4-1-2024

Students’ Experiences When Using Real-Time Automated Captions and Subtitles in Live Online Presentations: A Phenomenological Study

Anymir Orellana Ed.D.
Nova Southeastern University, orellana@nova.edu

Elda Kanzki-Veloso Ph.D.
Nova Southeastern University, veloso@nova.edu

Georgina Arguello Ed.D.
Nova Southeastern University, deheredi@nova.edu

Katarzyna Wojnas
Nova Southeastern University, kw1583@mynsu.nova.edu

Follow this and additional works at: https://nsuworks.nova.edu/tqr

Part of the Curriculum and Instruction Commons, Disability and Equity in Education Commons, Educational Psychology Commons, Educational Technology Commons, Online and Distance Education Commons, and the Quantitative, Qualitative, Comparative, and Historical Methodologies Commons

Recommended APA Citation

This Article is brought to you for free and open access by the The Qualitative Report at NSUWorks. It has been accepted for inclusion in The Qualitative Report by an authorized administrator of NSUWorks. For more information, please contact nsuworks@nova.edu.
Students’ Experiences When Using Real-Time Automated Captions and Subtitles in Live Online Presentations: A Phenomenological Study

Abstract
According to the Universal Design for Learning (UDL) framework, as a text-based alternative to auditory information in videos or presentations, captions can make the content equally accessible, and multilingual subtitles can promote a cross-linguistic understanding of the content. We conducted a phenomenological study to understand the common meaning of the participants’ experiences when using real-time automated captions/subtitles during live online class presentations. Twenty-four remote student participants were placed in three study groups. All participants were fluent in spoken and written English, eight could read in one or more additional languages, and none had a hearing disability. We used Microsoft PowerPoint Present Live via Zoom to deliver the online presentation to each group with real-time automated captions/subtitles, and then we conducted a focus group session with each group. Ten themes emerged and were clustered into three overarching themes: challenges, benefits, and interactions with subtitles. Overall, participants described a positive experience, perceiving the captions/subtitles as useful and accurate. Participants found the tool easy to use and highlighted the benefits of using captions/subtitles, such as providing access to live instruction for a wide audience and reinforcing learning for diverse student types. While they were able to troubleshoot connectivity and technological issues encountered, they experienced an apparent split-attention effect and noted limitations in the tool’s inability to recognize different dialects. Findings contribute to educational research related to accessible live instruction in multilingual settings and could aid educators in selecting and integrating tools with real-time captioning/subtitling, in line with the UDL guidelines.

Keywords
Universal Design for Learning, real-time, captions, multilingual subtitles, phenomenology, automatic speech recognition, online presentation, inclusive instruction, PowerPoint Present Live

Creative Commons License
This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 4.0 International License.

Acknowledgements
This study was funded by Nova Southeastern University's 2021-2022 President's Faculty Research and Development Grant, scholar track, awarded to the project titled “Students’ Experiences Using Live Captions and Subtitles in Class Presentations.”

This article is available in The Qualitative Report: https://nsuworks.nova.edu/tqr/vol29/iss4/7
Students’ Experiences When Using Real-Time Automated Captions and Subtitles in Live Online Presentations: A Phenomenological Study

Anymir Orellana, Elda Kanzki-Veloso, Georgina Arguello, and Katarzyna Wojnas
Nova Southeastern University, Fort Lauderdale, Florida, USA

According to the Universal Design for Learning (UDL) framework, as a text-based alternative to auditory information in videos or presentations, captions can make the content equally accessible, and multilingual subtitles can promote a cross-linguistic understanding of the content. We conducted a phenomenological study to understand the common meaning of the participants’ experiences when using real-time automated captions/subtitles during live online class presentations. Twenty-four remote student participants were placed in three study groups. All participants were fluent in spoken and written English, eight could read in one or more additional languages, and none had a hearing disability. We used Microsoft PowerPoint Present Live via Zoom to deliver the online presentation to each group with real-time automated captions/subtitles, and then we conducted a focus group session with each group. Ten themes emerged and were clustered into three overarching themes: challenges, benefits, and interactions with subtitles. Overall, participants described a positive experience, perceiving the captions/subtitles as useful and accurate. Participants found the tool easy to use and highlighted the benefits of using captions/subtitles, such as providing access to live instruction for a wide audience and reinforcing learning for diverse student types. While they were able to troubleshoot connectivity and technological issues encountered, they experienced an apparent split-attention effect and noted limitations in the tool's inability to recognize different dialects. Findings contribute to educational research related to accessible live instruction in multilingual settings and could aid educators in selecting and integrating tools with real-time captioning/subtitling, in line with the UDL guidelines.

Keywords: Universal Design for Learning, real-time, captions, multilingual subtitles, phenomenology, automatic speech recognition, online presentation, inclusive instruction, PowerPoint Present Live

Introduction

Educators strive to deliver instructional strategies that effectively address the diversity of learners in their classrooms (e.g., learners with or without disabilities, native and non-native speakers of the language of instruction). A framework that can be used to design instruction and curricula to address diversity is the research-based framework Universal Design for Learning ([UDL] CAST, 2018a; Meyer et al., 2014). UDL addresses diversity as a function of learner variability in how they engage in learning, perceive and comprehend information, and express what they know. Research findings related to the effectiveness of UDL on learning have been mixed partly because of the variability in how researchers have reported
relationships between specific UDL guidelines and their interventions (Ok et al., 2017); however, overall, the findings suggest that UDL can be an effective framework to design flexible learning environments to benefit all types of learners (Al-Azawei et al., 2016; Capp, 2017; Seok et al., 2018; Shreffler et al., 2019). The UDL framework is also referenced in US-based education policies, such as those found in §1003(24) of the Higher Education Opportunity Act (2008), where it is defined as a “scientifically valid framework for guiding educational practice” (p. 182).

According to the UDL guidelines, because learners perceive and comprehend information in multiple ways, options for perception should be provided “to ensure that key information is equally perceptible to all learners by … providing the same information through different modalities (e.g., through vision, hearing, or touch); [and] providing information in a format that will allow for adjustability by the user” (CAST, 2018b, para. 1). For auditory information, alternatives can be offered by using “text equivalents in the form of captions or automated speech-to-text (voice recognition) for spoken language” (CAST, 2018c, para. 2). Additionally, “for new learners of the dominant language … or for learners of academic language… the accessibility of information is greatly reduced when no linguistic alternatives are available [and] providing alternatives, especially for key information or vocabulary is an important aspect of accessibility” (CAST, 2018d, para. 1). Thus, according to the UDL guidelines, as a text-based alternative to auditory information in live videos or presentations, captions can make the content equally accessible to many and multilingual subtitles can promote a cross-linguistic understanding of the content.

Captions and multilingual subtitles can be generated by a human or with automatic speech recognition (ASR) technology (SRT). The accelerated advancements in the field of ASR have allowed many to access easy-to-use, free, or inexpensive SRTs and have enabled application developers to continually enhance or integrate features of automated captions/subtitles in various tools such as mobile devices, videoconferencing systems, computer browsers, and presentation applications. Furthermore, the COVID-19 pandemic has caused an increased demand for groups to work, meet, or learn remotely using live videoconferencing systems (Hacker et al., 2020). Thus, it could be anticipated that the use and access of tools with the feature of real-time automated captions/subtitles could become pervasive in the online classroom. However, the proper use of such feature requires an understanding of its limitations and effectiveness, how educators can effectively use it and how easily, and how students with and without disabilities experience and learn using it in different instructional scenarios, including live online presentations.

