AN E-MATE CASE STUDY: ADDING CLOSED CAPTIONS TO *FUNDAMENTALS OF LIGHT AND LASERS, COURSE 1*

BACKGROUND

Consider this. Universal design advances in construction like sidewalk ramps were implemented to accommodate those in wheelchairs, yet have proven beneficial to people pushing baby strollers, pedaling bicycles, riding skateboards, and rolling hand-trucks.

Universal design for learning is likewise about accessibility, not disability. Providing closed captions in videos, whether the videos are in an interactive e-textbook, online course, website, or shown in a traditional face-to-face classroom has the potential to benefit all learners, not just the deaf or hearing impaired.

To increase the learning potential of the interactive e-textbooks developed under E-MATE, E-books and Mobile Apps for Technician Education (NSF ATE DUE# 1205113) the project team explored adding closed captions to videos in *Fundamentals of Light and Lasers, Course 1*. The e-textbook was developed in collaboration with OP-TEC, the National Center for Optics and Photonics Education, and includes 63 videos totaling 310 minutes or approximately 5 hours. The videos range in duration from 39 seconds to 21 minutes. Twenty-three of the videos demonstrate and discuss hands-on photonics and optics lab activities, and 40 of the videos provide on-demand math tutorials. All videos include technical terminology and most also include mathematical equations. Creating transcripts and subsequently encoding closed captions for all of these videos was a seemingly overwhelming undertaking.

The team experimented with a free trial of popular commercial voice recognition software over a year ago, and the results were entertaining at best and would have been flagged by Google as "gibberish". A quick Google search confirmed what we already experienced, DIY options to automatically generate transcripts are about 70% accurate at best. Creating our own transcripts, given limited staff and time and the technical nature of the content, was not a viable option either.

COMMERCIAL CLOSED CAPTIONING SERVICES

Two vendors that provide transcription services in education kept surfacing during further online research: **3Play Media** in Massachusetts and **cielo24** in California.

The team participated in a free **3Play Media** webinar that discussed federal and state closed caption accessibility guidelines for education and also touched upon the company's transcription generation services. A follow-up call was scheduled, and a free trial demo was set up.

The team also scheduled a call with **cielo24** to learn more about the company and its services. Another free trial demo was arranged.

To compare transcription accuracy, the team uploaded the same video from *Fundamentals of Light and Lasers*, a Module 1-5 lab video 9 min., 37 sec. (9:37) in duration and within the free trial time limits for both vendors. The demos also gave the team the opportunity to compare the two interfaces for uploading videos, editing transcripts, and downloading captions and transcripts. Both vendors allow users to upload glossaries of terms with videos, an important feature for technical content, and both offer multiple transcript and caption download file formats.

3PLAY MEDIA

3Play Media has a 99% or higher accuracy standard and follows a three-step transcript generation process that blends technology with human talent:

