAUDIBLE] Q&A.

So let's start by talking about what is live captioning. Live captioning is obviously going to be different than your typical closed captioning because it's happening in real time for live events such as this or meetings that you may have during work. It's typically done by an automatic software or by a stenographer CART solution, so a human. And an example of that would be for this webinar, we're currently using a human. You can expect a slight latency with live captioners or with live captioning because of the computer processing, what is being said, or the stenographer typing.

Some important terms to know when it comes to live captioning is CART, which stands for Computer Assisted Real-time Translation. And essentially, this means that there's one, typically remotely, typing out the captions. And then you'll [INAUDIBLE] live automatic or automatic speech recognition, and this is an actual software that's translating it. So think of Google captions on slides or YouTube captioning.

It's important to talk about caption quality when it comes to live captioning, just because it's going to be different than the quality that you have for closed captions. So typically with closed captions, you can expect a 90% accuracy rate. With live captions, it'd be a little bit lower, and it depends on whether you're using an automatic solution or a human and then the environment that the event is happening in. And the accuracy can range from 80% to 90%.

I mentioned earlier the latency. So the latency in closed captioning-- that you see the captions as they're being spoken. They're time-coded to the audio. In live captioning, the latency can range from four seconds to 11 seconds, when this gives time for the software to [AUDIO OUT] up or the stenographer to finish typing what is being said. This is totally normal. And the longer the latency, typically, the more time there is for an accurate description, but it can also range.

And then, unlike closed captioning, there isn't a governing body for live captioning. There aren't any laws that specifically mention, although there's guidelines that do mention it. And there isn't really anything that says, oh, it has to be 80% accurate, or you need to use a solution that's 95% accurate.

Some best practices around live captioning-- so number one, you want to make sure you have a strong network connection, just so that it doesn't cut off or you lose connection with your captioner or captioning solution. You want to make sure that there is good quality audio. And typically, we recommend using a microphone. A good microphone will cost you around \$50. And at 3Play, we use the [INAUDIBLE] microphone, I highly recommend it.

You also want to make sure that there's little to no background noise. So if you're in a room, you can hang blankets to sort of get rid of an echo. Or just make sure that you're in a room that is sealed, and you don't hear anything going on. That will ensure that the captioner or captions can capture what is being said.

And then you want to make sure that there is a single speaker. So if you have a panel, make [AUDIO OUT] only one person is speaking, and they have clear speech and pronunciation.

There's various ways for you to stream your live captions. So a lot of video platforms now allow you to stream them natively. But if that's not an option, you can [INAUDIBLE] use a URL that links to an external caption place where the people can follow along with the captions as it's being presented.

We typically get are live captions good enough? And the answer is typically no, if you are planning on publishing the recording afterwards. Number one, because of that latency, the captions aren't time-coded to the video, so the captions won't match up when the speaker is speaking. So you would need to go back and set the appropriate time codes.

And then, number two, it's like I mentioned earlier, live caption accuracy can range from 80% to 95%. And typically, we want to aim for a 99% accuracy when it comes to closed captions or captions for prerecorded video. That just really ensures they're accessible, and people can comprehend them, and there's no errors that can distract your viewer.

If you want to learn more about the state of automatic speech recognition or automatic captions, I will link to our state of automatic speech recognition report, which will dive into the findings of different ASR solutions that are out there in the market and compare the accuracy so that you can make an educated decision on which to use.

And then the last part of whether captions are good enough is that you always need to have the human element to go back in and edit the final transcript, again, for the accuracy, part and also because it's really important for the quality of the captions, SEO, and branding. And we'll talk a little bit about that next.

So there's a lot of benefits to captioning, whether you do it live or closed captioning for recorded video, and number one is accessibility. This is one that you hear most often, because captions were created as an accessibility accommodation. In America, there 48 million Americans with hearing loss. And so adding captions ensures that your content and events are accessible.

There's also the SEO element. And this would apply more to pre-recorded video. And SEO stands for Search Engine Optimization. And essentially, a study from Facebook found that videos with captions had 135% greater organic traffic than ones without video.

Captions are also really important for branding, as I mentioned earlier, because they show that you're a brand that wants to make your content accessible. And it also helps with improving memory and behavioral intent and even brand recall.

Then there's the comprehension part. So if you are publishing a lot of educational content or hosting events that are educational, a study by USFSP actually found that 98.6% of students find captions helpful. And many of them were saying that it's because it helps them to focus on the content. So you can imagine if you're hosting a really complicated chemistry lecture or something like that, having captions will reinforce what the professor is saying and also help the students understand and focus better what is being presented.