In this phenomenological study, we sought to understand the common meaning of the participants’ experiences when using real-time automated captions/subtitles during live online class presentations. The participants were remotely located and used Microsoft PowerPoint Present Live (PPT Live) via Zoom to experience the presentation with real-time automated captions/subtitles. In this study, we referred to captions as the transcription of the presenter’s spoken words without sound description or speaker identification, and to subtitles as the transcription in a different language.

Background

Video captions are the equivalent text that describes “not only spoken content but also non-speech information such as sound effects, music, laughter, and speaker identification and location” (Morrin et al., 2014, What are Captions section). Video captioning has been shown to benefit many, including those who are deaf or hard of hearing (DHH), hearing adults wanting to increase their retention of verbal information, and individuals learning a second language (Dallas et al., 2016; Gernsbacher, 2015; Linder, 2016; Morris et al., 2016). Captions not only
help individuals with cognitive disabilities but also foster literacy among children and adults (Morrin et al., 2014). Captions are necessary for those who are DHH and 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.).

Subtitles are different from captions. Captions are in the same language as the spoken one, typically used for accessibility purposes, such as aiding viewers with hearing impairments, or in environments where audio cannot be heard. Subtitles, on the other hand, are translations of the spoken text into another language, commonly used in movies to help viewers who speak a different language understand the dialogue (Morrin et al. 2014). In general, captions are needed for accessibility accommodations, and subtitles in other languages are not (Henry, 2022a).

Video captions/subtitles can be generated live in real-time or added in post-production time, and they can be generated by a human or with SRT. SRT-generated automated captions and subtitles typically have error rates that are unacceptable in most lecture environments and in many unconstrained real-world environments (Kushalnagar et al., 2014). Nonetheless, to address the diversity of learners and promote inclusive live classroom lectures or presentations, providing real-time human-generated captions/subtitles can be expensive or not available on demand; thus, the use of SRTs could be a cost-effective approach when it would be otherwise necessary to hire dedicated staff (Revueulta et al., 2010).

Captions/Subtitles with Automatic Speech Recognition (ASR) Technology, or SRT

Research in ASR has grown for over 60 years and has contributed to developing a broad range of valuable and pervasive speech recognition products (Li et al., 2015). In 2015, Linn explained that the biggest hurdle to be solved was for the SRT to comprehend what a person is saying in non-controlled environments. Similarly, according to Linn (2015), SRT was still not very good in environments that were noisy, crowded, or had echoes, especially when using subpar equipment like low-quality microphones. Additionally, SRT struggled when people spoke rapidly or softly, had accents, or in comprehending speech from children and elderly speakers (Linn, 2015).

In 2017, Microsoft’s speech and dialog research group announced they had reached a 5.1% error rate, or a word accuracy of 94.9%, for the English language with their speech recognition system. This error rate was described as an industry milestone that surpassed the accuracy achieved in 2016 (Huang, 2017). Hannun (2017) described the advancements in conversational speech by the SRT as remarkable yet noted that assertions of its performance being on par with humans were overly generalized. According to Hannun, some areas in ASR that needed improvement were accents and noise, semantic error, single-channel multi-speaker context, and deployment.

The years 2020 and 2021 showed accelerated advancements in ASR technology, which not only enhanced the performance of existing applications like the virtual assistants Siri and Alexa, but also broadened the range of markets that ASR technology serves (Jetté & Miller, 2022). Jetté and Miller envisioned ASR developments by the year 2030 and concluded that many were already under way, including the improvement of accuracy; the development of multilingual models to create applications capable of understanding anybody and in any language; richer transcripts; and fast, affordable, reliable, and private ASR.

Although SRTs have limitations to overcome to fully comprehend what a person is saying, students have found SRTs beneficial for learning. For example, students in Shadiev et al.’s (2017) study found speech-to-text recognition technologies helpful in aiding their
comprehension and attention, enhancing learning, and awareness. Students have also found SRT beneficial for cross-cultural learning activities (Shadiev et al., 2018; Shadiev et al., 2019). SRT can also provide a transcript of what the instructor presents in real-time (Revuelta et al., 2010) and help non-English-speaking students to better understand a lesson, take notes, and confirm what was being said in the lecture (Huang et al., 2015; Huang et al., 2016).

The continuous advances in ASR have allowed many to easily access real-time automated captions/subtitles with common, inexpensive, or free applications with different but comparable capabilities. Video platforms, presentation applications, and videoconferencing applications have enabled users to access automated captions, transcripts, or multilingual subtitles generated in real-time with SRT. For example, in 2009 YouTube (Alberti & Bacchiani, 2009) launched the automated captions feature for recorded videos, and in 2021 automatic livestream captioning became available for all YouTube creators (Clark, 2021; Google, n.d.-c). Features in the Google Chrome browser (Google, n.d.-d; Scharff & Kompalli, 2021) provide real-time automated captions for recorded audio and video, played with or without volume. Android-based mobile devices, as well as iPhones, allow captioning and translation of live speech (Apple Inc., n.d.; Google, n.d.-e). Videoconferencing systems such as Google Meet (Google, n.d.-b), Zoom (Zoom Video Communications, Inc., 2022a), Webex (2022), and Microsoft Teams (Microsoft, n.d.-b; O’Sullivan, 2022; Tung, 2020) provide real-time automated captions of presenter’s and participants’ words during an online videoconference. Presentation applications such as Google Slides (D’Aquila & Klein, 2018; Google, n.d.-a) and Microsoft PowerPoint ([PPT] PowerPoint Team, 2018) provide real-time automated captions/subtitles for the presenter’s spoken words.

Automated Real-Time Captions/Subtitles in PPT Live

In 2018, the PowerPoint Team (2018) announced the real-time automated captions/subtitles feature as one 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” (para. 2). The feature is one of the cloud-enhanced Microsoft 365 features powered by Microsoft Speech Services and, to provide the service, the speech utterances are sent to Microsoft (Microsoft, n.d.-a). PPT uses cloud-based speech recognition for real-time captioning of the presenter’s spoken words (Microsoft, n.d.-a). As of late January 2019, the feature has been available for Microsoft 365 subscribers worldwide for PPT on Windows 10, PPT for Mac, and PPT Online.

In January 2020, Microsoft announced the PPT Live feature in Microsoft 365 (Microsoft Education, 2020) that became available on PPT for the web by June 2020 (Johnson, 2020). The PPT Live feature allowed real-time automated captions, the real-time translation of 12 spoken languages into more than 60 languages, the possibility of individual viewers, online or on-site, 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). For online viewers wanting to hear the presenter, PPT Live could be used with a conferencing system that transmits audio, such as Zoom or Microsoft Teams (Microsoft, n.d.-c).

PPT is an application commonly used for presentations in educational settings; thus, it could be expected that educators with a Microsoft 365 license would be inclined to use it for their live lectures or presentations, on-site or online. Presenters can use PPT to show captions or subtitles to all in the audience or use PPT Live to give the individual viewer in the audience a choice to access the captions/subtitles and in their preferred language. As with any use of technology for learning, effective integration of PPT Live requires proper planning. In general, the process of selecting proper media and technologies for learning is complex and involves consideration of many interacting variables. Among the criteria to select appropriate
technologies for learning, it is important to determine the cost-benefit, the ease-of-use, and the reliability of the technology (Bates, 2022; Smaldino et al., 2019). When integrating SRT, such as PPT Live’s caption/subtitle feature, it is also necessary to determine the accuracy and intelligibility of the generated captions and subtitles.

