Enhancing Accessibility of Engineering Lectures for Deaf & Hard of Hearing (DHH): Real-time Tracking Text Displays (RTTD) in Classrooms

Mr. Gary W Behm, Rochester Institute of Technology (CAST)

Gary W. Behm, Assistant Professor of Engineering Studies Department, and Director of NTID Center on Access Technology Innovation Laboratory, National Technical Institute for the Deaf, Rochester Institute of Technology.

Gary has been teaching and directing the Center on Access Technology Innovation Laboratory at NTID for five years. He is a deaf engineer who retired from IBM after serving for 30 years. He is a development engineering and manufacturing content expert. He develops and teaches all related engineering courses. His responsibility as a director of Center on Access Technology Innovation Laboratory include the planning, implementation and dissemination of research projects that are related to the need of accessibility. He received his BS from RIT and his MS from Lehigh University. His last assignment with IBM was an Advanced Process Control project manager. He managed team members in delivering the next generation Advanced Process Control solution which replaced the legacy APC system in the 300 mm semiconductor fabricator. Behm has fifteen patents and has presented over 30 scientific and technical papers at various professional conferences worldwide.

Dr. Raja S Kushalnagar, Rochester Institute of Technology

Raja Kushalnagar is an Assistant Professor in the Information and Computing Studies Department at the National Technical Institute for the Deaf at the Rochester Institute of Technology in Rochester, NY. He teaches information and/or computing courses, and tutors deaf and hard of hearing students in computer science/information technology courses. His research interests focus on the intersection of disability law, accessible and educational technology, and human-computer interaction. He is focused on enhancing educational access for deaf and hard of hearing students in mainstreamed classrooms. He worked in industry for over five years before returning to academia and disability law policy. Towards that end, he completed a J.D. and LL.M. in disability law, and an M.S. and Ph.D. in Computer Science.

Mr. Joseph S. Stanislow, National Technical Institute for the Deaf

Joseph Stanislow is a faculty in the Information and Computing Studies Department at the Rochester Institute of Technology/National Technical Institute for the Deaf. Joe has been teaching and tutoring computer and information technology to AOS/AAS/AS/ BS students for over 14 years. He has a Bachelor’s degree in Electrical Engineering Technology from the Rochester Institute of Technology and a Master of Science in Computer Science from Stevens Institute of Technology. Before joining the NTID faculty he worked for AT&T Bell Laboratories as an electrical engineer, a physical designer, and a computer programmer for 20 years.

Dr. Aaron Weir Kelstone, RIT/NTID

background provides insight related to the use and implementation of access technology in performance and presentation environments to support development of prototypes for use in performances.
Enhancing Accessibility of Engineering Lectures for Deaf and Hard of Hearing Students: Real-time Tracking Text Display in Classrooms

Abstract
The introduction of Real-time Text Display (RTD), in which typists transcribe audio and display it to students in real-time has greatly increased accessibility of lectures for deaf and hard of hearing (DHH) students as evidenced by their increased graduation rates from post-secondary programs. However, significant but subtle barriers in current static displays of RTD persist, especially in engineering, which makes heavy use of detailed visuals and explanations via sequential steps. Hearing students are able to look at the visuals and simultaneously listen to the spoken explanation and combine the two effortlessly. By contrast, DHH students have to constantly look away from the static image of RTD on a display to search and observe details in the lecture visually. As a result, they spend less time watching lecture visuals and gain less information than their hearing peers.

We discuss the implementation and evaluation of an accessible technology system, Real-time Tracking Text Display (RTTD) that addresses the issue of accessibility to information that DHH engineering students face every day in the classroom. The system minimizes the distance between RTD and presenter by tracking the presenter’s movement and then displaying the RTD at a fixed distance right above the presenter. An evaluation showed a significant preference in favor of RTTD over static RTD in terms of reducing the distance between the presenter and images. Future enhancements should focus on improving the flow and/or the predictability of the moving RTTD.

Challenges
Historically, DHH students are an under-represented and under-served minority in higher education because they do not receive adequate information in class. As a result these students are often unprepared for traditional STEM classrooms. Most DHH students cannot understand spoken lectures without the aid of aural-to-visual access. Prior to the introduction of the earliest accessibility laws in 1974, less than 4% of DHH individuals completed college in the 1960s. Although substantial progress has been made in the past 40-50 years in terms of accessibility laws and technology, the disparity in graduation persists: the graduation rate is 16% of DHH as compared to 30% for their hearing peers.

Currently there are over 31,000 DHH students enrolled in college and this enrollment number is up 15,000 over the past 10 years. While careers demand highly qualified engineers with various skills, including visual skills, most classrooms are not designed to utilize students’ visual skills and are not fully accessible by DHH students, including engineering classrooms. When classrooms maximize the benefits of visual learning, the barriers of traditional pedagogy, such as using spoken English as the primary mode of communication are partially ameliorated. Specifically, DHH students also struggle in solving engineering problems. While a university may provide resources such as sign language interpreters, note takers, real-time text (RTD) services and specialized mentoring faculty, the instructional environment within classrooms may not provide optimized learning environment for DHH students.
Hearing students are able to look at the visuals and simultaneously listen to the spoken explanation effortlessly. By contrast, DHH students have to constantly look away from RTD to search and observe details in the lecture visual as shown in Figure 1. They risk losing information, which can slow down or even derail learning. As a result, they spend far less time watching lecture visuals and comprehend less information than their hearing peers. The ability to follow lectures with dispersed visuals is difficult without accessible technology.