1. Computer recognition. When a video is uploaded for transcription a computer uses automatic speech recognition software to generate an initial transcript.
2. Editing/cleanup. A US-based professional transcriptionist reviews the computer generated transcript and edits as necessary.
3. Quality control. The edited transcript is put through a two-phase Quality Assurance (QA) process:
    1. The transcript is rendered by an error-flagging algorithm
    2. A human quality assurance manager reviews the transcript to verify accuracy of difficult words and to correct punctuation.
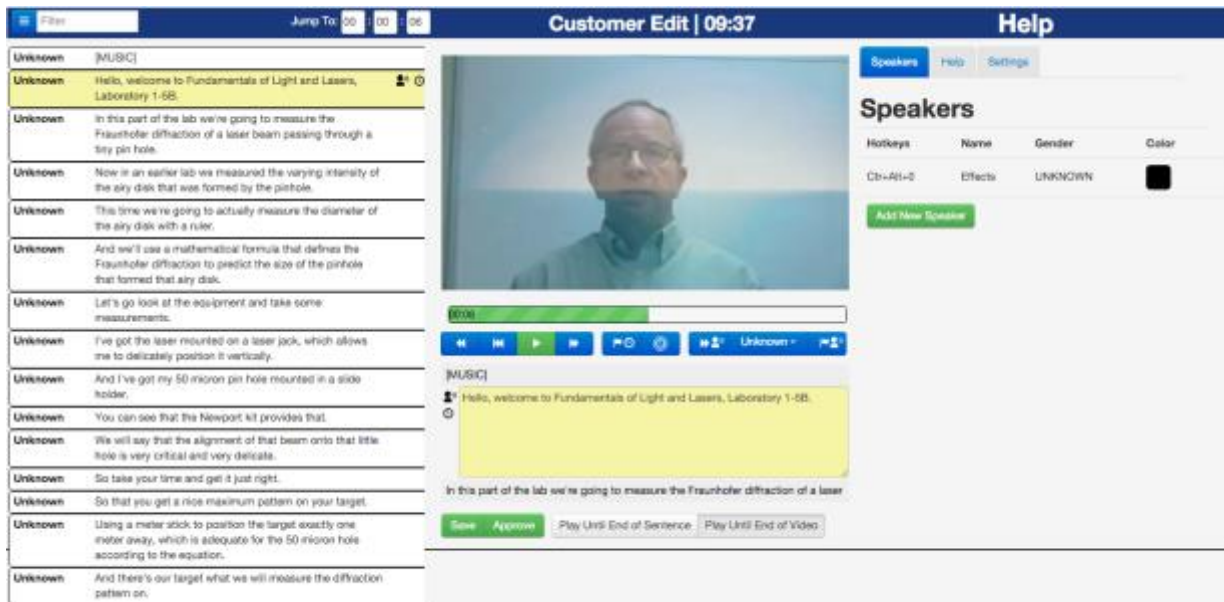
3PLAY MEDIA TRANSCRIPT EDITING INTERFACE
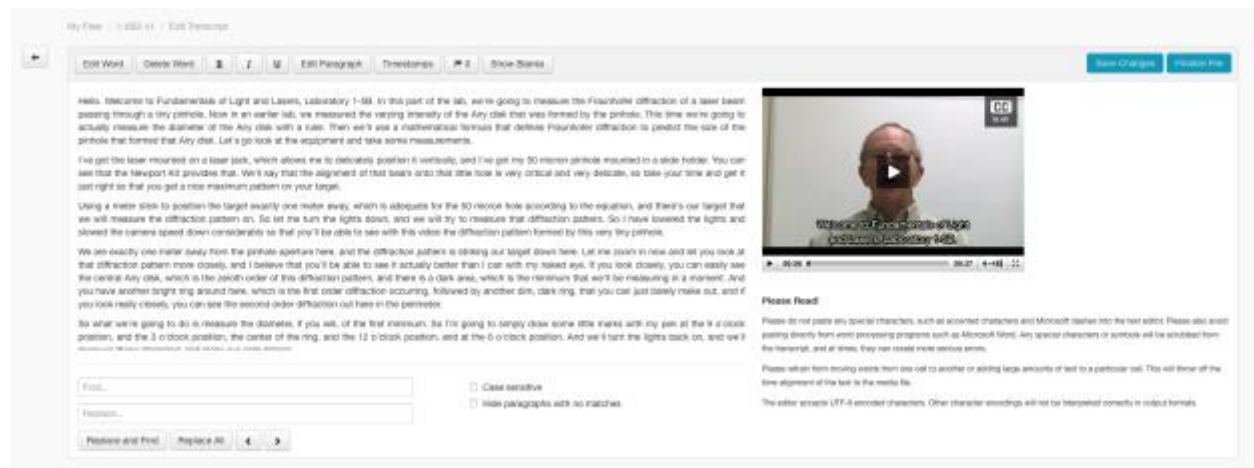
"Domain expertise" is also offered. For example, upon upload, users can indicate that video content has a lot of mathematical equations. Speaker identification settings are set before uploading.

