Winter 11-13-2013

Improving Deaf Accessibility to Web-based Multimedia

Brent N. Shiver
DePaul University, bshiver@gmail.com

Improving Deaf Accessibility to Web-based Multimedia

Brent Shiver

Submitted to the Faculty of the College of Computing and Digital Media at DePaul University

Principal Advisor:	Rosalee Wolfe, PhD
Committee Advisors:	John McDonald, PhD Evelyn Lulis, PhD Scott Grissom, PhD

November 13, 2013

Accepted by the faculty of the School of Computing, DePaul University, in partial fulfillment of the requirements for the Doctor of Philosophy Degree.

Doctoral Committee:

_____
Dr. Rosalee Wolfe, Chair

_____
Dr. John McDonald

_____
Dr. Evelyn Lulis

_____
Dr. Scott Grissom

November 13, 2013

Acknowledgements

This work would have been impossible without guidance and support from my advisor, Dr. Rosalee Wolfe. She is one of the most caring and selfless people I know. She always provided excellent advice when I needed it, assisted with proofreading, and helped pull me out whenever I got stuck. She also never stepped away from efforts to learn and practice sign language so we could have conversations together without an interpreter present. It is statistically improbable that I was able to find an advisor at DePaul who also happens to have so much interest in and respect for Deafness and ASL. I have a deep appreciation and gratitude for all that she has done for me over the years.

Furthermore, I would like to express appreciation to the doctoral committee, Dr. John McDonald, Dr. Evelyn Lulis, and Dr. Scott Grissom, for taking the time to collaborate and review my work. They were able to bring me a step closer to making Internet more accessible to the deaf population. Without their involvement, I wouldn't have been able to complete this dissertation.

I want to thank the ASL Synthesizer Team for taking a chance with me and welcoming me as a member. I have really enjoyed making contributions to the project especially with the usability tests. Special thanks to Dr. Karen Alkoby, who introduced me to Dr. Wolfe and the team. She, who is also deaf, was the one who convinced me to take a chance with the doctoral program. I want to also thank Dr. Jerry Schnepp, Dr. Jorge Toro, and Dr. Mary Jo Davidson who allowed me to assist with their research and helped make my life easier by answering all my silly questions. On several occasions Jerry would tease me by sending me lmgtfy.com links. I will miss you guys.

Education would have been next to impossible without the sign language interpreters provided by DePaul. They were expected to translate spoken words from difficult topics into sign language and voice for me. I was fortunate to have interpreters with a great deal of experience and talent. A very special appreciation goes to Joshua Ayoub, who interpreted for countless of meetings with Dr. Wolfe and my dissertation proposal defense. He will also interpret at my defense. He often would make changes to his schedule so that he would be available to interpret for me.

I would like to offer appreciation to my former employer, Allston Trading, for generous tuition support over the years. I would like to acknowledge the fellowship awards received from District of Columbia NASA Space Grant Consortium. Finally, special thanks to College of Computing and Digital Media for tuition, research, and travel support during my tenure.

I am truly blessed to have many wonderful friends who help make my life more exciting and interesting. They helped me maintain my sanity and bring balance to my life. It is truly wonderful to be able to have friends whom I can communicate freely and at ease with. This is a special salutation to my friends.

Very special thanks to Shannon Kennedy who gave extra support and encouragement during the last mile. Whenever I had a bad day with the study she would remind me that the finish line is just around the corner. We share similar upbringing and experience as deaf individuals so she was able to relate with my challenges. She was always patient and understanding at times when I had to shift my complete attention to research and writing. Thank you, Shannon, for believing in me and motivating me to stay on track.

I am indebted to my loving parents, Buck and Kyle Shiver, who have always had faith in me. When I became deaf at age two, they were told by educators, doctors, and other so-called professionals on deafness not to expect much. They chose to not to believe them and never, for once, let my deafness stand in the way of gaining the education and knowledge to get where I am today. My parents learned how to sign so we could all communicate easily and were always available so that I could ask for help with homework. They also fought with schools and the school system so I could get a sign language interpreter for classes and extracurricular activities. They, with my sister Amanda included, always gave unconditional love, support, encouragement, hope, and everything. I cannot thank them enough for all that they did for me.

Abstract

Improving Deaf Accessibility to Web-based Multimedia

Brent Shiver

DePaul University, 2013

Internet technologies have expanded rapidly over the past two decades, making information of all sorts more readily available. Not only are they more cost-effective than traditional media, these new media have contributed to quality and convenience. However, proliferation of video and audio media on the internet creates an inadvertent disadvantage for deaf Internet users. Despite technological and legislative milestones in recent decades in making television and movies more accessible, there has been little progress with online access. A major obstacle to providing captions for internet media is the high cost of captioning and transcribing services.