And then the last part is engagement. 41% of videos are incomprehensible without sound or captions. So adding captions to your video will ensure that people have the flexibility to view your videos wherever they are and just really understand what is going on in your video.

Now, I want to cover a little bit around the accessibility laws. But I want to preface this by saying that I'm not a lawyer. So if you want to ensure that you're following proper procedures, I highly recommend you consult your legal counsel to see which laws apply to you.

The laws of the United States are the ones here I have on the screen, which I'll go through. Number one is the Rehabilitation Act of 1973. And the two sections that particularly apply to accessibility are Section 504 and Section 508. And Section 504 is a broad antidiscriminatory law that requires equal access for individuals with disabilities.

And Section 508 applies to federal communications and information technology. And Section 508 specifically mentions something called Web Content Accessibility Guidelines, which I'll cover in the next slide. And these essentially talk about adding closed captions to your videos.

The Americans with Disabilities Act is a broad antidiscriminatory law as well here in the United States. And in particular, Title II and Title III are going to apply around accessibility laws. Title II is for public entities, and Title III is going to be for private entities, so things like hotels, airports, or private businesses.

Traditionally, the ADA was applied to physical locations, so libraries, cinemas. But because of the advent of the internet and because we're relying on the internet more to conduct everyday business and things that we do, it's being applied to online as well. For example, a couple of years ago, Netflix was sued for lack of closed caption on their videos. And the courts found that because you can enjoy Netflix wherever, on your phone, if you're at a coffee shop, it is considered a place of public accommodation and therefore should be made accessible.

The CVAA and FCC are going to apply more for people in entertainment. So the CVAA specifically says that any content that previously aired on television must be published online with captions. And it also phases in something called audio description, which is an accommodation for blind or low-vision viewers. And then the FCC is going to require captions and have certain standards for broadcast video.

I mentioned the Web Content Accessibility Guidelines. And it's important to focus on these, because they're really helpful for making your online content accessible. They're often referred to as WCAG, so you might have heard that.

And there's different versions. So the one that is typically referred to in laws such as Section 508 of the Rehabilitation Act is going to be WCAG 2.0. But there is the most recent one, which is, like, 2.4. So if you are following these guidelines to ensure your website is accessible, I always recommend to go to the most recent one because it will include things like mobile accessibility and changes to the web.

There's different levels around WCAG. So level A is going to be the easiest to achieve. And then level AA is going to require a little bit more work, but it's going to be the most accessible. In terms of video accessibility, level A is going to require transcripts for audio-only content. So think podcasts, captions for pre-recorded video, so anything you upload to YouTube, and then audio or text alternatives for audio description.

Level AA is going to require captions for pre-recorded video, captions for live, so any events you do, and then audio description for pre-recorded video. And then level AAA is going to require sign language track for your event or for your videos, extended audio description, and then live transcripts for audio-only content.

Now, quickly, I just want to dive into who we are here at 3Play and what resources we have available for you. So if you know us for [? educational ?] content, we can also help you make your videos accessible. We offer services in closed captioning and transcription. We have a live automatic solution which can now display your captions natively on any video player. We also offer subtitles and translation. And then audio description, as I mentioned, is just an accommodation for blind or low-vision which essentially describes what is going on in the video.

We focus here on helping you make your videos accessible and future-proof and provide you with the solutions to grow as you grow with your video content. We can scale with you. We have the highest deadline compliance in the industry. And we can process large quantities of files.

We allow you to upgrade any time. So if you use our live automatic captioning solution, you can easily upgrade to full transcription, where we'll go back and edit your transcript. You get access to an account manager to help you make sure that you're staying on top of accessibility laws and your goals.

And then we also just provide you a lot of flexibility. We work with a lot of customers who have very complex video platform and different solutions. And so we really work to help accommodate that.

We publish a ton of free resources that I recommend you check out on our website. We publish weekly blogs around what is going on in accessibility and free white papers and checklists, monthly webinars such as this, where we bring in industry experts to talk about accessibility. And then we have a free video accessibility course, where you can just dive in deeper into the elements of accessibility, so things like the law, how to [? publish ?] [? is ?] closed captioning, all those.

Some upcoming free events that are happening this week-- we're celebrating Global Accessibility Awareness Day with actor Mickey Rowe, who he was one of the first autistic actors to be featured on a major role in Broadway. And then deaf activist Michael Agyin, who has a really great TED Talk around growing up deaf in Compton and the barriers that society has for deaf people and what we can do to help improve that.

