

Online Presentations with *PowerPoint Present Live* Real-Time Automated Captions and Subtitles: Perceptions of Faculty and Administrators

Anymir Orellana

Georgina Arguello

Elda Kanzki-Veloso

Nova Southeastern University, USA

Abstract

Captioning of recorded videos is beneficial to many and a matter of compliance with accessibility regulations and guidelines. Like recorded captions, real-time captions can also be means to implement the Universal Design for Learning checkpoint to offer text-based alternatives to auditory information. A cost-effective solution to implement the checkpoint for live online presentations is to use speech recognition technologies to generate automated captions. In particular, *Microsoft PowerPoint Present Live* (MSPL) is an application that can be used to present with real-time automated captions and subtitles in multiple languages, allowing individuals to follow the presentation in their preferred language. The purpose of this study was to identify challenges that participants could encounter when using the MSPL feature of real-time automated captions/subtitles, and to determine what they describe as potential uses, challenges, and benefits of the feature. Participants were full-time faculty and administrators with a faculty appointment in a higher education institution. Data from five native English speakers and five native Spanish speakers were analyzed. Activities of remote usability testing and interviews were conducted to collect data. Overall, participants did not encounter challenges that they could not overcome and described MSPL as an easy-to-use and useful tool to present with captions/subtitles for teaching or training and to reach English and Spanish-speaking audiences. The themes that emerged as potential challenges were training, distraction, and technology. Findings are discussed and further research is recommended.

Keywords: Online presentation, real-time, captions, subtitles, speech recognition, universal design for learning

Orellana, A., Arguello, G., & Kanzki-Veloso, E. (2020). Online presentations with Power Point Live real-time automated captions and subtitles: Perceptions of faculty and administrators. *Online Learning*, 26(2), 34-51.

Captioning videos is beneficial to many, including individuals who are deaf or hard of hearing, hearing adults wanting to retain what is heard, and persons learning a second language (Dallas et al., 2016; Gernsbacher, 2015; Linder, 2016; Morris et al., 2016). Captioning is also a matter of compliance with accessibility regulations and guidelines, such as the Americans with Disabilities Act (United States Department of Justice Civil Rights Division, n.d.), the Rehabilitation Act Section 508 (U.S. General Service Administration, n.d.), and the Web Content Accessibility Guidelines 2.0 (World Wide Web Consortium, n.d.).

From an instructional and learning perspective, captioning is of particular applicability when aiming to implement the Universal Design for Learning (UDL) principle of providing “multiple means of representation” (CAST, 2018a; Meyer et al., 2014). UDL is an evidence-based “framework to improve and optimize teaching and learning for all people based on scientific insights into how humans learn” (CAST, 2018a, para. 1). UDL promotes inclusive pedagogy that beneficially supports diverse students and reduces the need for specific accommodations. According to CAST (2018a), the UDL Guidelines “offer a set of concrete suggestions that can be applied to any discipline or domain to ensure that all learners can access and participate in meaningful, challenging learning opportunities” (para. 1).

Offering alternatives to auditory information can allow all learners to access the content equally (CAST, 2018a), for example, with the use of “text equivalents in the form of captions or automated speech-to-text (voice recognition) for spoken language” (CAST, 2018b, para. 2). Figure 1 depicts UDL Checkpoint 1.2 “Offer alternatives of auditory information” within the UDL Principle “Provide multiple means of representation,” UDL Guideline 1 “Provide options for perception.”