To respond to this problem, a possible solution lies in automatic speech recognition (ASR). This research investigates possible solutions to Web accessibility through utilization of ASR technologies. It surveys previous studies that employ visualization and ASR to determine their effectiveness in the context of deaf accessibility. Since there was no existing literature indicating the area of greatest need, a preliminary study identified an application that would serve as a case study for applying and evaluating speech visualization

technology. A total of 20 deaf and hard-of-hearing participants were interviewed via video phone and their responses in American Sign Language were transcribed to English.

The most common theme was concern over a lack of accessibility for online news. The second study evaluated different presentation strategies for making online news videos more accessible. A total of 95 participants viewed four different caption styles. Each style was presented on different news stories with control for content level and delivery. In addition to pre-test and post-test questionnaires, both performance and preference measures were conducted.

Results from the study offer emphatic support for the hypothesis that captioning the online videos makes the Internet more accessible to the deaf users. Furthermore, the findings lend strong evidence to the idea of utilizing automatic captions to make videos comprehensible to the deaf viewers at a fraction of the cost. The color-coded captions that used highlighting to reflect the accuracy ratings were found neither to be beneficial nor detrimental; however, when asked directly about the benefit of color-coding there was support for the concept. Further development and research will be necessary to find the appropriate solution.

Table of Contents

List of Tables

List of Figures

1 Introduction

Internet technologies have expanded rapidly over the past two decades, making all types of information more readily available. It has created a shift from traditional media such as mail, newspapers, books, television, and movies, towards e-mail, online news, e-books, and online videos. Not only are they more cost-effective than traditional media, new media has contributed to quality and convenience, as well as fostered creativity and opportunities. However, proliferation of video and audio media on the Internet creates an inadvertent disadvantage for deaf Internet users.

Despite technological and legislative milestones in recent decades making television and movies more accessible (Ellcessor, 2011), there has been little progress with online access for the deaf and hard-of-hearing. A major obstacle to providing captions for Internet media is the high cost of captioning and transcribing services. A recent development which has the potential for positive change is the passage of the Twenty-First Century Communications and Video Accessibility Act of 2010, which was signed into law by President Obama (U.S. Congress, 2010). Part of the law's purpose is to help make online content more accessible for the deaf. However, even with assistance from the legislative front, it is virtually impossible to manually caption every single video, or audio clip, on the Internet due to the staggering cost.

A possible alternative lies in automatic speech recognition. This study investigates possible solutions to Web accessibility through presentation strategies relying on the utilization of automatic speech recognition technologies. It surveys previous studies that

employ visualization and automatic speech recognition to determine their effectiveness in the context of deaf accessibility.

1.1 Captioning History

Long before captioning became a reality, deaf people imagined screens accompanied by captions on the bottom. Emil S. Ladner Jr., a deaf high school student, wrote in the American Annals of the Deaf in 1931 (Downey, 2008):

> Perhaps, in time, an invention will be perfected that will enable the deaf to hear the "talkies," or an invention which will throw the words spoken directly under the screen as well as being spoken at the same time.

It wasn't until the 1970s that captions on television sets finally materialized. Some shows were captioned, but it was not popular with hearing audiences. The visibility of captions was distracting to hearing users and it could not be turned off. This led to invention of closed captioning (CC)[1] which enabled users to control the visibility of captioning on television sets. It is encoded and broadcasted using line 21 of Television Vertical Blanking Interval signal, authorized by the FCC in 1976 to be used exclusively for captioning. The National Captioning Institute (NCI) offered first-ever closed-captioned television series on March 16, 1980. The first series captioned included *Masterpiece Theatre, Once upon a Classic* by PBS, *The Wonderful World of Disney* by NBC, and *ABC Sunday Night Movie* by ABC. In 1982, real-time captioning was made available by NCI. Despite the significant milestone with captioning, the bulky TeleCaption CC decoders by Sears, Roebuck and Co. were not cheap. In 1990, Congress passed the Television Decoder Circuitry Act, which mandated all new TVs 13 inches, or larger, manufactured for sale in U.S., include caption-decoding

[1] A glossary of terms and acronyms is contained in Appendix A.

technology available at a nominal cost. No longer were deaf consumers required to purchase separate caption decoders that often cost more than television set.