CIELO24

The **cielo24** transcript generation approach is a layered, multi-pass workflow designed to achieve scale, while maintaining quality and keeping costs low. When a video is uploaded, **cielo24's** proprietary software chunks content into 20-second snippets that are distributed to a crowd of curated human transcriptionists. The snippets and their associated transcripts are then digitally stitched back together for a full end-to-end review and edit by the most qualified worker (based on a **cielo24** quality ranking system). The edited video is then submitted for a final quality control review to ensure the transcript meets the vendor's 99%+ accuracy standard (Professional level). **cielo24** has security measures in place throughout the process, such as auto-expiring URLs to prevent unauthorized access.

CIELO24 TRANSCRIPT EDITING INTERFACE

THE RESULTS

Both vendors produced the high accuracy rates advertised, and both provided the transcripts in the time promised. Both vendor interfaces provided positive user experiences for video uploading and transcript editing, and contacts from both vendors were very responsive to questions before and during the demo trials. Costs were also comparable.

The most significant difference in service offerings is that **cielo24** does not encode captions. For readers to be able to toggle closed captions on/off in the iBooks e-textbook, the team learned videos must be in .m4v format with the closed captions encoded. As neophytes to closed captioning, however, the team did not know how difficult and time-consuming DIY caption encoding is or isn't.

The team therefore opted to purchase **3Play Media** services first because caption encoding services could be purchased should DIY prove too challenging. The encoding service is $10 per video, regardless of video length. This approach also allowed the team to move forward with transcription generation for the 14 Module 1-1 videos while still researching and testing options to do the encoding ourselves.

DIY ENCODING

Compressor, an app available through the Apple App Store surfaced as an affordable $50 and positively reviewed option. The team purchased the app.

The biggest unknown was synchronization or alignment of the closed captions with the video. Is using Compressor a simple matter of selecting the video file, caption file, and output format and clicking a button? Or is there another step: manually aligning the timing of closed captions with the video. And, if so, how time and labor intensive is it to get the timing right?

There are multiple closed caption file formats. Online research disclosed SCC (.scc) or Scenarist Closed Captions is the format required for Apple Devices. During the week the Module 1-1 videos were being transcribed, the team experimented with Compressor using the video uploaded for the free trial demo and its resultant transcript. The captions were downloaded in .scc format from the **3Play Media** interface for testing with Compressor.

After spending 5-10 minutes exploring the Compressor interface overviews provided under Help in the application, the team made its first encoding attempt, which seemed to require a few simple steps:

1. Select the destination for the final encoded video
2. Click **Add File** and upload video
3. Select an output format, in this case **Apple Devices**
4. In the **Inspector** at right:
    a. Under **File Properties**:
        i. Select **English** as language in dropdown
        ii. Select United States as country in dropdown
    b. Under **Metadata**, click **Choose** button to upload the .scc file
5. Click **Start Batch** button

It took just 2:48 to process the encoding for the 9:37 video and its associated .scc file. Better yet, the synchronization was spot on. The team, however, did not know if this was beginner's luck, or if encoding and synchronization would prove this simple and seamless for all videos.

TESTING IN IBOOKS

The processed encoded .m4v file was then dragged and dropped into an iBooks Author Media widget and then previewed and tested in iBooks. The captions were indeed there and, by tapping the speech bubble icon in the video playback controls, the reader could toggle them on/off.

IBOOKS AUTHOR FILE TO TEST ENCODED .M4V VIDEO IN MEDIA WIDGET
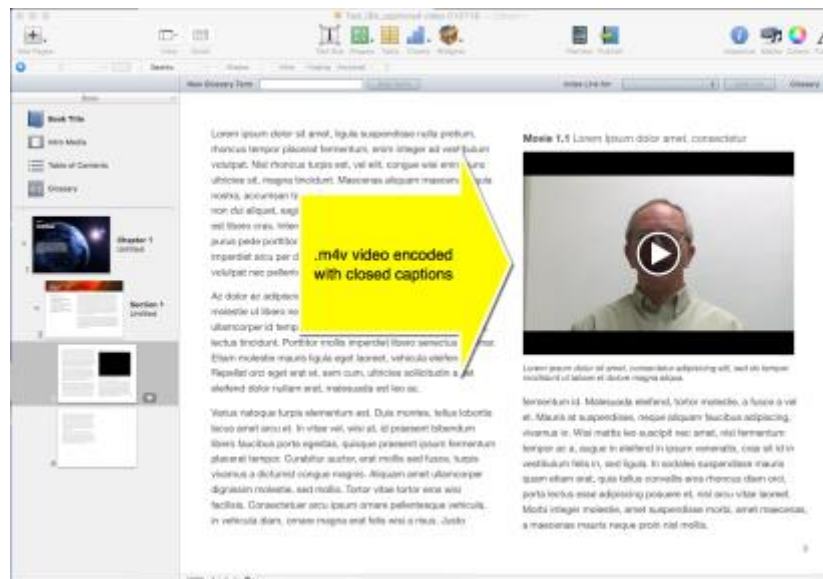
UPDATING THE E-BOOK WITH CLOSED CAPTIONS