To assist educators in the proper selection of SRTs for learning, prior to the current study, we conducted research to determine the effectiveness and faculty’s perceptions of PPT Live’s feature of automated real-time captions/subtitles in English and Spanish (see Orellana et al., 2021, 2022). We computed the word accuracy of the captions and the intelligibility of the subtitles as translation of the captions in English and Spanish (see Orellana et al., 2021). We conducted a qualitative moderated usability testing interviewing five English- and five Spanish-speaking faculty to identify challenges, potential uses, and benefits when presenting with PPT Live (see Orellana et al., 2022). It is worth noting that, aside from Orellana et al.’s studies (2021, 2022), no other peer-reviewed studies that used the PPT Live’s feature of captions/subtitles had been found in the literature by the time this current study was conducted.

Based on our previous study findings, we cautiously concluded that automated real-time captions/subtitles generated with PPT Live in English/Spanish could be useful to implement UDL guidelines to provide a text alternative to the spoken language when the viewer could hear the presenter and is fluent in the presenter’s spoken language (see Orellana et al., 2021, 2022). Additionally, we recommended further exploring the use of automated captions/subtitles in “more realistic scenarios where the presenter is familiar with the content, uses slides with visuals and related text, and speaks spontaneously” (Orellana et al., 2021, p. 38). The goal of this current study was to further the research related to accessible multilingual presentations, in line with the UDL guidelines, by exploring how students experience automated captions and subtitles in a live scenario that resembles an online class presentation experience.

Research Purpose and Questions

The purpose of this study was to understand the common meaning of the participants’ experiences when using real-time automated captions/subtitles during live online class presentations. Participants acted as remote learners in an online class listening to a live presentation delivered in English, and they chose to follow the presentation in one or more preferred languages.

The main research question that guided the study was, “What are participants’ experiences when using real-time automated captions/subtitles in live online class presentations?” We examined the following sub-questions:

1. What meaning do participants give to their experience when using the feature of real-time automated captions/subtitles in live online presentations?
2. What meaning do participants give to their experience when reading real-time automated captions in the presenter’s language or subtitles in other languages during live online presentations?

Researchers’ Relationship to the Topic

This study reflects our interest in inclusive educational practices through technology, in line with the UDL framework, extending our prior research related to the effectiveness and faculty perceptions of automated real-time captions and subtitles (see Orellana et al., 2020, 2021). Our areas of expertise span instructional design, media and technology, counseling and
mental health, psychology, research, and educational leadership. Our team included full-time faculty members, an administrator in bilingual and international programs, and a doctoral student in clinical psychology. We have an extensive background, with each of us having over 15 years of experience in teaching across various formats—online, on-site, and hybrid—across different academic disciplines and levels, serving a diverse student body. In addition to serving a diverse student body, the diverse linguistic backgrounds within our team, with three members being non-native English speakers and one being a bilingual native English speaker, were key in shaping our interest in exploring how SRTs can be effectively integrated into educational settings, particularly those that are multilingual.

Method

Design

We conducted a phenomenological study to explore the common meaning of participants' experiences when using PPT Live’s features of real-time automated captions and multilingual subtitles. We expected to explore participants’ interactions—among themselves and with the technology—that typically occur in an online classroom during a live presentation delivered by an instructor. We also anticipated that these interactions would shape and provide context to the participants’ collective experience.

Phenomenology helps convey what participants have in common as they experience a phenomenon (Creswell & Poth, 2018; Smith et al., 2009). Phenomenological methods have been associated with educational communications and technology (ECT) for years to provide insight into studying users' experiences with technology (Valentine et al., 2018). By conducting a phenomenological study, researchers can understand the collective group experience of students using ECT within a specific social context, such as the classroom. Furthermore, phenomenology is a suitable approach for studies with ECT because it allows researchers to explore how a group of individuals experience and interact with the phenomenon (Cilesiz, 2011). In the context of ECT, phenomenology helps convey what participants have in common as they experience a technology that offers them the option to engage in a live online class presentation with real-time automated captions and multilingual subtitles. Additionally, phenomenological studies can be successfully conducted with variable groups and still allow for the individual voices of participants to emerge (Arroll, 2015; Coyle, 2014; Love et al., 2020).

Participants

Participants were students from two colleges from the institution setting of the study. The institution was a private not-for-profit university in the southeastern United States. The institution was considered a majority-minority university with a diverse student population from more than 100 countries and was also recognized as a Hispanic-serving institution, with more than 25% of its students identified as Hispanic. The colleges delivered various online programs and, due to the COVID-19 pandemic, they had also been delivering online classes that were typically delivered on-site.

Upon IRB approval, the principal investigator (PI) emailed the instructors from two selected colleges a recruitment message and flyer to distribute to their students. Participants’ inclusion criteria were being able to read and speak English, being over 18 years of age, residing in the U.S. or one of its territories, having a computer with an Internet connection, and being enrolled in a program in one of the two colleges.
Forty-one students emailed the PI indicating interest in the research study, 35 consented to participate, and 24 participated in the study: 22 females and 2 males, between 20 and 69 years old. All were fluent in the English language; eight indicated that they could read fluently in another language (Spanish, 6; other, 2); and one indicated that they could read fluently in two other languages (French and other). Seventeen were enrolled in on-site programs (16 master’s and one undergraduate), and 7 in online doctoral programs. All participants reported being almost very confident in using the basic technologies needed for the session (i.e., Zoom for viewing and listening to an online class presentation, a mobile device for reading QR codes, an Internet browser for opening two windows side by side on a computer screen). Before participating in the study, 16 participants had never used PPT Live, five had used captions/subtitles as a viewer able to select and read the subtitle language of their choice, and four as a viewer reading the captions on screen.

Data Collection Procedures

The 24 participants were grouped in three study presentation sessions based on their preferred day: Group 1, 4 participants; Group 2, 5 participants; and Group 3, 15 participants. We conducted a focus group session (FG) after each presentation to capture participants’ immediate recollections of their experience. The use of FG allows a rich and unique experience that cannot be achieved in an individual interview (Hall, 2020). According to Farnsworth and Boon (2010), the use of FG can provide real-time observations of interactions between group members which can be shaped by their experiences and opinions.

We used PPT to create the presentation, PPT Live to generate the automated captions/subtitles, and Zoom to connect the presenter, FG moderators, and participants. We chose these technologies because they were accessible to the participants and commonly used in online or on-site classes. During each presentation session, the PI welcomed participants, read the consent and purpose of the study, introduced the presenter, played a 3-minute video tutorial demonstrating how to connect to the PPT Live presentation with their devices and how to follow the captions/subtitles in real-time, and asked the participants to follow the presentation with subtitles in any of the languages that they could read fluently. Once all participants were connected to PPT Live, the presenter delivered in approximately 20 minutes the presentation titled “Working with Individuals & Families in Haitian Culture: Cultural Considerations.”

Upon completion of the presentation session, the PI asked the participants to take approximately 5 minutes to report what devices they used for reading captions/subtitles, the subtitle and transcript languages they selected, their perception of the subtitles' intelligibility, the usefulness and ease of use of the captions/subtitles feature, and their confidence in using the feature. Upon participants' reporting the information, one of the researchers moderated the FG via Zoom following the FG protocol. Each FG lasted approximately 45 minutes and was recorded to allow the compilation, verification, and analysis of the Zoom-generated transcripts.

Presentation Session with PPT Live via Zoom

Participants could join the PPT Live presentation with captions/subtitles with their mobile device by reading the QR code, or by entering the session link with the mobile browser or in a separate computer browser window next to the Zoom window. Participants were able to listen to the presenter via Zoom, view the presentation via their computer screen on Zoom or via their mobile device, and read the transcripts in their preferred language with their computer second browser window or with their mobile. Figure 1 shows the view of how participants would access the transcripts when using their mobile, and Figure 2 shows the view...
when using another browser next to their Zoom window. By default, the transcript language is the presenter’s language that each viewer could change at any point in time by selecting the icon located on bottom left of the transcript screen.