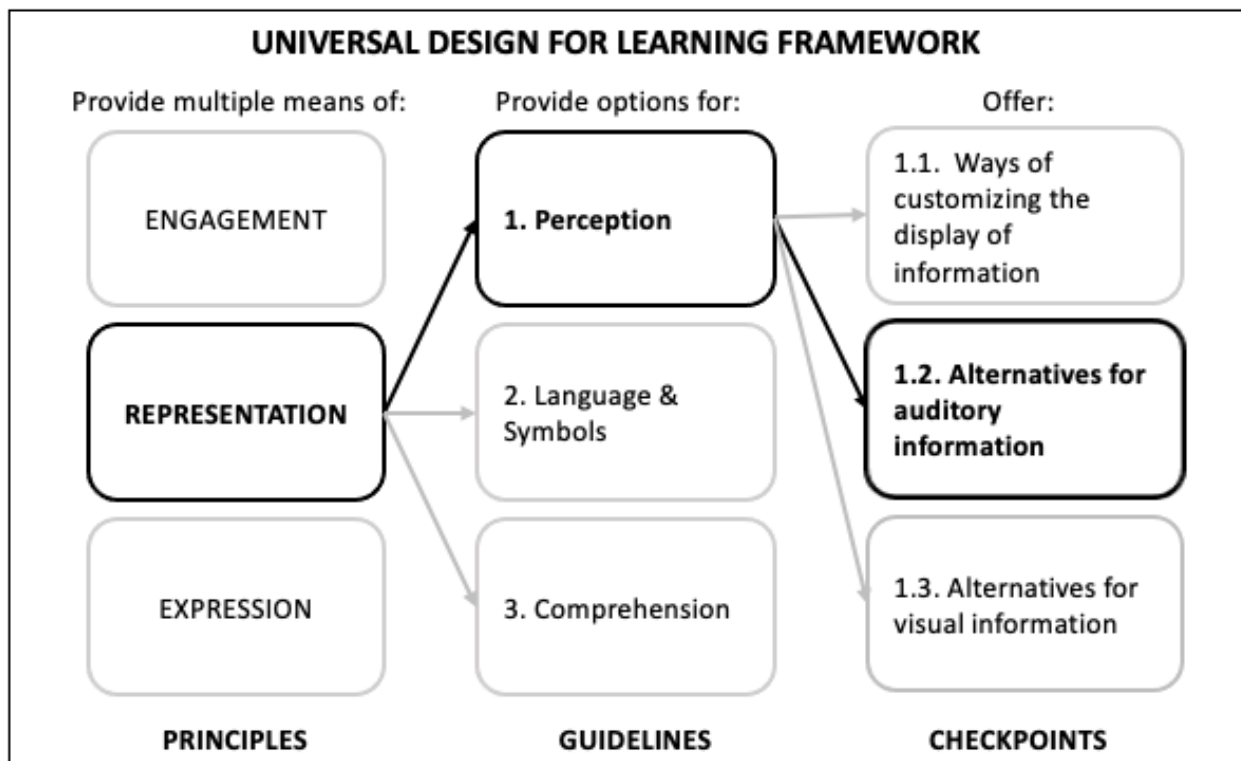


Fig. 1 Checkpoint 1.2 “Offer alternatives for auditory information” outlined under Universal Design for Learning principle “Provide multiple means of representation,” Guideline 1 “Provide options for perception” (CAST, 2018a).

Figure created by authors.

As suggested by UDL Checkpoint 1.2, real-time captions can be an alternative to auditory information when the presenter is speaking live and online. However, real-time captioning can be expensive if a human transcriber is to caption every presentation in every live session. A cost-effective solution can be speech recognition technology (SRT) to generate real-time captions and to provide a transcription of the speech (Revuelta et al., 2010). Students have found SRT beneficial in educational settings, such as in their English-language lectures (Huang et al., 2015; Huang et al., 2016) and for cross-cultural learning activities (Shadiev et al., 2018).

In 2020, *Present Live* (MSPL) became available as a *Microsoft PowerPoint* (PPT) presentation feature that allowed real-time automated captions, the translation in real-time of 12 spoken languages into more than 60 languages, the possibility of individual viewers to follow the presentation in their preferred language using their own devices, and the ability to compile a transcript of the presentation (Microsoft Education, 2020). PPT is a commonly used tool for presentations in educational settings and it can be anticipated that those with a Microsoft Office 365 license would be inclined to use it. As online instructors and administrators aim to implement UDL guidelines for inclusive learning opportunities, reach out to multilingual audiences, and comply with regulations and guidelines regarding accessibility, the following questions arise: “Would online instructors use a tool like MSPL to offer captions as a text-based alternative for auditory information when they are presenting online in real time?” “Would online instructors use MSPL to translate their spoken words when they are presenting online in real time to reach students who speak or are learning a different language?” “Would online instructors be able to use MSPL effectively, and would they find the features of captions/subtitles useful?”

Review of Related Literature

Captions are typically referred to as the transcription of the presenter’s speech in their language along with background sounds and speaker identification, whereas subtitles are referred to as the translation of the speech into a different language (3PlayMedia, n.d.; Myers, 2019; Take Note, n.d.). Closed captions/subtitles can be turned on and off by the viewer, as opposed to open ones that are always visible on screen (Bureau of Internet Accessibility, 2019). Captions/subtitles can be generated in real time or added to the recorded video offline in post-production time, and they can be generated by a human transcriber or with speech recognition technology (SRT).

Gernsbacher (2015) documented more than 100 empirical studies that showed how captions benefit a diverse population, including individuals who may be deaf or hard of hearing, hearing adults wanting to retain what is heard, and persons learning a second language. Linder (2016) surveyed 2124 students without hearing disabilities from 15 institutions enrolled in different course modalities—online, face-to-face, and hybrid—to determine how they used and perceived closed captions and transcripts for recorded videos. Respondents indicated that using captions helped them focus, retain information, overcome poor audio, access the content in quiet environments, comprehend complex vocabulary, overcome difficulty with hearing, and better comprehend English as their second language (Linder, 2016). Among the benefits of displaying on-screen subtitles are the comprehension of viewers who speak a different language, reaching a larger audience, and allowing viewers to learn a foreign language.

Dallas et al. (2016) examined the relationship between students' exposure to captions and information recalling. Dallas et al. analyzed data from 216 randomly selected undergraduate students without a hearing disability or not having English as a second language. Dallas et al. found that those exposed to captions performed better on information recall, although sophomores scored lower compared to seniors and African Americans scored lower compared to Caucasians. In general, Dallas et al. concluded that "closed captions may be beneficial for learning video-based information [and that] faculty members are encouraged to turn on closed captions when showing course-related videos in class or for online courses" (p. 62).

Morris et al. (2016) surveyed 66 students regarding their "perceived advantages or disadvantages of their experience with captioning in the current [online] course" (p. 233). Morris et al. found that 99% reported that captions helped clarify content, the spelling of keywords, and note taking. Additionally, although a 99% accuracy was reported from the captioning vendor, students noted "issues and missing spaces between words were observed, and these errors were a potential distraction, possibly limiting the value of the captions" (p. 235).

Speech recognition, also known as "automatic speech recognition (ASR), computer speech recognition, or speech-to-text, is a capability that enables a program to process human speech into a written format [and] focuses on the translation of speech from a verbal format to a text" (IBM Cloud Education, 2020, What is Speech Recognition section, para. 1). The industry standard for caption and transcript quality is at least 99% accuracy rate and, on the other hand, according to Enamorado (2019a), "typically, automatic speech recognition produces about 60-70% accurate transcripts, which means that 1 out of 3 words is wrong" (Automatic Speech Recognition section, para. 3). Additionally, Enamorado (2019b) compared the accuracy rates of two vendors and found that their measured accuracy rates fell between 84.7% and 94.4%.