The team used Compressor to encode all 14 Module 1-1 videos and their associated .scc caption files. The shortest video in Module 1-1, 39 seconds, took Compressor 14 seconds to encode with the .scc caption file. For videos in the 1-2 min. duration range, processing took between 20 and 30 seconds. The longest video, 14:55, took 3:29 to encode. Alignment or synchronization of captions and videos was spot on for all videos, which validated that the same result encoding the first demo video was not beginner's luck. Encoding .scc caption files generated by **3Play Media** with videos using Compressor was indeed fast, easy, and aligned.

Closed caption encoding using Compressor did increase all file sizes for the Module 1-1 videos processed, so the team explored other encoding and compression options with Pellet Productions. Pellet was able to decrease file sizes using other tools; however, we learned that the Pellet .m4v videos had open captions only when viewed in iBooks. Open captions are burned into the video and are therefore always on. Furthermore, these open captions were fixed at the top of each video and covered essential content. The team felt it was important for users to be able to toggle closed captions on and off and that the captions be displayed at the bottom of the screen to block minimal content. Pellet did further research and learned that Compressor is the only encoding tool that allows you to generate closed caption videos for iBooks that allow users to toggle captions on and off. The team therefore encoded all videos using Compressor. All original videos were replaced with Compressor encoded closed caption videos. Surprisingly, overall e-textbook size decreased slightly. This was great news because it meant the e-textbook with closed captioned videos did not exceed the iBooks Store size limit.

Since the team also plans to provide closed captions for the video tutorials on the project website, which are hosted on a YouTube channel to which the website links, the team also experimented with this process. The video used for the free trial demo was uploaded to the project YouTube Channel. The captions were downloaded from **3Play Media** in SRT (.srt) format. SRT stands for SubRip Subtitle and is the format YouTube requires. **3Play Media** allows users to download captions multiple times and in multiple formats for no additional cost. Following instructions easily found on the **3Play Media** website, the video and captions were added to the E-MATE YouTube channel. One quick attempt was all it took.

LESSONS LEARNED

The lessons learned throughout this research and experimentation are as follows:

1. The team is still scratching the surface of best practices and possibilities, but the speed and ease with which we got this far is very encouraging.
2. Don't be afraid to dive in and start adding closed captions to your videos because you have no knowledge, understanding, or experience. Google and go!

WHAT THE TEAM LIKES ABOUT 3PLAY MEDIA SO FAR:

- The interface is very intuitive.
- 99% accuracy is great, but it is nice to have the opportunity to catch and correct that 1%.
- Users can download captions and transcripts over time and in multiple formats, not just once.
- Users can be added to the online account, and their access can be restricted. This will make it possible to add student workers and subject matter experts to review and edit transcripts only.
- E-mail notifications are sent when a transcript is ready for review/editing.
- Turnaround options are available by file, not project, upon video upload.
- Pricing is based on turnaround timeframe needs.
- Real time billing updates are available by clicking Invoices.
- "Math" can be selected during the upload process for content that has a lot of equations.
- Online documentation, video tutorials, and resources cover a depth and breadth of closed captioning related topics and are of high quality.
- Glossaries of technical terms can be uploaded.
- Translation services are available at an additional cost.
- Users can import and edit their own existing transcripts.