**Figure 1**
Transcript (caption subtitle) language options in a PPT Live session using a mobile. The transcript language could be changed by clicking the icon on the bottom left of the mobile screen.

**Figure 2**
Two browser windows next to each other: the left with the Zoom session to connect the audience with audio and the right with the PPT Live presentation and transcript language screen with the subtitles in English. The transcript/subtitle language can be changed by selecting the icon on the bottom left of the screen. Note: The transcript screen would be next to, or under, the slides depending on the browser window's width.
Data Analysis

We followed Colaizzi’s method (1978) to collect the groups' interactions and to analyze the data to achieve interaction connections between the groups. A phenomenological design was appropriate for analyzing the data collected through focus groups, as it allows the researchers to understand the shared experience of participants (Ayton et al., 2023; Bliss, 2016; Bush et al., 2019; Ho, 2006). We reviewed the transcripts of each FG session for accuracy, coded the data, and analyzed for repetitive words or phrases to create categories using open coding to organize similarities and differences based on participants’ responses. Themes started to emerge from the created categories. We followed the steps below to ensure a comprehensive analysis of the data and maintained rigor in our interpretation of participants' experiences:

1. Familiarization. We thoroughly read and re-read transcripts to gain a better understanding of semantic content and language use.
2. Noting Key Ideas. We took notes and highlighted key ideas or statements that stood out.
3. Extracting Significant Statements. We identified significant statements or phrases capturing participants' experiences related to research questions through open coding.
4. Formulating Meanings. We analyzed each significant statement individually to identify underlying meanings or concepts.
5. Identifying Commonalities. We sought commonalities, patterns, or emerging themes within the identified meanings.
6. Clustering Themes into Categories. We grouped similar meanings together based on shared characteristics to form categories that capture distinct aspects of participants' experiences.
7. Developing Descriptions. We wrote descriptions and provided rationale for each theme, including quotes from participants' experiences.
8. Verification. We utilized member checking or peer debriefing to confirm the accuracy of the analysis in representing participants' experiences.

Ethical Considerations

We recruited the participants upon IRB approval. Because participants were adult students with different professional and academic schedules, we provided alternative research session days and times to accommodate their schedules. To ensure proper remote access to the research sessions, we selected a population of students that had an account and experience with Zoom, had internet connection, and had the minimal technological skills and equipment to access PPT Live’s captions/subtitles. To minimize participants’ difficulty with PPT Live’s captions/subtitles, before each presentation session we showed a 3-minute video tutorial created by the PI to demonstrate how to access, follow, and select the transcript language with their devices.

To minimize distractions related to the presentation content, we chose a general topic that would be appealing to participants. To maximize the accuracy of the captioning of the presenter’s spoken words, we ensured that the presenter (a) had and could properly use the equipment (i.e., microphone, computer, fast and reliable internet connection) and applications (i.e., PPT Live, Zoom); (b) was the same who delivered the same presentation to the three groups of participants; (c) was a native English speaker; and (d) had rehearsed the presentation several times so that she took approximately 20 minutes to deliver the content, spoke in her
normal accent and pace and intelligibly, and avoided phrases and special terms that PPT Live might not caption properly.

Findings

Participants’ Use of PPT Live’s Caption/Subtitle Feature During the Presentation Sessions

Participants indicated that they selected various subtitle languages to read during the presentation session: nine indicated English and another language, one Spanish only, and 14 English only. The subtitle languages selected by participants were Spanish (7), French (1), Italian (1), Portuguese (1), and other (Malay, Indonesian, Romanian, and Vietnamese). Participants indicated that during most of the session time, they followed the presentation in English (19), Spanish (4), and Malay and Indonesian (1). When asked how well they were able to understand the subtitles in the language that they followed during most of the presentation session time (from 1 = Not at all to 5 = Very well), 23 participants rated as a 4 or 5; and one rated as Not at all, being English their selected language. When asked how useful they found PPT Live’s feature of captions/subtitles and how easy it was to use (from 1 = Not at all to 5 = Very much) participants’ average rating was 4.12 and 4.54, respectively.

To connect to PPT Live to read the captions/subtitles, six participants only used the computer that had Zoom. Eighteen participants used one or more personal devices to connect to PPT Live to read the captions/subtitles, in addition to the computer they used to connect to the Zoom session, as follows: iPhone with iOS 11+ (8), smartphone with Android OS (8), Apple iPad with iOS 11+ (2), Windows computer/laptop (1), and Apple MacBook or iMac (6). Table 1 shows the participants’ reported confidence levels when using the technology during the presentation session.

Table 1
Participants’ Confidence Level When Using the Technologies During the PPT Live Presentation (N = 24)

<table>
<thead>
<tr>
<th>Technology</th>
<th>Average</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zoom to view and listen to the presentation</td>
<td>4.739</td>
<td>24</td>
</tr>
<tr>
<td>Mobile (smartphone or tablet) for reading the QR code to connect</td>
<td>4.47</td>
<td>17</td>
</tr>
<tr>
<td>Two internet browser windows (e.g., Google Chrome, Safari, MS Edge) side</td>
<td>4.8</td>
<td>17</td>
</tr>
<tr>
<td>Another computer with the internet browser (e.g., Google Chrome, Safari,</td>
<td>4.75</td>
<td>12</td>
</tr>
</tbody>
</table>

Emergent Themes

Ten themes emerged to explain participants’ experience when using real-time automated captions/subtitles in live class presentations. Three overarching themes emerged:
challenges, benefits, and interactions with subtitles. Following is a description of the overarching and the subsidiary themes along participants' responses that support the themes.

Challenges

This overarching theme emerged from challenges that the participants encountered when using PPT Live’s caption/subtitle feature during the live presentation, along with the strategies they employed to address challenges. We defined challenges as encompassing technical problems that arise when using live cloud-based technologies (e.g., internet or network connectivity issues); technological limitations inherent to the caption/subtitle feature capabilities (e.g., accuracy, placement of subtitles and transcript in mobile and computer, scrolling feature, font size) or behavior (e.g., transcript speed, synchronization); and difficulties related to the multimedia presentation with live captions/subtitles.

Connectivity. This theme emerged from participants' comments about experiencing connectivity problems while accessing the captions/subtitles during the live presentation. The theme highlights technical challenges participants faced regarding internet stability and network coverage when accessing captions/subtitles during live presentations.

Participants reported instances where unstable internet connections led to brief disconnections from the live session, temporarily disrupting their ability to follow along. Participants employed various strategies to re-establish their connection and continue participating in the presentation. For example, one participant recounted their experience of losing connection and the subsequent steps taken to rejoin the session: "I also had technical difficulties. Like I lost connection, and when I lost connection, it told me to reset it, and then I clicked it, reset it, and then it asked me for a code on the top of the slide, but I couldn't find the code."

Others shared similar experiences of connectivity hiccups that momentarily halted their access to the presentation: "I did experience a hiccup on mine and had to refresh, it froze up and hit refresh one time, but I could see it being useful." Another participant attributed the connectivity issues to their location, suggesting environmental factors played a role: "Mine froze up too, I just thought it might have been my home Wi-Fi, but it came right back itself;" and "[my connectivity issues were] because I’m in Puerto Rico right now, or maybe because where I am personally now, which is kind of the mountains."