In general, captions generated by SRT are not 100% accurate and often need a human to edit them for full compliance with accessibility regulations and guidelines. Typically, ASR "is good, but not good enough to remove humans from the process" (Enamorado, 2019b, Why is it 99% Accuracy and Not 100%? section, para. 2) and "is often fast, cheap, but highly inaccurate" (Enamorado, 2019a, Automatic Speech Recognition section, para. 1). Despite the typical low accuracy of text generated with SRT, students have found SRT beneficial in their English-language lectures to aid learning, to help them better understand a lesson, to allow them to take notes, and to confirm what was being said in the class (Huang et al., 2015; Huang et al., 2016). Huang et al. (2016) summarized studies that looked at how STR supports the learning of non-native English speakers and concluded that the literature showed that for the most part, students found SRT helpful during real-time lectures and as compiled transcripts. The use of SRT in the classroom can also aid awareness, attention, and meditation (Shadiev et al., 2017).

Shadiev et al. (2018) investigated the use of speech-enabled language translation (SELT) technology, which consists of SRT and computer-aided translation, to facilitate cross-cultural understanding and intercultural sensitivity. Shadiev et al. (2018) computed the accuracy and intelligibility of the 10 different languages among 21 multilingual students representing 13 nationalities. Shadiev et al. (2018) found that the texts generated were meaningful and valuable to participants in their cross-cultural learning activity and suggested "applying SELT to support student interaction in their native language" (p. 1425).

The use of SRT to generate real-time captions can also be a cost-effective solution for classroom presentations where it would otherwise be necessary to hire dedicated staff (Revuelta et al., 2010). According to Revuelta et al., the essential use of ASR technology inside the classroom is to transcribe what the instructor presents in real time. Regarding presentation tools that allow for automatic real-time captioning, PPT is a presentation application that uses cloud-based SRT for real-time captioning of the spoken words of the presenter (Microsoft, n.d.-b). The feature of PPT live captions/subtitles is “one of the cloud-enhanced features in Microsoft 365 and are powered by Microsoft Speech Services” and, to provide the service, the speech utterances are sent to Microsoft (Microsoft, n.d.-b; Important Information About Live Captions & Subtitles section, para. 1). In 2018, the PowerPoint team announced this new feature powered by artificial intelligence that would allow PPT to support “12 spoken languages and display on-screen [real-time] captions or subtitles in one of 60+ languages” (PowerPoint Team, 2018, para. 2). As of late January 2019, this feature has been available for Office 365 subscribers worldwide for PPT on Windows 10, PPT for Mac, and PPT Online. The Microsoft Education Team (2019) claimed that a benefit of this feature is having a “speech recognition that automatically adapts based on the presented content for more accurate recognition of names and specialized terminology” (Present More Inclusively with Live Captions & Subtitles in Microsoft PowerPoint section, para. 2).

MSPL for Office 365 was announced in January 2020 (Microsoft Education, 2020) and became available in PPT for the web by June 2020 (Johnson, 2020). A MSPL presentation can be shared with anyone who has internet; viewers anywhere can join the live presentation on their devices and read live captions/subtitles in their preferred language as the speaker is presenting. The live presentation can be delivered to an audience onsite or to an online audience connected to a conferencing system by sharing the screen (Microsoft, n.d.-a).

Purpose and Research Questions

The MSPL feature of real-time automated captions/subtitles can be a means to implement the UDL guideline that suggests that a way to reduce barriers is to provide a real-time, text-based alternative to auditory information. Additionally, with MSPL the viewers can follow the presentation in their preferred language. The purpose of this study was to identify challenges that participants could encounter when using the MSPL feature of real-time automated captions/subtitles, and to determine what participants describe as potential uses, challenges, and benefits of the feature. For the study, captions were referred to as the transcription of the presenter’s speech in their same language without background sounds or speaker identification, and subtitles as the translation of the speech into a different language. In particular, the focus of the study was on captions and subtitles in English and Spanish. Participants were English- and Spanish-speaking full-time faculty and administrators with a faculty appointment in a higher education institution. To address the purpose of the study, the following questions were addressed:

1. What challenges do participants encounter as presenters and as viewers with MSPL?
2. What do participants describe as potential challenges, benefits, and uses of real-time captions and subtitles with MSPL?

Methods

Setting of the Study

The institution that served as the setting of the study was a private not-for-profit university considered a majority-minority institution. The university had been recognized as a Hispanic Serving Institution with a diverse student population from more than 100 countries and more than 25% of its students identified as Hispanic. Participants of the study were affiliated to a college of the institution that offered online and onsite undergraduate and graduate programs of study. The college served students in the U.S. and several international locations, including Latin America and the Caribbean. As a result, to assist this population of students, the college offered some of the graduate programs of study in Spanish with Spanish-speaking faculty and doctoral dissertation committee chairs and members.

Data Collection

The researchers followed basic activities of *usability testing* to collect data. Barnum (2010) summarized usability as encompassing “the product’s effectiveness and efficiency for users, as they work with the product ... [and] the elusive quality of user satisfaction, which is based on users’ perceptions entirely” (p. 1). The researchers’ intent of following usability testing activities for the study was not to formally test MSPL as a product, nor to inform product developers or to conduct rigorous experimental designs that typically address three dimensions of usability (i.e., effectiveness, efficiency, and satisfaction). The intent was to use basic activities of usability testing as a study framework to identify challenges that participants could encounter as presenters and as viewers with MSPL.

Participants were observed working with MSPL performing the task of delivering and viewing a presentation with captions/subtitles meant to be “real and meaningful to them” (Barnum, 2010, p. 1). This observation activity is what Barnum describes as *usability testing*. Specifically, the researchers conducted activities of a moderated qualitative usability testing to gain an in-depth description of potential uses, challenges, and benefits of the MSPL feature of real-time captions/subtitles based on the experience and narrative of participants. According to De Blecker and Okoroji (2018) “qualitative usability studies are focused on gaining in-depth understanding based on narrative data, while quantitative studies collect numerical data to produce statistically relevant metrics” (Qualitative and Quantitative Usability Studies section, para 1). Furthermore, because of the restrictions of meeting onsite during the COVID-19 pandemic, the researchers scheduled a Zoom session with each participant to carry out what Barnum refers to as a *moderated remote usability testing* by “observing [via Zoom] in one location and the user [participant] in another location” (p. 2). After the testing session, participants were interviewed to determine how they described potential challenges, benefits, and potential uses of live captions/subtitles with MSPL.