We address two major barriers in using RTD. **First**, the audio-to-visual nature of RTD creates a simultaneous visual for the student, who is forced to choose and switch between reading the RTD and the current visual (e.g., slides or whiteboard). If the student looks at the RT, they will miss the whiteboard or vice-versa as shown in Figure 1. **Second**, as the presenter moves around, this creates distance between the visual information and RTD, which forces the student to move his/her eyes or head constantly. As a result, the student will focus only on the RTD or visuals to his/her detriment.

**Background**
Real-time text is usually delivered through a display system. One popular integrated captioning and display system is C-Print, which was developed at the National Technical Institute for the Deaf, a federally funded institution for deaf and hard of hearing students. The system has significantly improved access to lectures for DHH individuals in many programs around the country\(^9,10\). It also benefits individuals with other disabilities, such as those with a visual impairment or a learning disability\(^a\).

The basis of C-Print is printed text of spoken English displayed on a laptop or a mobile device in real time, which is a proven and appropriate means of acquiring information for some individuals who are DHH. A trained operator, called a C-Print captionist, produces a text display of the spoken information in classroom or other settings. At the same time, one or more students read the display to access the information.

\(^a\) (https://www.rit.edu/ntid/cprint/)
People often falsely assume that the use of captions in the classroom enables full access to real-time text for DHH people. This assumption minimizes other information accessibility issues, such as simultaneous visual streams or content complexity. For example, DHH students have to constantly switch gaze between the C-Print display and the lecture slides as shown in Figure 2. So, for current classrooms, captions should be presented in a way that reduces the effects of visual dispersion and cognitive overload\textsuperscript{11}.

**Proposed Solution: RTTD**

We implemented a system called Real-Time Tracking Text Display (RTTD) that addresses the challenges described in the previous section. The goal of the RTTD design is to minimize visual dispersion. The system involves three components: C-Print, Kinect 2, and the RTTD system. The system can be easily set up and turned on within a few minutes by the captionist.

**Kinect for Windows**

A Kinect for Windows device, developed by Microsoft, is a physical device with a depth sensing technology. It has a built-in color camera, an infrared (IR) emitter, and a microphone array, enabling it to sense the location and movement of people as well as their voices. The depth sensor consists of an infrared laser projector combined with a monochrome CMOS sensor, which captures video data in 3D under any ambient light conditions. The sensing range of the depth sensor is adjustable. The Kinect for Windows Software Development Kit (SDK) 2.0 enables developers to create applications that support gesture and voice recognition, using Kinect sensor technology on computers running Windows platforms\textsuperscript{b}.

\textsuperscript{b}(http://www.microsoft.com/en-us/kinectforwindows/)
Real-Time Tracking Text Display (RTTD)

The RTTD system is a software application developed by the authors that integrates the C-print application with a Kinect 2 device, so that the display of real-time text tracks the person’s position. The RTTD system is designed to be portable and it consists of a portable computer with Kinect 2 and a tripod that can be carried in an over-the-shoulder bag as shown in Figure 3. Many classrooms have a projector and the RTTD system can connect directly to the projector. If the classroom does not have a projector, the portable projector can be used. The best position for the Kinect 2 with tripod is placed in the front row of student desks so that it can track the movement of the instructor as shown in Figure 4a. The RTTD real-time transcript is produced by C-Print above the freely moving presenter, within the viewer's peripheral vision as shown in Figure 4b.
The RTTD uses the Kinect 2 system to localize of the presenter, as shown in Figure 5. The system then projects the captions at a fixed distance from the presenter’s location, usually over their head. This enables the viewer to easily see both the transcript and presenter’s information simultaneously. They can detect changes, easily follow and read the real-time transcript without missing the presenter's information. Blocked views occur when the transcript is displayed in a static location next to the presenter's slides or demonstration. The view of the transcript is easily blocked whenever the presenter moves in front of the real-time transcript.

**Evaluation of RTTD**
To address these issues our first testing of the system involved a mock lecture in a typical classroom, with a real teacher and college students. We adopted this approach due to time and schedule constraints for access to actual courses. We recruited a math teacher to present slope and y-intersection instruction for 30 minutes. The teacher selected instructional materials to ensure it was easy to understand and highly visual. The lecture was divided into two equal segments of 15 minutes each. For the first segment, the captions were displayed using the RTTD system, and for the second segment, the captions were displayed in a fixed location on the right of the whiteboard – **Real-Time Fixed Text Display (RFTD)**.