Despite these challenges, overall, participants did not express significant frustration, highlighting their ability to reconnect by refreshing their screens or waiting for the connection to re-establish. The experiences shared by participants reflect broader considerations about the reliability of technology in facilitating seamless access to digital content.

Application Features. This theme emerged from participants' discussion regarding difficulties trying to keep up with the presentation when reading transcript languages. These difficulties were inherent to the behavior of PPT Live’s feature of captions/subtitles. One participant expressed, “It was hard for me catch up while I switched the language, there was a gap in catching up with the presentation.” As participants experienced difficulties related to the elapsing time when switching languages, they discussed features that they thought may be worth adding or modifying in PPT Live to address the issue. For example, a participant explained the need for a feature that had been available in the application, such as the ability to scroll the screen to read the transcripts: “The only thing that I would think of that might improve the system is so that we can scroll up and see what has already been said about five seconds ago 10 seconds.” Another participant identified having difficulty with slow translation: “The same language was easy but switching to Spanish there was lagging on the comments, and it wouldn't be at the same speed.” On the other hand, one participant experienced the subtitles as too fast: “The subtitles were going way too fast and at some point, due to the context
of the paragraph, for the sentence every word everything, so I had to let it go... but only if you have two languages.”

The insights shared by participants suggest the importance of adaptable and user-friendly captions/subtitle features to enhance the accessibility and effectiveness of the presentation. The feedback points to specific areas for improvement of the tool’s performance from the view of the participants which could impact their user experience, including language-switching delays, subtitle navigation, translation responsiveness, and subtitle pacing.

**Inaccurate Translation due to Dialectal Variations.** This theme emerged from participants’ comments about inaccurate translations resulting from dialectal variations. This limitation is inherent to the capabilities of PPT Live’s feature of captions/subtitles. Participants noted inaccuracies in translations due to dialectal variations in languages, particularly Spanish. Spanish was the other language that several participants could read fluently (n = 6). Spanish-speaking participants noted that the Spanish-language dialect from their respective country was not available as a subtitle language and expressed concerns that some of the translated words were inaccurate due to dialectal variations. For example, one participant stated: “I am from Peru [and] there are words that were different that I don’t understand, there are words in my language that were different.” Another participant stated, “The words are different in Spanish in Puerto Rico.” Some participants expressed that PPT Live was unable to recognize dialects of Spanish or other languages, potentially leading to not being able to understand parts of the presentation. One participant explained that “Spanish-speaking individuals speak different Spanish than in Spain, same with Puerto Rico and maybe even Argentina, I know that there are very different words as well [and] that can be a problem.”

These experiences highlight an awareness among participants of the linguistic diversity within languages, particularly in the Spanish language, and the challenges it presents for accurate translation in digital tools. Additionally, the participants’ reflections suggest a need for enhanced technological solutions that can more effectively accommodate different dialects, ensuring greater accessibility and inclusion.

**Troubleshooting.** This theme emerged from participants’ expressing their strategies for maintaining engagement with the live presentation captions/subtitles in situations such as reconnections or when switching transcript languages. Participants detailed various troubleshooting techniques they employed, including refreshing the screen, switching devices, scrolling through the transcript screen to catch up on missed content, and taking screenshots of the transcripts for later review. Some participants shared their initial lack of knowledge regarding these troubleshooting methods, noting moments of realization about solutions like screen refreshment. Examples of participant comments include: “I guess I didn't know that I could just refresh my screen; maybe they should have told us at the beginning of the session,” and “I just refreshed, and it brought me back to the exact slide.” Additionally, the exchange of troubleshooting tips among participants through Zoom’s chat feature was noted, with advice such as “Just refresh [the mobile device screen or computer browser].” Overall, these actions highlight participants’ resourcefulness and the role of peer support in resolving technical challenges.

**Reading the Subtitles While Following the Presentation can be Difficult.** This theme emerged from participants discussing difficulties when listening to the presentation in a different language from the presenter’s while reading and looking at visuals in the PPT slides and reading the subtitles. One participant explained, “Only was difficult to follow [the subtitles] when listening to a different language than the presenter because I’m a slow reader in a language that is not my own.” Another participant explained the difficulty with understanding multiple languages:
It was just really hard because you know both languages, so when your brain is trying to process something that you are hearing, and then your other part of your brain is trying to process what you’re reading [and] it could be hard.

Participants’ descriptions indicate that they found it challenging to process auditory and visual information simultaneously, especially when these were in different languages.

Benefits

This overarching theme emerged from the benefits or advantages of using captions/subtitles in live presentations with PPT Live, as reported by participants.

**Access to Live Instruction for Various Types of Learners.** This theme emerged from participants discussing how the tool could allow access to live instruction to diverse types of learners. Participants identified the diverse types of learners as those with or without disabilities, speakers and non-speakers of the language of instruction, and those in face-to-face settings or online. Participants highlighted the utility of the tool in supporting diverse learning preferences and needs. One participant articulated how the tool serves as a multifunctional educational aid, especially for students with disabilities.

I think if we use it for students who were speaking the same language as the presenter or also, like, I said, that fact that it reinforces learning because you’re hearing the person talk and you’re reading the PowerPoint, or you’re looking at the PowerPoint slides, and then you’re reading the transcription, so I think it also keeps the student [with disabilities] focused.

This insight highlights the tool's role in reinforcing learning through simultaneous auditory and visual engagement, catering to various sensory preferences and enhancing focus.

Another participant expressed the benefit of using the real-time captions in face-to-face settings, emphasizing the tool's capacity to diminish feelings of intimidation and foster a more inclusive learning environment: “If they are coming into a classroom using one on one computers and they have a teacher speaking, [each student can] feel less intimidated … [when they] are able to [read the translation] in real time…. The student can see the subtitle in real time; it is extremely beneficial.” This perspective highlights the tool's potential to personalize the learning experience, allowing students to engage with live instruction in a manner that aligns with their individual comprehension needs and language proficiencies. Through these accounts, participants illustrated the tool's versatility in addressing the varied demands of live instruction across diverse educational contexts.

**Reinforcing Learning for all Types of Learners.** This theme emerged from participants’ discussion on how captions/subtitles could be helpful for all learners in the reinforcement of spelling, grammar, and proper grammatical usage. Participants appeared excited explaining that real-time captions/subtitles could reinforce learning of spelling and grammar, not only for students learning a different language but also for those who speak the same language as the subtitles. One participant explained, “It would be fantastic for any learner of another language that's trying to learn in a spelling, also diverse population we don't have any access to.” Other participants thought that real-time captions were beneficial for students who are struggling with grammar and how it can be a useful tool to reinforce the learning of the accurate and proper grammatical usage: “For students who have problems with grammar, it would be useful to see it [because] it reinforces learning.” Another participant explained a scenario where captions could be beneficial for reinforcing learning: “If you're given a lecture,
even if you are in that same language as that person, sometimes, they might miss a word or two and sometimes they just need to see it."

Through these reflections, participants illustrated the role of captions/subtitles in reinforcing language learning. The feature was recognized as a tool for improving spelling and grammar comprehension, offering a visual reinforcement of auditory information, and catering to the needs of learners across a spectrum of language proficiencies.

**Useful and Accurate Captions/Subtitles and Convenient and Easy-to-use Tool.** This theme emerged from participants expressing appreciation of a captioning/subtitling tool that provided mostly accurate transcriptions in real-time, making them a useful resource during live presentations. Participants expressed that the tool was easy to use and that the real-time captions/subtitles were accurate, useful, and easy to follow. One participant explained: “[It was easy to follow]; it would stop and pause and have little dots per second and then just keep flowing I really liked that mark.” According to other participants, it was beneficial to use the captions in their first language: “It is very beneficial, I think, being able to still read in your first language.”