Preparing the Moderated Remote Usability Testing

The researchers followed five steps recommended by Barnum (2010) to prepare the moderated remote testing:

1. Recruit participants. The study population included native English-speakers and native Spanish-speakers who were full-time faculty, or administrators with a faculty appointment, in a

college of the institution that served as the setting of the study. As employees of the college, all native Spanish-speaking participants were fluent in English. Additionally, as employees of the institution, all participants had licensed access to Microsoft Office 365 online to present with MSPL.

The researchers used purposive sampling to recruit 12 participants. When using purposive sampling in qualitative studies, a sample size from 7 to 12 is appropriate (Malterud et al., 2016; McCracken, 1988; Young & Casey, 2019). Similarly, for qualitative usability testing, “a small number of participants is sufficient to provide valuable results” (Bleecker & Okoroji, 2018; Qualitative and Quantitative Usability Studies section, para 2). For qualitative usability testing studies, there can be as few as 3 to 5 and as many as 12 to 15 participants (Bleecker & Okoroji, 2018).

An invitation to participate in the study was emailed to 57 potential participants. The first six English speakers and the first six Spanish speakers who accepted the invitation and met the inclusion criteria were recruited. Inclusion criteria were having experience using PowerPoint and Zoom, a headset or a microphone, fast and reliable internet connection, web camera, and a computer with a recent version of a browser (i.e., Mozilla Firefox, Google Chrome, or Microsoft Edge). Participants were encouraged to bring a smartphone or tablet with iOS version 11+ or Android version 8+.

2. Assign team roles and responsibilities. Two researchers fully fluent in Spanish and English (i.e., R1 and R2) met with each participant via Zoom. R1 moderated the session, guided the participant with a Walkthrough Protocol, troubleshoot, and compiled captions and subtitles in one language. R2 observed and took notes of the test session, collected text of captions and subtitles in the other language, completed the Walkthrough Checklist, and noted if and when the participant had issues completing each step.

3. Prepare other materials. The researchers prepared a consent form, a Walkthrough Protocol, an Interview Protocol, and a 6-minute video tutorial on using MSPL.

4. Create the qualitative semi-structured Interview Protocol. The protocol consisted of the researcher’s script and four open-ended questions about if and how the participant would use the MSPL features of captions/subtitles and about any challenges that they thought they would need to overcome when using these features. Participants were also asked to describe their role in the institution.

5. Test the test. Two faculty members with characteristics of participants validated the materials and completed the testing activity and interview.

Conducting the Moderated Remote Usability Testing

Each participant received an email with a unique link to a password-protected Zoom meeting. Before starting the usability testing, R1 made sure that the participant had the necessary equipment (i.e., microphone, browser, and/or mobile device) and asked the participant to test their network speed using the *Speedtest* website <https://www.speedtest.net>. Upon starting the testing, each participant viewed the 6-minute video tutorial, presented with MSPL three or more slides with the content of their choice, and connected as viewers to an MSPL with their devices

or with a different browser when R1 acted as the presenter.

Conducting the Interview

Upon completing the testing session, R1 interviewed the participant regarding their experience with MSPL. The video and transcripts of the Zoom interview session were recorded for each participant. Each interview lasted from 25 to 30 minutes and consisted of the following open-ended questions:

1. Demographic questions: Would you please briefly describe your primary role? In what programs(s) or courses are you involved (as teacher or administrator)? Can you describe your students or audience (e.g., general characteristics, needs, skills, online)? When you present live, do you mostly do it online or onsite?

2. Interview question 1: Would you please describe one scenario or more where you would use live captions or translated subtitles in a presentation? Describe the characteristics of the audience and the setting or type of presentation. Your audience can be onsite or online with Microsoft Teams or Zoom.

3. Interview question 2: How often would you use these features?

4. Interview question 3: How do you think a particular audience would benefit from following the presentation in their preferred language?

5. Interview question 4: Tell me about your experience with real-time subtitles in PowerPoint Live?

6. Interview question 5: Do you have anything you'd like to add or ask?

Data Analysis

One English- and one Spanish-speaking participant could not successfully present during their initial session nor a second scheduled session. Hence, data from ten participants were analyzed (i.e., five native English speakers and five native Spanish speakers). The notes in the Walkthrough Checklist taken during the usability testing were analyzed to describe the challenges that participants encountered as presenters and as viewers with MSPL.

The qualitative interview data were analyzed to determine how participants described potential challenges, benefits, and potential uses of live captions/subtitles with MSPL. A general inductive approach (Thomas, 2006) was followed. The inductive approach is used to develop “categories into a model or framework that summarizes the raw data and conveys key themes and processes” (Thomas, 2006, p. 240). Open coding was used to assign descriptive labels that came from the text of transcripts. The text was then grouped into categories and reduced until it could no longer be reduced. This process allowed the creation of essential categories that later emerged into major themes.

Triangulation (Denzin, 2009) was employed to ensure the trustworthiness of the analysis process and validity of the research process (Creswell & Poth, 2018). The interview transcripts from Zoom were downloaded, checked against the recorded Zoom session, revised accordingly, and then sent to each participant to adjust for inaccuracies.