We compared two C-Print display settings: 1) use of RTTD with a projector to display the real-time text above the presenter as shown in Figure 6, and 2) use of a projector alone to display the captions to the side of the lecture as shown in Figure 7.
The C-Print output is projected through the RTTD system onto a whiteboard. All participants had a clear view of the instructor and the whiteboard. The projector was mounted at the rear of the classroom and connected to a desktop computer and a Kinect sensor to capture movement of the instructor across a width of fifteen (15) feet. A C-Print captionist typed the real time text that accurately represented the teacher’s spoken content. The text was then projected using two lines of text directly above the presenter.

We recruited 13 participants (10 male, 3 female) for the study, who were paid for their time in the study. Their mean age was 24.55 years. Their mean time in college was 3.32 years, and their mean experience with using captions in the classroom use was 2.45 years. Participants’ experience in using captions in the classroom was relatively consistent across individuals and they had sufficient experience to divide their gaze between the captions, teacher and math information on the whiteboard.

After viewing the lecture and the captions displayed via 1) RTTD and 2) RTFD, the students were instructed to complete a survey that consisted of three questions that used a five-point Likert scale. The questions were:

1) How easy was it for you to read the real-time text?
2) How helpful were the captions in understanding the lecture?
3) How likely are you to recommend this captioning system to other students?

We asked the students to respond to the questions twice. First after they viewed the captions through the RTTD system, and second, after they viewed the captions through the RTFD system. The students were also asked to write down their thoughts and feedback on the strengths and weaknesses of viewing the captions. Since the participants could only watch the lecture once, the lecture topic and questions were chosen to minimize the impact of variation in student backgrounds and their understanding of the content of the material.

**Outcomes**

A paired-samples t-test was used to evaluate whether there was a statistically significant difference between showing the captions in a static location versus showing the captions above the presenter. First, we compared the ease of readability (Question 1) when presented with traditional Real-Time Text Display (RTD) and our Real-Time Tracking Display (RTTD). There was not a significant difference in the scores for ease of readability of RTD (M = 3.308, SD = 0.947) and RTTD (M = 4.0, SD=0.707); t (13) = -1.966, p = 0.069, as shown in the table below.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>CI 95%</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTD was easy to read</td>
<td>13</td>
<td>3.308</td>
<td>2.793 – 3.823</td>
<td>0.947</td>
<td>2.0</td>
<td>5.0</td>
</tr>
<tr>
<td>RTTD was easy to read</td>
<td>13</td>
<td>4.000</td>
<td>3.616 – 4.384</td>
<td>0.707</td>
<td>3.0</td>
<td>5.0</td>
</tr>
</tbody>
</table>

We repeated the test to compare the helpfulness (Question 2) in understanding the lecture when presented with traditional Real-Time Text Display (RTD) and our Real-Time Tracking Display (RTTD). There was a significant difference in the scores for helpfulness of RTD (M = 3.000, SD = 1.080) and RTTD (M = 4.000, SD=0.751); t (13) = -2.944, p < 0.05, as shown below.
Next, we repeated the test to compare the students’ likelihood of recommending the system to other students (Question 3) after using the traditional Real-Time Text Display (RTD) and our Real-Time Tracking Display (RTTD). There was a significant difference in the scores for recommendation preference of RTD ($M = 2.923$, $SD = 0.954$) and RTTD ($M = 4.077$, $SD=0.954$); $t (13) = -2.560$, $p < 0.05$, as shown below.

In summary, the participants did not significantly differ with regard to the ease of reading the captions whether presented with the traditional RTD or when it tracked the presenter with RTTD. However, the participants reported that RTTD was significantly more helpful than RTD and this, in turn, explains why they are significantly more likely to recommend RTTD over RTD. Finally, it is important to mention that a small sample size was used in this study. It is likely that a larger sample would have resulted in clearer differences between the two display methods.

The students reported that while it was easy to move their gaze between the display and presenter, it was hard to predict where the display would be, as the presenter moved his head often. The student comments suggest that following the movement of the captions needs to be smoothed to be easier to follow. Some students also said two lines are not enough. The most likely explanation is that they are not able to read fast enough, as they may not have had strong English comprehension. This suggests that the number of caption lines displayed should be adjustable in order to better serve the needs of students.

**Conclusion**

This paper has identified perceptual challenges faced by DHH students who use real-time text in engineering courses. Our survey showed that DHH students benefit from RTTD and would recommend it to their peers. In addition, we showed that this approach reduces visual dispersion and empowers students to sit anywhere they want. RTTD also creates a more inclusive and versatile classroom environment.

The takeaway from this study is that participants find the first version of the RTTD system significantly more helpful, but not necessarily easier to use. We will address ease of use in our future work, with matriculated students in engineering courses. If this approach is proven to be beneficial, it will allow DHH students to focus on both instructor/presentation and text display simultaneously.
Based on our findings, we plan to continue this research to identify ways to improve the RTTD system. Specifically, we will use eye-trackers as a tool to accurately measure and collect data on eye movements while following the projected real-time text. While the early focus of this work is in the classroom and on deaf students in particular, there are other universal access applications for this technology related to panel presentations, conference settings, and theatre productions.

References