Participants who had mixed feelings about the feature emphasized the importance of its accuracy and reliability, especially when switching languages. One participant remarked on the tool's performance, saying, “I was very impressed because she [the presenter] spoke quickly, and it did pick it up, you know, very accurately too.” They further appreciated the tool's ease of use: “I thought it was very easy to use and very user-friendly; the subtitles seem to follow her speech very clearly and thoroughly.” Another participant pointed out an unexpected benefit of using real-time captions to enhance focus: “I actually really enjoyed the presentation because my focus was on reading what was actually happening, instead of being distracted by the background or like whatever was going on, movements, like excessive movements.” These reflections underscore the tool's potential to not only accurately transcribe speech but also to improve participants' engagement with the presentation content.

The real-time captions were particularly valued for aiding in language comprehension, with participants noting that hearing and reading simultaneously helped them better understand and remember the content. One participant felt that it is useful to read what is being said: “I think sometimes, at least for myself, hearing it and reading it, it tends to stick a little better; so, for me specifically, that was a benefit.” Another participant explained that it can be helpful to understand their second language while listening to the presenter speak as well as reading the subtitles: “I listen to movies in English because English is my second language anyway, so I understand English more when I watch English movies [with subtitles], this is the same.”

Participants thought that the translations were accurate and easy to understand. One participant expressed, “I did find Romanian subtitles and I saw that they translated very accurately in real time, so I would say, I had a very good experience with this.” Overall, participants valued the captioning/subtitling tool for its ability to provide accurate, real-time transcriptions in an easy-to-use format, enhancing accessibility and engagement.

**Interaction With Subtitles**

This overarching theme emerged from participants’ testing or playing with PPT Live’s caption/subtitle feature during the live presentation.

**Participants’ Choice of Subtitle Language was Based on Their Language Fluency and Curiosity.** This theme emerged from what participants reported as reasons for having chosen the subtitle languages. The reasons were mainly based on their language proficiency or because they were curious to explore other languages. Some participants selected the language they could read, and others indicated that they were curious to read the subtitles in other languages.
Many participants enjoyed the ability to toggle between different languages. English speakers and those with other first languages, like Romanian, found it fun and useful to switch languages during the presentation. This feature was also seen as fascinating, even for those who did not speak a second language, as it allowed them to see translations in various languages. Participants who spoke only English indicated that the captions were easy to follow: “I followed the presentation in English, so it was easy to follow; it was great to follow in the same language.” Many participants reported that they used the toggle feature to determine which language they preferred. Some commented that they found useful and fun being able to switch between multiple languages:

- “It was my first time with this kind of presentation as well, similar to the other participants, I also toggle back and forth, so English is my second language, my first is Romanian.”
- Regarding the ability to change between languages, one participant commented, “I changed it back to a different language, and it was fun.”
- Another participant commented on the real-time captions’ usefulness, despite not speaking another language: “I don't speak another language but a fascinating to like flip through the languages and see all the different ways, you can read the same sentence.”

The participants who could read multiple languages indicated that the translations were accurate and checked for accuracy by reading the subtitles and toggling the languages back and forth: “I wanted to see if it was the same in another language, I changed it to Romanian, it was accurate.” Some participants identified that it was easy for them to follow the presentation with the subtitles, while others felt that the subtitles were too fast. One participant that found the presentation easy to follow stated, “I really liked the subtitles and the speed of it because it was quick, and I’m a fast reader and fast processor, so I really enjoyed that. I watch TV with subtitles all the time, so, for me, I enjoyed that.” Comparatively, one participant experienced the subtitles as too fast: “The subtitles were going way too fast and at some point, due to the context of the paragraph, for the sentence every word everything, so I had to let it go… but only if you have two languages.”

Overall, the subtitle language selection feature was embraced by participants for its ability to cater to diverse language proficiencies and curiosities, offering a personalized approach to engaging with live presentations. The experiences shared by participants underscore the value of technologies that accommodate individual learner needs.

Reading the Subtitles While Following the Presentation can be Distracting. This theme emerged from participants’ discussion on how toggling between subtitles languages distracted them from the presentation content. Many participants indicated being distracted when reading subtitles in a language different from the presenter’s language. One participant experienced this as impacting the amount of information learned from the presentation: “I want to be honest that while trying to juggle both, I didn't really get most of the information for the presentation.” Others agreed, “It can be hard to follow the presentation when you switch the language,” “Yeah, it can be distracting to do both because you can lose information,” “When you do both languages, yeah, it can be a distraction,” and “I agree, I did the same thing.” Overall, participants expressed concerns about the potential for distraction when toggling between subtitle languages, impacting their ability to absorb information from the presentation. Their experiences suggest challenges in managing attention between reading subtitles and following spoken content.
Discussion

The group dynamics influenced and shaped participants’ individual experiences with the presentation content and their use of technology in real time. During the FG sessions, we observed that participants were engaged, eager to share with others, and excited to implement what they had learned in the research session. Participants’ experience could be explained by two further themes that emerged when they were asked about their overall experience when using the PPT Live’s feature of automated real-time captions/subtitles in live online presentations: An “exciting” experience and the usefulness of the tool. Some quotes that support these themes are, “I enjoyed it as well, and I thought that it was kind of interesting to be able to see everything that she [the presenter] said it actually recorded;” “I’m very impressed with everything, thank you so much;” “I would say, had a very good experience with this.”

A surprising theme that emerged that supports the participants’ positive experience was related to the content of the presentation: the content was easy to follow, the actual topic was fascinating, and I learned a lot. Although we selected a general topic that we considered would be of interest to the participants, we did not expect them to highlight such positive aspects about the presentation’s topic or content. Perhaps participants’ engagement in the experience was what Schraw and Lehman (2001, as cited in Jones 2020) described a situational interest, or “interest and enjoyment students experience at any one moment as they engage in an activity” (p. 22). Situational interest is the student’s “emotional engagement with classes or tasks,” and it leads to cognitive engagement (Kahu et al., 2017, p. 60). When students are “more situationally interested in an activity [they] tend to be more motivated and engaged in that activity” (Hidi & Renninger, 2006, as cited in Jones, 2020). Additionally, introducing novelty and piquing students’ curiosity about the content are strategies that can increase situational interest (Jones, 2020). Perhaps the novelty of the tools and the participants' curiosity heightened their situational interest in the activity, which in turn helped maintain their engagement and excitement throughout the research sessions.

Challenges

While discussing challenges, participants did not express frustration but instead gave a positive meaning to their experience. The meaning became apparent as participants explained how they overcame the technical and technological challenges while being critical and suggesting improvements to the tool’s features.

Participants were able to troubleshoot connectivity issues quickly and overcome technological limitations they encountered while using the PPT Live’s caption/subtitle feature. Overall, troubleshooting effectively and finding solutions on their own could have instilled a feeling of competence in using digital tools, and troubleshooting by following peers’ advice highlighted a sense of collective problem-solving. It was not surprising that participants overcame challenges given their self-reported confidence levels in using basic technologies (see Table 1).