According to Gibbs (2018), “Coding is a way of indexing or categorizing the text to establish a framework of thematic ideas about it” (p. 54). The transcripts were coded to develop categories using a member checklist process that consisted of coding separately and then meeting to reach an agreement in the categories that emerged. First, the text was segmented into sentence

fragments, sentences, phrases, and paragraphs and assigned a descriptive label (i.e., code) to each qualitative data unit (i.e., text from the interview transcript). Then, codes were grouped into categories to connect the codes and attribute meaning to the data units. The process of open coding was exhausted until all the categories were created.

Once the categories were developed, they were checked to achieve consistency amongst them. Categories were then combined to create axial codes that allowed for the central meaning of each category. Subcategories were then created, which identified the core meaning of the open codes. Last, themes started to emerge from the axial codes. This process was repeated twice for accuracy of themes. Finally, the themes that were extracted were reviewed with the transcript to confirm the meaning. A prolonged engagement technique (Lincoln & Guba, 1985) was followed by meeting several times to understand better the analysis process and the themes that emerged.

Results

Characteristics of the Participants

An inclusion criterion was that the participant had experience presenting with PPT for any of their job-related roles and using Zoom to remotely participate in the study. In general, participants used multiple delivery methods (e.g., onsite, online, and hybrid) when performing their roles. However, because of the COVID-19 pandemic at the time of the study, they had been delivering all their synchronous meetings and class sessions online via Zoom.

Out of the ten participants, two identified themselves as administrators with a faculty appointment. All ten taught online graduate students and two also taught onsite undergraduate students. All taught in English and two Spanish-speaking participants also taught hybrid courses in Spanish to students in Colombia, Puerto Rico, or the Dominican Republic. Seven indicated that they served as doctoral dissertation chairs or members to national or international online students who spoke their same language.

In general, participants described their typical graduate student population as non-traditional working adults who were mainly technologically proficient. Six who taught in English indicated that they interacted with students who had English as their second language (e.g., Spanish, Haitian Creole, Portuguese), Spanish being the predominant language. In general, the participants highlighted that these diverse students were proficient in English, but their primary language of listening and speaking was some other language.

Potential Uses and Benefits of MSPL

The central theme that emerged as potential uses of MSPL was the possibility to deliver online presentations for training and teaching, especially during the COVID-19 pandemic. This was a surprising theme perhaps because of participants' experiences during the pandemic. For example, participants stated that most presentations moved to Zoom because of COVID-19, which made perfect timing for the use of MSPL. Following is an excerpt from a participant's Zoom interview transcript:

Because of COVID I think most of them [students] had experience using Zoom. Maybe we could use for some sessions ... PowerPoint live, and all the students are from the U.S. so we will not use Spanish I could see using PowerPoint live to present the workshop and use subtitles. I would use them in English.

Two other themes emerged as potential uses and benefits of MSPL captions/subtitles: English-speaking audiences would be able to verify the information from the speaker, and captions/subtitles would be beneficial to several audiences (e.g., English-speaking students, English- and Spanish-speaking doctoral dissertation chairs, and students with English as a second language). The following are excerpts from participants' Zoom interview transcripts that support the themes:

Teaching classes to students that speak Spanish. I think students would like it, those students who want to have their primary language, their first language, but would also like to get exposure to English.

I have a student who is English speaking ... from Jamaica, and I have a student who is Spanish speaking. He is proficient in English, but I think that ... I might ask him or let them know that we can do this, and he may opt to do the subtitles on his device in Spanish. It would benefit them to have the captions for subtitles so that they'd be able to make sure that they're getting all the information that you're providing.

Any dissertation-related presentation could have been done using it so they can still see this good. You know the Puerto Rican, or any international student as well, would benefit from this.

Another theme that emerged was the benefits of using MSPL as a friendly and easy-to-use tool that allows access to the presentation using any device and helps the viewer confirm what the speaker is saying. In relating how friendly and easy MSPL was to use, one participant stated, "I did not find it distracting as a presenter to have the subtitles underneath, which, you know, you might think that would be distracting to have the constantly appearing under your presentation, but I didn't find it distracting at all." Another participant shared how subtitles can help students who may not understand teachers or other students who speak in a different accent than their own to connect with what the speaking is saying. Participants also described that MSPL would be beneficial to special education students who are hard-of-hearing, non-English-speaking international students and doctoral dissertation chairs, and English-speaking students in conference settings.

Potential Challenges When Using MSPL

The following themes emerged as potential challenges when using MSPL captions/subtitles in live presentations:

1. **Training.** Participants described that the presenter would need training and a "refresher course." The audience would need tips on accessing the application, connecting to the presentation, and accessing the transcripts of the captions and subtitles.
2. **Distraction.** Participants described that when the speaker constantly checks for accuracy, they may cause the presentation flow to stop and thus, cause a potential distraction; additionally, talking too fast may cause many errors and, therefore, distraction when reading the captions/subtitles. One participant explained the distraction that may arise from using captions and subtitles by stating, "Our challenges would be that

perhaps the captions are coming too fast for some people who may need to have them at a slower pace.”

3. Technology. Participants indicated that adult learners and faculty, who are not technology savvy, might need extra training. Participants also wondered how onsite viewers would be able to read the captions/subtitles if they were not connected to the MSPL presentation, and how transcripts could be forwarded to those, online or onsite, who could not connect to the MSPL presentation.

Technology Used by Participants During Usability Testing

All participants used a laptop as presenters, out of which only one was an Apple Mac computer, and the rest were Windows-based. As for the browsers that participants used, eight used Chrome, one used Microsoft Edge, and another used Firefox.

To connect to the MSPL presentation when acting as a viewer, one participant used a second browser window and nine used a smartphone (seven used iPhones with iOS 11 or higher and two used a device with Android OS 8 or higher). The average speed of participants' network connection, measured as megabytes per second (Mbps), was 200.6 for download speed, varying from 31.66 to 400.53; and 100.33 for upload speed 4.64 to 531.18.