And then we have a free ACCESS at Home. It's a virtual conference. And this one focuses around going back to normal in a safe and accessible way. So I definitely encourage you to register for those.

All right, well, that's all I have for the portion of the presentation and now would love to dive into your questions. Let me-- all right.

So the first question we have here is "Does video player equal streaming platform?" Yes, for live captioning, that also means streaming platform. And a lot of video players will offer both, like YouTube or Brightcove, that kind of thing.

So the next question we have-- which, Ryan, you may be able to help answer this better-- is "Has your live automatic captioning changed in the last two months? In the past, I found the setup to be complex for Zoom. How has it changed recently?"

**RYAN
MARTINEZ:**

Absolutely, so I can jump in on that. So the closed-captioning process requires the same sort of general steps of your Zoom event or your Zoom platform needs to be linked to 3Play with username and password credentials. After that process happens, any event that you schedule in Zoom is automatically visible in 3Play.

But 3Play doesn't have the ability to automatically provision live resources just when we see that scheduled event. So it does require somebody to log in to actually select that event to receive closed captions, which is the signal for 3Play to provision our live infrastructure to begin listening for audio for that event.

But aside from that process of a 3Play user logging in to specify which event needs captions, the rest of the workflow happens seamlessly. And that the host of the meeting simply sees the option to add closed captions to the meeting. And the rest of the workflow happens from there.

SOFIA LEIVA: We had a couple questions around WCAG, so I'm happy to answer those. We had a question whether auto-captions are considered WCAG level A or AA compliant. And like I mentioned, there isn't a governing body around the compliance of live captions. Basically, the WCAG just says that in order to meet level AA, you need to provide live captions. Typically, with automatic captions, the accuracy is going to be lower. So it is preferable to use a human solution. But that's just up to you and to your budget needs.

The next question we have is-- and, Ryan, maybe you'll be able to talk a little bit about this-- is "Do you have time to go over how to set up live captioning?"

RYAN MARTINEZ: So for the purpose of today, we may not be able to kind of screen share and walk through the whole process. But if there are certain platforms that you're interested in, generally speaking, 3Play offers four different platform integrations-- YouTube, Brightcove, Facebook, and Zoom. And so of those four integrations, we're able to deliver in-player captions to each of those four players-- again, YouTube, Brightcove, Facebook, and Zoom.

Aside from that, 3Play has a workflow that is very flexible that essentially uses RTMP streaming protocols and 608 captioning coding, which both are standards in livestreaming. So as long as the platform where these live events are hosted can accept both our RTMP and 608, 3Play can actually package both video and captions into a single stream and send it to an endpoint that you dictate.

So between the four integrations, those are all platform-specific. But our RTMP/608 workflow is by far our most flexible option because you can use it on so many different video platforms, again, assuming that they accept those standards.

SOFIA LEIVA: The next question we have is "What is your price for human-written live captions?" And currently, we don't offer a human solution. Our pricing for live automatic is around-- how much, Ryan?

RYAN MARTINEZ: It's \$0.60 per minute for automated. And so the solution in that-- that equates to about \$36 an hour. I had mentioned in one of the chats earlier, though, that 3Play has a seamless upgrade path. As Sofia mentioned earlier, if your content is going online, you really don't want to post it online with live-quality captions. Though the accuracy is fine for the live event, it's not recommended for online content.

And so 3Play has a very seamless path for which you can upgrade that content to 99%. And we actually deduct the per-minute rate for the live event from the regular cost of our traditional transcription service. So it winds up being more or less a free live event if you ultimately choose to upgrade. The cost to you is no different for getting those 99% accurate captions.

SOFIA LEIVA: The next question we have is "What's the accuracy of 3Play's ASR for live audio captioning?"

RYAN MARTINEZ: Sure, so that's a really great question. The live accuracy can vary. As Sofia had mentioned, a lot of it depends on network connection, single speakers, background noise, or lack thereof. But 3Play's live accuracy does achieve anywhere from 90% to 95% on the high end. But once again, even when you consider true accessibility for content online, that 95% is still markedly different from a full 99% accurate standard.

So, though, in a best case scenario-- 90%, 95%-- live accuracy I've seen can range as low as 60% to 70% if there is a lot of background noise or multiple speakers. But again, considering this idea of live as an accommodation versus what is considered truly accessible is important.

SOFIA LEIVA: Thank you. The next question we have is "Can a terminology list/glossary be plugged into your ASR engine? And how does it deal with proper nouns or personal names?"