It is worth noting that some troubleshooting techniques that participants discussed (e.g., refreshing the browser or the phone screen) were not related to any of the basic technological skills we had asked them to self-report. We did not anticipate all the issues that participants could have encountered because of the several technologies that interact for PPT Live to function properly via Zoom (e.g., presenter’s and participant’s connection to internet, browsers, mobiles, computers, Zoom). Participants’ end-user-experience with PPT Live, as a cloud-based application, could have been impacted by issues occurring in any of the involved technologies, as explained by Cignoli (2017): “While the cloud provider can ensure adequate resources for
Some participants argued that PPT Live’s inability to recognize different dialects could be a challenge. Perhaps the theme of inaccurate translation due to dialectal variations would not have emerged as a challenge had participants been from the same geographical region, or the discussion would not have ensued if one participant had not indicated that dialectical variations “can be a problem.” It was apparent that participants’ experience heightens their awareness of the linguistic diversity within their own language and the challenges it poses for technology. Additionally, participants might interpret their experience as indicative of a broader issue of accessibility and inclusion in digital tools. This realization could lead to acknowledging the complexities involved in creating technology that accommodates such diversity.

However, it was surprising that participants would expect the tool to recognize all dialect variations for all languages. Some SRT-based translation tools, such as PPT, include variations of some languages as options for the spoken language to be captioned/subtitled (e.g., Spanish from Mexico, International, or from Spain; English from the United States or England). As speakers of a second language and users of translation tools ourselves, we could see that it might be a technological challenge for these tools to recognize all existing dialectical variations of all spoken languages.

Participants’ expressing difficulty to focus on the presentation content while hearing the presenter and reading the subtitles in a different language that they were not fully fluent in appears to align with the concept of the split-attention effect in the Cognitive Load Theory (CLT) Sweller, 2010). According to Sweller, the split-attention effect occurs when the learner must split their attention between multiple sources of visual information that are essential for learning and, thus, causing extraneous cognitive load. Sweller defined extraneous cognitive load as the load caused by “inappropriate instructional designs that ignore working memory limits and fail to focus working memory resources on schema construction and automation” (p. 26). From the Cognitive Theory of Multimedia Learning (CTML) standpoint, extraneous cognitive processing— analogous to extraneous cognitive load—refers to cognitive processing that does not serve the instructional goal (Mayer, 2020). Mayer exemplifies a scenario of extraneous processing when a learner would need to visually scan back and forth between captioned words at the bottom of a screen and an animation above the screen.

It cannot be inferred that the presentation design had an impact on extraneous cognitive processing and, thus, learning. However, a split-attention effect may have occurred in our multimedia presentation scenario where participants were exposed to multiple visual sources of information (i.e., images and text in the PPT slides, the presenters’ talking head, and the transcript in their mobile or another screen). Additionally, although the participants did not report major challenges with the use of the tools, using various software and hardware tools at the same time (i.e., PPT Live, Zoom, computer, mobile) may have added to participants’ split attention. For example, in addition to the effort of reading the subtitles, because of the PPT Live’s subtitles being presented on a screen or device different from Zoom’s some participants may have been viewing duplicate presentation slides while viewing the presenter on the Zoom screen at the same time they tried to read the transcript on another device/screen (see Figures 1 and 2).

Overall, participants’ experiences shed light on the cognitive challenges of multitasking in learning environments, especially in multilingual settings. Participants’ struggles with processing information in different languages simultaneously raise important considerations for the design of educational content and the use of subtitles in teaching. Participants’ experiences point to broader educational implications, suggesting a need for pedagogical
strategies that consider the cognitive demands placed on learners and seek to minimize potential distractions, thereby enhancing comprehension and retention.

**Benefits**

The meaning of participants’ experiences was beyond satisfaction with a functional tool. It encompassed an appreciation for how technology can enhance learning, promote inclusivity, and facilitate clear and effective communication across diverse groups. This meaning became apparent through participants’ enthusiasm and engagement in discussions about the tool’s benefits and their evident appreciation for its impact across different learner types and environments.

Participants appeared excited when describing how remote audiences could have live access to the “expert” or “our thoughts as educators” and how they could apply live captions/subtitles as students, teachers, or practitioners to reach out to a broader and more diverse audience. Perhaps their excitement was because experiencing PPT Live real-time captions/subtitles via Zoom was something new for most. Although all participants had been exposed to Zoom, and most had taken courses offered via Zoom during the pandemic (n = 17), most (n = 16) had not been exposed to PPT Live real-time captions/subtitles.

Perhaps participants thought the subtitle features could be helpful to others in reinforcing the learning of a language because of their own positive experiences with subtitles. For example, some had been exposed to subtitles in movies and felt that reading subtitles while hearing a different language helped them reinforce/learn/understand better what was being said: “I listen to movies in English because English is my second language anyway, so I understand English more when I watch English movies [with subtitles], this is the same.” Subtitles in movies have been found beneficial not only for the deaf but also for individuals who are learning to read and those in the process of learning a new language (Gernsbacher, 2015).

In addition to the participants’ perceived usefulness of captions/subtitles, we considered the participants’ perceived caption/subtitle accuracy as a benefit. Because of the current limitations of SRT and previous findings (see Orellana et al., 2021), we did not expect fully accurate captions with PPT Live and took measures to maximize the captioning effectiveness. During the presentation sessions, we noted inaccuracies as we read the PPT Live transcripts in English and in other languages that we could read. Perhaps the participants did not stress the inaccuracies because they could hear and understand spoken English to help them overcome the inaccuracies in the transcript as they heard the correct words in English and other languages they knew.

**Interaction With Subtitles**

Participants attributed significance to their experiences, valuing both the accessibility provided by subtitles and the opportunity for linguistic exploration they afforded. Their engagement with selecting subtitle languages during live presentations highlighted the dynamic relationship among technology, learning, and individual choice in educational contexts. This significance was particularly evident as participants reported their enjoyment and appreciation for the subtitle feature.

The ability to switch between subtitle languages sparked considerable interest among the participants. Many chose their subtitle language based on their own linguistic proficiency or driven by curiosity. However, it was unexpected that participants would frequently toggle between subtitle languages out of sheer interest or curiosity during the live presentation. This behavior of switching languages for exploration might have diverted their attention from the
presentation's content, potentially leading to increased cognitive load, a point further elaborated in the "Challenges" section.

**Limitations**

Limitations of the study were the following:

1. We were interested in the experience of participants and not in their learning with captions/subtitles or the effectiveness of the presentation design. Participants were not asked to retain content or learn from the presentation.
2. Participants were proficient in the presenter’s language.
3. We did not plan nor design an accessible presentation session for participants with hearing or visual disabilities.
4. The instructional context was limited to a one-way lecture-type instructional scenario where the audience was discouraged from speaking or interacting with the presenter because of the limitations of PPT Live via Zoom.
5. We asked the presenter not to help the participants troubleshoot as perhaps an instructor would in a class. We purposefully disallowed this type of situation that would have resulted in interactive or non-instructional unrehearsed dialogues.

**Conclusions and Recommendations**

We explored how students without hearing disabilities and fluent in the presenter’s language experience using automated captions and subtitles during one-way live online presentations. By using FG as a data collection method and Colaizzi’s data analysis approach, we interpreted and gave meaning to the participants’ collective experience. While participants were able to troubleshoot connectivity and technological issues encountered, they experienced an apparent split-attention effect and noted limitations in the tool's inability to recognize different dialects. However, overall, participants' experiences were positive, suggesting their ability to look beyond the immediate technical limitations of the tool and focus on the broader benefits and implications of live captioning and subtitle technologies. Their experiences also suggest their appreciation for educational practices that embrace inclusivity, minimizing barriers related to language and geographical location. This approach aligns closely with the principles of UDL, highlighting the importance of accessible and inclusive educational environments.