Challenges Encountered by Participants During Usability Testing

All participants were able to complete the testing as presenters and as viewers without significant challenges. Few ran into technical challenges before starting the testing session. If the participant was not able to resolve during a first session, a second session was scheduled.

One participant ran into several technical issues during the first session: not being able to log in to the institution portal using Chrome, a “freezing” Zoom session, problems with Bluetooth microphone, MSPL not yielding the QR code or link for viewers to connect, and MSPL suddenly stopping. The participant tested with several browsers and computers (e.g., Microsoft Edge and Chrome with a Windows computer, and Chrome and Firefox with an Apple Mac). During the last try with Microsoft Edge, the “Present Live” icon was not available and MSPL appeared unstable. During a second scheduled session, the participant completed the testing session using Firefox and a Windows computer.

Limitations of the Study

The limitations of the study were as follows:

1. The study was limited in the diversity of the sample. Additional information may have been learned from experiences of users from other institutions or educational settings.
2. The researchers conducted a qualitative usability testing study with a small sample size suitable for qualitative research. The researchers did not seek to conduct a quantitative usability testing study to collect numerical data or obtain statistically relevant metrics of MSPL.
3. Although MSPL allowed real-time captions/subtitles in various languages, English- and Spanish-speaking faculty were available for the researchers to recruit through purposive

sampling. Additionally, the researchers, who were native Spanish speakers fluent in English, needed to be able to read the captions/subtitles in both languages.

4. Participants used MSPL to present online only due to COVID-19 and, thus, were not able to comment on their experiences using MSPL in an onsite context.
5. Participants used MSPL in a testing scenario and, thus, they were not able to comment on their experiences using MSPL in their typical presentation scenario.

Discussion of the Findings

Findings were expected to help educators select presentation tools, such as MSPL, that allow automated real-time captioning when implementing UDL guidelines, specifically Checkpoint 1.2, which suggests that offering alternatives to auditory information can enable all learners to access the content equally. Diverse learners can benefit from real-time captions/subtitles, including those with hearing disabilities, have English as their second language, or want to retain the information by reading what they hear. Findings could also help faculty and administrators decide on tools to comply with accessibility regulations and guidelines.

By the time of the study, only *Google Slides* (Google, n.d.) as a presentation application allowed for real-time automated captions. MSPL was selected for the study as a licensed application that was readily available to the participants of the study. Additionally, unlike *Google Slides*, MSPL generated real-time captions/subtitles in various languages other than English and allowed the viewer to select the language of their choice. It is worth noting that MSPL was not formally tested as a product, nor were rigorous experimental designs for usability. Thus, findings were not meant to be used to inform product developers nor for product endorsement.

Challenges Encountered by Participants During Usability Testing

During the testing session, participants did not encounter technical challenges when using MSPL that they could not overcome by themselves or with the assistance of the testing moderator, nor did they describe potential challenges that they thought could not be resolved with proper training or tools. All who completed the testing session were connected to a stable internet with a network speed higher than the highest minimum recommended broadband by Zoom (Zoom Video Communications, n.d.) for the presenter, corresponding to high-definition video (i.e., speed rates of 3.8 Mbps for upload and 3.0 Mbps for download), and also higher than the recommended broadband as an attendee, corresponding to 1.2 Mbps download speed for high-definition video. The network download speed was higher than the 6 Mbps as the minimum recommended by the Federal Communications Commission (2020) for high-definition video teleconferencing. On the other hand, those who could not complete the testing, but were still able to stay connected via Zoom, had networks with download speeds of 2.07 Mbps and 5.99 Mbps, and upload speeds of 0.0Mbps and 0.13Mbps, respectively.

Given the performance values of participant's networks, the following can be concluded: (a) download and upload speeds as low as 31.66 Mbps and 4.64 Mbps, respectively, were appropriate to hold a Zoom session as an attendee and to present using MSPL features of captions/subtitles, and (b) upload speeds lower than 1 Mbps can prevent the proper use of MSPL

as a presenter. Overall, it can be concluded that the network speed was a significant obstacle that prevented the proper use of MSPL.

Participants' Descriptions of Potential Challenges, Benefits, and Uses of MSPL Real-Time Captions/Subtitles

Overall, participants described MSPL as an easy-to-use and helpful tool to provide captions/subtitles and reach English and Spanish-speaking audiences. It was surprising that only one participant mentioned accessibility as a reason for using captions and that none emphasized the inaccuracies of the captions. One participant voiced the benefit of MSPL for a Spanish-speaking student in their class, saying, "It's kind of [an] exciting idea to be able to speak in English, and other students see it in English, but for him to be able to have that choice of having an English or Spanish [translation] is a great idea." All participants described the features of captions/subtitles as a "benefit for all" for various scenarios (e.g., presentations, training), primarily online, and to multiple types of audiences (e.g., English, and non-English speaking students, and Spanish-speaking dissertation chairs). For instance, regarding the benefits of using MSPL, one participant stated, "I mean, this is something that is professional development for me. I mean, this is useful stuff." It can be concluded that MSPL can help provide a text-based alternative to auditory information presented live, as suggested by UDL Checkpoint 1.2.

After a more in-depth review of the interview transcripts and after further discussion, it was apparent that the pandemic influenced how participants perceived the uses and benefits of MSPL. For example, all mentioned benefits for class and meeting presentations online only, in a world where no traveling would be possible as in the time of a pandemic.