RYAN MARTINEZ: So that's also an awesome question, and relates to a feature that we've rolled out, actually, within the last couple of months. We actually call it Wordlist, but it's exactly what you're describing there. Essentially, because we're doing speech with automated speech recognition, that automated speech recognition software performs better if it knows what to listen for. So you can include things like acronyms, proper nouns.

We actually have a customer in the Boston area, which is where 3Play is, and our mayor is Mayor Marty Walsh. And we had an event recently that was uploaded related to COVID protocols and some of those announcements. And Mayor Marty Walsh was something that was included in those word lists. Because without it, you could spell it M-A-R-T-Y, T-I-E-- lots of variations there.

It's still imperfect. So in some cases, it may make mistakes on proper nouns or acronyms. But that, again, is why 3Play's upgrade path and our editing interface is so important. Because if you need to come in and make simple edits, you can do that. And if you want to upgrade the full event to 99%, we'll take care of those proper nouns for you.

SOFIA LEIVA: The next question we have is, "If I use a different solution for my live captioning, can I upload that transcript to 3Play for a full review? Or how would that work?"

RYAN MARTINEZ: Absolutely. So it wouldn't be a situation where-- I should clarify, it wouldn't be sort of a full review. In that situation, if there were mistakes, or the transcript was incomplete, we would request that the media file actually be uploaded so we could put it through our full process, which includes a round of automated speech recognition and then two rounds of human editors to clean that up and achieve the 99%.

If, however, you had a full transcript where it was verbatim, word for word, you could upload that into our system along with the media file, and we will actually timecode that transcript and create many different downloadable caption files. So if your ultimate goal is to put that transcript in a format to where it can be posted online or order additional services like audio description off of that existing transcript, that is an option.

SOFIA LEIVA: Thank you. The next question we have is "Do you provide a web page where the live ASR captions appear if we hire you for a live event?"

RYAN MARTINEZ: Yes, so that's awesome. That's an awesome use case. So we've briefly touched on this in the presentation earlier. But when you schedule live captions, whether you're using one of the four platforms we offer-- YouTube, Brightcove, Zoom, or Facebook-- or you're using that RTMP workflow that I had mentioned, that flexible workflow, the event itself-- let me go ahead and just double-check.

Oh, where was that question, Sofia? I apologize. I just want to make sure to hit on every point of that. I was using it for reference.

SOFIA LEIVA: My bad.

**RYAN
MARTINEZ:**

SOFIA LEIVA: "Do you provide a web page where the live ASR appear if we hire you for a live event?"

**RYAN
MARTINEZ:** Yeah, so sorry. I lost my train of thought-- a lot of moving parts to that. So when you schedule your captions, we're going to push to that end point that you tell us. So whether it's in that player or that URL that you provide via that RTMP link, we'll push captions to that end point.

But every single live event is provisioned with a captions embed, which is basically a rectangle that we push the text for that event to. So for displaying on an external web page, you simply need to embed your video element and embed this 3Play captions embed underneath the video. And that allows for users to view the captions in players that don't support in-player caption options.

SOFIA LEIVA: Thank you. The next question we had here is "What's the confidentiality situation? Like, my company's policy doesn't allow external and/or cloud solutions that haven't been vetted thoroughly. Where does the data end up? And do you use a third-party resource?"

**RYAN
MARTINEZ:** Sure. So in situations like if you were to use a Zoom integration, or you would use our Brightcove integration, that all automatically creates video resources within that video platform. However, in order to use those integrations, this has to be video platforms that you own, that you have credentials for, to where you're actually logging into those platforms through 3Play with your ordinary credentials. So in a lot of cases, there isn't as much privacy concern there because, presumably, you've signed a contract with that video platform that acknowledges some of the same things we're talking about here.

When it comes to the transcripts that 3Play hosts in our system, we are an AWS company, so it is all hosted in the cloud. But our editors, as they're working on your content, all sign NDAs and confidentiality agreements. And they face strict penalties should they break those. And 3Play also, because we maintain our marketplace and our editing infrastructure, we maintain a lot of that responsibility as well.

So if there's some interest in seeing some of that documentation, that's all something that we can share around. So if you want to reach out directly, we'd be happy to provide that information if confidentiality is a concern.

SOFIA LEIVA: Well, that's all the time that we have for today. Thank you so much, Ryan, for joining us here. And thank you, everyone, for attending the webinar. If you have any more questions around our live solution or 3Play Media in general, feel free to reach out to us on our website. And I hope everyone has a wonderful rest of your day. Bye.

**RYAN
MARTINEZ:** Thanks, everyone.