To exemplify advancements in ASR in recent years, we presented a non-exhaustive list of tools that let users access real-time automated captions, transcripts, and multilingual subtitles. The tools are various, and their use could become pervasive in the educational arena as a cost-effective approach to allowing equal access to content in online or on-site meeting scenarios. Although we did not describe the tool’s capabilities (e.g., the number of languages it could caption or translate, the ability to save transcripts, the cost, or limitations of the free versions), it can be expected that these tools would improve their automated caption/subtitle features, other tools would emerge, or the existing ones would incorporate these features. For example, by the time this article was being written, Zoom had introduced the feature of automated subtitles in various languages (Ball, 2022; Zoom Video Communications, Inc., 2022b).

It is worth noting that, as of the time we were conducting this study, automatic captions with PPT Live were not confirmed to be fully accurate, and, thus, they did not meet accessibility requirements. If offering human-generated captioning services is not possible
during online live sessions to meet accessibility accommodation requirements, a recording of the live session with captions/subtitles edited for accuracy could be provided whenever allowed that, in turn, would benefit all learners.

Our findings support previous conclusions that PPT Live’s automated captions/subtitles could be used as text-based alternative to auditory information as suggested in the UDL guidelines, mainly when the audience can hear the presenter and understand the presenter’s language (see Orellana et al., 2021, 2022). However, when aiming to address learner diversity to offer multiple options of perception and of language, educators should understand the limitations and affordances of automated captions/subtitles. Research is warranted to better understand how to effectively integrate automated captions/subtitles for learning in live online, hybrid, and onsite classrooms. Following, we present non-exhaustive recommendations for research.

Automated Captions and Subtitles

In our remote study scenario, the presenter was a native English speaker, and student-participants were proficient in English with some able to read one or more additional languages. A question that arises is how effective would live online learning interactions with automated captions/subtitles be in scenarios where the accuracy of captions/subtitles may be impacted (e.g., the presenter is delivering instruction in their non-native language, has an accent, is speaking in a loud place, does not use proper technology, speaks special terms) or when the learners are not fluent in the presenter’s language or cannot hear, making it difficult for them understand the content by only reading the inaccurate captions/subtitles. Furthermore, with PPT Live, real-time automated captions/subtitles can be accessed by on-site on online viewers (Microsoft Education, 2020) and further research could focus on determining effective on-site or hybrid instructional scenarios where students can follow the presentation in their preferred language.

According to Hacker et al. (2020), the use of live videoconferencing technologies had nearly doubled and continued to rise as jobs and schools shift to a permanent work-from-home approach because of the pandemic. Further research is needed to understand possibilities and limitations to interact in these types of virtual settings in academic and non-academic scenarios for individuals with different learning styles, with impairments or disabilities, or who encounter language and communication barriers. Research can also help instructors effectively create online inclusive scenarios using interactive live videoconferencing tools with students in unstructured non-rehearsed meetings, such as informal encounters or office hours, and in interactive multilingual learning scenarios where the participants select their preferred language to speak and read.

Live Multimedia Presentations

According to Mayer (2018), there is a need to continue the research growth on replicating or creating boundaries of existing principles of instructional design or creating new principles for new venues. Further research could focus on testing principles from cognitive psychology theories (see Mayer, 2020; Sweller, 2010) to design effective multimedia presentations for learning in live online scenarios. For example, Mayer et al. (2020) present five evidence-based principles to increase the effectiveness of instructional videos and one when instructional videos are not effective. The subtitle principle is among the five principles to increase effectiveness: “People learn better from a video documentary in their second language when the words are printed (or printed and spoken) rather than spoken” (p. 847). Mayer et al. refer to evidence suggesting that “when students are viewing an instructional video
in their second language, it would be useful to add subtitles and make sure the pace is slow enough not to overload working memory” (p. 848). On the other hand, the seductive principle explains that “people do not necessarily learn better when interesting but extraneous video is added to a multimedia lesson [such as the presenter’s talking head]” (p. 849). Mayer et al. recommended further research to determine the conditions under which the principles apply. Research could focus on exploring how these principles apply in live instructional presentations via videoconferencing, as opposed to recorded instructional videos.

Furthermore, according to Skulmonski and Xu (2022), CLT has been challenged by interactive learning media where learners need to invest on extraneous load to allow the occurrence of certain types of germane processing aimed at making sense of the material. Hence, guidelines are needed to determine how much extraneous load is acceptable for learning. Future research could investigate acceptable amount of extraneous load when using SRT’s for automated captions/subtitles for learning, based on Hollender et al.’s (2010) adapted CLT model where the concept of extraneous load is reformulated to add the load due to digital interactions caused by the usability of the tool.

Additionally, learners may experience distractions or possible extraneous load that may affect learning in live instructional presentations via videoconferencing. For example, viewers may lose focus on the instructional presentation when reading inaccurate/unintelligible automated captions/subtitles, keeping up with lagging captions/subtitles, viewing talking-head videos of the presenter and of other participants, or frequently reading out-of-sync transcripts of previously spoken words while hearing live speech. Further experimentation could contribute to determining the effect on learning of such distractions during live instructional presentations via videoconferencing.

In addition to effective multimedia presentations for learning in live online scenarios, an inclusive presentation session would require adhering to effective practices for live online or on-site accessible presentations (e.g., see Henry, 2022b; National Center on Accessible Educational Materials, n.d.; Vogler, 2020). Further research could focus on how the practices for accessible presentations and multimedia learning principles could be implemented for learning in a live classroom with students with and without disabilities.

References


CAST. (2018b). *Universal Design for Learning Guidelines version 2.2 guideline 1, provide options for perception.* https://udlguidelines.cast.org/representation/perception


Hall, J. N. (2020). Focus groups: Culturally responsive approaches for qualitative inquiry and program evaluation. Myers Education Press, LLC.


Henry, S. L. (Ed.). (2022b, August 31). Making events accessible: Checklist for meetings, conferences, training, and presentations that are remote/virtual, in-person, or hybrid. https://www.w3.org/WAI/teach-advocate/accessible-presentations/#material


Jones, B. D. (2020). Motivating and engaging students using educational technologies. In M. J. Bishop, E. Boling, J. Elen, & V. Svihla (Eds.), Handbook of Research in Educational
Communications and Technology (pp. 9-35). Springer. https://doi.org/10.1007/978-3-030-36119-8_2


Microsoft Education. (2020, January 6). Engage your audience with Live Presentations in PowerPoint [Video]. YouTube. https://www.youtube.com/watch?v=Lzfjwn05LZg


Author Note

Anymir Orellana is a Professor at Nova Southeastern University’s Abraham S. Fischler College of Education and School of Criminal Justice. She teaches online graduate courses in English and Spanish in related areas of instructional technology, instructional media, instructional design, and distance education. Please direct correspondence to orellana@nova.edu

Elda Kanzki-Veloso is a Professor at Nova Southeastern University’s College of Psychology, counseling department. Her main responsibilities include teaching and coordinating the practicum and internship placement process for the department.

Georgina Arguello is an Assistant Dean for Academic and Faculty Affairs for Bilingual Programs and an Associate Professor at Nova Southeastern University’s Abraham S. Fischler College of Education and School of Criminal Justice. She teaches online and blended doctor-level courses for domestic and bilingual students.

Katarzyna Wojnas completed her master’s and is completing her doctorate in clinical psychology at Nova Southeastern University's College of Psychology.

Acknowledgement: This study was funded by Nova Southeastern University’s 2021-2022 President’s Faculty Research and Development Grant, scholar track, awarded to the project titled “Students’ Experiences Using Live Captions and Subtitles in Class Presentations.”

Copyright 2024: Anymir Orellana, Elda Kanzki-Veloso, Georgina Arguello, Katarzyna Wojnas, and Nova Southeastern University.
Article Citation