Recommendations

As more presentation applications with SRT-based real-time captions/subtitles become available and the existing ones improve their technologies, the possibilities of using them in the day-to-day presentations in classrooms or training are likely to increase. Although studies show the potential value of SRT for increasing inclusiveness, accessibility, and communicative ability with multilingual audiences, more research is needed to support the usefulness and effectiveness of presentation tools such as MSPL in classroom settings. A venue for this line of inquiry is through a better understanding of students' experiences in various scenarios (e.g., online, onsite, and hybrid) and for different types of students (e.g., with and without learning or hearing disabilities, undergraduates, graduates, native and non-native English-speakers).

The ten participants resided in the United States and were English and Spanish native speakers. Further research is recommended with a larger sample size and with participants who speak other languages and from different institutions. Furthermore, participants did not fully act as a presenter with an authentic audience and further research is recommended in more realistic scenarios where the presenter speaks freely to their typical audience and with more and relevant presentation slides. It can also be beneficial to include an audience connected from other countries or places where participants might need to overcome different technological and technical barriers.

Participants perceived MSPL as an easy-to-use tool and all agreed that training would be needed before its use. If and when a new tool is to be introduced and training provided, it is recommended that the reality of the participants is considered because it could influence how the

usefulness of the tool is perceived. Participants could dismiss the potential benefits of the tool because of more significant issues taking precedence in their lives, such as the pandemic.

Ease of use and perceived usefulness of a tool are essential factors to consider when deciding to use a tool like MSPL. It is also important to evaluate if the tool generates quality captions and subtitles measured by their accuracy and intelligibility. Hence, a comprehensive evaluation of the usefulness of MSPL should include determining the quality of the captions/subtitles it generates to determine to what extent MSPL can “accommodate individuals in the audience who may be deaf or hard of hearing” (Microsoft, n.d.-b, para. 1) and allow those who speak a different language from the presenter to comprehend the subtitles effectively.

Technical (e.g., poor network speed rates, poor microphones) and technological challenges (e.g., outdated software, hardware, versions of mobile devices and browsers) encountered by participants led to reflection about the working-from-home situation confronted by many because of the pandemic. If leaders of institutions expect faculty and staff to work from home efficiently, they must foresee these challenges and provide proper tools, training, and assistance.

Sudden instability of MSPL is also a significant issue that prevents its use and cannot be resolved by the user. It is not uncommon for cloud-based services to become unavailable because of outages or become unstable because of updates or maintenance. After conducting the study, it was noted that the interface of MSPL had changed regarding placements and labels of options and the placement of the presentation link for viewers to connect to the presentation. Changes in the interface and functionality of applications also affect training materials, such as printed tutorials or videos. Thus, it is recommended that training materials be revised frequently. Users are given training “refreshers” before using the tool, and that users be aware that technology “can go wrong” and should have an alternative plan.

Finally, having conducted a remote usability testing via Zoom presented challenges and opportunities for the researchers. Challenges included moderating the session remotely and troubleshooting without physically being able to assist the participant. On the other hand, the opportunities outweighed the challenges: Being able to record the interview video with Zoom allowed for validation of what was heard and observed; obtaining Zoom’s automatic transcripts, although not 100% accurate, facilitated the data collection and analysis; the possibility of scheduling individual sessions without the need of physical rooms or the commute saved time and resources; and using MSPL in real time with Zoom allowed participants to experience MSPL as presenters to a remotely located audience and as remote viewers connected to the presentation.

Declarations

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

The authors assert that approval was obtained from an ethics review board (IRB) at Nova Southeastern University, USA.

The authors declared that they received no financial support for the research, authorship, and/or publication of this article.

References

- 3PlayMedia. (n.d.). *The ultimate guide to closed captioning*.
<https://www.3playmedia.com/learn/popular-topics/closed-captioning/>
- Barnum, C. M. (2010). *Usability testing essentials*. Morgan Kaufmann.
- Bureau of Internet Accessibility. (2019, April). *Checklist for creating accessible videos*.
<https://www.boia.org/blog/checklist-for-creating-accessible-videos>
- CAST. (2018a). *Universal Design for Learning guidelines version 2.2*.
<http://udlguidelines.cast.org>
- CAST. (2018b). *Universal Design for Learning guidelines version 2.2., Checkpoint 1.2: Offer alternatives for auditory information*.
<https://udlguidelines.cast.org/representation/perception/alternatives-auditory>
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches*. Sage.
- Dallas, B. K., McCarthy, A. K., & Long, G. (2016). Examining the educational benefits of and attitudes toward closed captioning among undergraduate students. *Journal of the Scholarship of Teaching and Learning*, 16(2), 56-65.
<https://doi.org/10.14434/josotl.v16i2.19267>
- De Bleecker, I., & Okoroji, R. (2018). *Remote usability testing: Actionable insights in user behavior across geographies and time zones*. Packt Publishing.
- Denzin, N. K. (2009). *The research act: A theoretical introduction to sociological methods*. Routledge.
- Enamorado, S. (2019a, June 3). *How accurate is your transcription service?*
<https://www.3playmedia.com/blog/how-accurate-is-your-transcription-subtitling-service/#:~:text=The%20industry%20standard%20for%20caption,is%20a%2099%25%20accuracy%20rate>
- Enamorado, S. (2019b, October 7). What is 99% accuracy, really? Why caption quality matters.
<https://www.3playmedia.com/blog/caption-quality/>
- Federal Communications Commission. (2020). *Broadband speed guide*.
<https://www.fcc.gov/consumers/guides/broadband-speed-guide>
- Gernsbacher, M. A. (2015). Video captions benefit everyone. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 195–202.
<https://doi.org/10.1177/2372732215602130>
- Gibbs, G. R. (2018). *Analyzing qualitative data* (2nd ed.). SAGE Publications Ltd.
<https://doi.org/10.4135/9781526441867>

- Google. (n.d.). *Docs editor help: Present slides with captions*.
<https://support.google.com/docs/answer/9109474?hl=en>
- Huang, Y. M., Liu, C. L., Shadiev, R., Shen, M. H., & Hwang, W. Y. (2015). Investigating an application of speech-to-text recognition: A study on visual attention and learning behaviour. *Journal of Computer Assisted Learning*, 31(6), 529–545.
- Huang, Y. M., Shadiev, R., & Hwang, W. Y. (2016). Investigating the effectiveness of speech-to-text recognition applications on learning performance and cognitive load. *Computers & Education*, 101(1), 15–28.
- IBM Cloud Education. (2020, September 2). *Speech recognition*.
https://www.ibm.com/cloud/learn/speech-recognition?mhsrc=ibmsearch_a&mhq=%22word%20error%20rate%22
- Johnson, D. (2020, June 17). *Live Presentations is now generally available*.
<https://www.microsoft.com/en-us/microsoft-365/blog/2020/06/17/powerpoint-live-generally-available/>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.
- Linder, K. (2016). *Student uses and perceptions of closed captions and transcripts: Results from a national study*. Corvallis, OR: Oregon State University Ecampus Research Unit.
<https://www.3playmedia.com/resources/industry-studies/student-uses-of-closed-captions-and-transcripts/>
- Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research*, 26, 1753-1760.
<https://doi.org/10.1177/1049732315617444>
- McCracken, D. G. (1988). *The long interview*. Sage.
- Meyer, A., Rose, D. H., & Gordon, D. (2014). *Universal design for learning: Theory and practice*. CAST. <http://udltheorypractice.cast.org>
- Microsoft Education. (2020, January 6). *Engage your audience with Live Presentations in PowerPoint* [Video]. YouTube. <https://www.youtube.com/watch?v=Lzfqwn05Lzg>
- Microsoft Education Team. (2019, January 23). What's New in EDU Live: Bett day 1. *Microsoft Education Blog*. <https://educationblog.microsoft.com/en-us/2019/01/whats-new-in-edu-live-bett-day-1/#bettday1-a>
- Microsoft. (n.d.-a). *Present Live: Engage your audience with live presentations*.
<https://support.microsoft.com/en-us/office/present-live-engage-your-audience-with-live-presentations-039aa2cc-67fa-4fb5-9677-46ed8a060c8c>

- Microsoft. (n.d.-b). *Present with real-time, automatic captions or subtitles in PowerPoint*.
<https://support.microsoft.com/en-us/office/present-with-real-time-automatic-captions-or-subtitles-in-powerpoint-68d20e49-aec3-456a-939d-34a79e8ddd5f?ui=en-US&rs=en-US&ad=US#OfficeVersion=Windows>
- Morris, K. K., Frechette, C., Dukes III, L., Stowell, N., Topping, N. E., & Brodosi, D. (2016). Closed captioning matters: Examining the value of closed captions for *all* students. *Journal of Postsecondary Education and Disability*, 29(3), 231-238.
- Myers, E. (2019, January 9). *Closed captions & subtitles: Which should you use?*
<https://www.rev.com/blog/subtitles-vs-captions>
- PowerPoint Team. (2018, December 3). *Present more inclusively with live captions and subtitles in PowerPoint*. <https://www.microsoft.com/en-us/microsoft-365/blog/2018/12/03/present-more-inclusively-with-live-captions-and-subtitles-in-powerpoint/>
- Revuelta, P., Jiménez, J., Sánchez, J. M., & Ruiz, B. (2010). Automatic speech recognition to enhance learning for disabled students. In J. Zhao, P. Ordoñez De Pablos, & R. Tennyson (Eds.), *Technology enhanced learning for people with disabilities: Approaches and applications* (pp. 89-104). IGI Global. ProQuest Ebook Central.
<http://ebookcentral.proquest.com/lib/novasoutheastern/detail.action?docID=3310777>
- Shadiev, R., Huang, Y-M., & Hwang, J-P. (2017). Investigating the effectiveness of speech-to-text recognition applications on learning performance, attention, and meditation. *Educational Technology Research and Development*, 65(5), 1239-1261.
<http://doi.org/10.1007/s11423-017-9516-3>
- Shadiev, R., Sun, A., & Huang, Y-M. (2019). A study of the facilitation of cross-cultural understanding and intercultural sensitivity using speech-enabled language translation technology. *British Journal of Educational Technology*, 50(3), 1415-1433.
<https://doi.org/10.1111/bjet.12648>
- Shadiev, R., Wu T-T., Sun A., & Huang Y-M. (2018). Applications of speech-to-text recognition and computer-aided translation for facilitating cross-cultural learning through a learning activity: Issues and their solutions. *Educational Technology Research and Development*, 66(1), 191-214. <https://doi.org/10.1007/s11423-017-9556-8>
- Take Note. (n.d.). *Closed captioning vs. subtitles: How to make the right choice*
<https://takenote.co/closed-captioning-vs-subtitles/>
- Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27, 237-246.
- United States Department of Justice Civil Rights Division. (n.d.). *Information and technical assistance on the Americans with Disabilities Act*. <https://www.ada.gov/>
- U.S. General Service Administration. (n.d.). *IT accessibility laws and policies*.
<https://www.section508.gov/manage/laws-and-policies>

World Wide Web Consortium. (n.d.). *Web Content Accessibility Guidelines (WCAG) 2.0*.
<https://www.w3.org/TR/WCAG20/>

Young, D. A., Casey, E. A. (2019). An examination of the sufficiency of small qualitative samples. *Social Work Research*, 43(1), 53–58. <https://doi.org/10.1093/swr/svy026>

Zoom Video Communications. (n.d.). System requirements for Windows, macOS, and Linux.
https://support.zoom.us/hc/en-us/articles/201362023-System-requirements-for-Windows-macOS-and-Linux#h_d278c327-e03d-4896-b19a-96a8f3c0c69c