However, the advent of new television technologies has effectively thrown deaf accessibility back to square one. Flat screen televisions, high-definition television shows, and Blu-ray players have become popular and they all require HDMI cables for optimal picture quality. Unfortunately, HDMI technical specifications were not designed to carry line 21 closed captioning signals. In order for HDMI devices to display captions, the captions have to be decoded before being transmitted through the HDMI cable. For instance, TiVo and cable boxes must have the caption decoder option enabled in order for captions to work. Virtually none of the Blu-ray and DVD players have caption decoding capability and instead support subtitles offered by the disc. They are usually identified as SDH (Subtitles for the deaf and hard-of-hearing) or English subtitles on disc cover boxes (Bartolome & Cabrera, 2005). In addition to television captioning, deaf individuals also benefit when subtitles are available for audio sources, such as movies, gaming, and scoreboards. Both CC and SDH usually include environmental cues that make it possible to indicate events such as doorbell, telephone ringing, music, whispering, and other unusual sounds. On the other hand, typical English subtitles usually do not come with such cues and are more geared towards international audiences; they are essentially a transcript of English spoken words.

Captioning services such as NCI create and incorporate captions into television shows before they are broadcast, but in some situations require captioning in real-time. These services, known as Communication Access Real-time Translation (CART), provide captions at events including classrooms, seminars, and meetings (Caption First, 2010). The process usually is two-fold: a skilled stenographer uses a device to translate speech into

shorthand strokes, and then a machine converts the shorthand into written English for display on screen. Although the CART services are very useful and provide speech accessibility to deaf participants, they are labor-intensive and expensive.

Recently, two web-based services have appeared with the goal of encouraging volunteerism in captioning media. These services may help pave the way to improved Web accessibility. Overstream (Hyatt, 2010) allows users to add captions to a set of supported video providers, and make them available to everyone. This is an excellent solution because it leverages the power of user community. Everyone contributes something and everyone benefits. However, this approach relies on volunteers to caption content manually. A similar feature is also supported by YouTube (Lowensohn, 2009), a popular video uploading and sharing website owned by Google, which allows owners to add captions. Similar to Overstream, this also requires manual intervention.

Taking it a step further, Google added the capability for users to upload English transcripts without time codes; its automatic speech recognition (ASR) technology would be applied to synchronize the captions with videos (Sutter, 2010). This is a significant time-saver over manually time-coding the transcripts to video. It employs Google Voice, a speech-to-text engine originally used to convert voice mails into text. Another promising development is Google's recent introduction of an automatic captioning feature available on some YouTube videos. It also uses the Google Voice engine to perform ASR transcribing tasks to make YouTube videos more accessible to deaf and hard-of-hearing users. Google recently received an Accessibility Award for their contribution to the betterment of Internet media access at the 2010 National Association of the Deaf Conference (National Association of the Deaf, 2010). Ken Harrenstien, a deaf engineer at Google involved with the project, believes this feature is a major milestone that could open doors to more accessibility on the

Internet, but acknowledges the accuracy issues that are inherent with ASR technology that tries to understand speakers from various backgrounds.

1.2 Challenges of Captioning on the Web

The major challenge of manually captioning is cost. Captioning a video costs approximately $10–$30 per minute (Custom Captions, 2010). The costs cover skilled captionists who not only transcribe the audio content, but also keep video's text and audio in sync. Transcriptions, on the other hand, are written texts of speech in its entirety and are not synchronized with video. Although they cost less than captioning, they remain expensive, usually costing $2–$5 per minute. Unless it makes business sense, or there is a legal obligation that requires the owner of the video to add captions, it is unlikely the owner will have the incentive to spend money on captioning.

A cost-effective alternative to manual captioning is automatic speech recognition (ASR) technologies. The term describes systems that translate audio content to text material. Applications of ASR include uses in the military and healthcare, as well as automated call centers, and for people with mobile challenges.

The first speech recognizer, developed in 1952 by Davis, Biddulph, and Balashek of Bell Laboratories, identified single spoken digits (Juang & Rabiner, 2005). The device utilized natural modes of resonance, also known as formant frequencies, as patterns to link unknown spoken digits to the best matching digit. This important milestone would not have been possible without work done by Fletcher and colleagues at Bell Laboratories during the first half of the 20th century. They found a relationship between the speech spectrum, which involved the power distribution of speech sound across frequencies, and

speech characteristics such as phonemes. In the 1930s, Homer Dudley of Bell Labs built a speech synthesizer called VODER (Voice Operating Demonstrator) that was demonstrated at the 1939 World Fair in New York City.

In the 1970s, Tom Martin founded Threshold Technology, the first commercial speech recognition company. They created the first actual ASR product called the VIP-100 System. Despite having only a few applications such as package sorting on a conveyor belt by FedEx, it gained the attention of the Advanced Research Project Agency (ARPA) of the U.S. Department of Defense. The agency eventually provided funding to the Speech Understanding Research (SUR) program during the early 1970s. One of the first benefactors of the program, Carnegie Mellon University, created the Harpy system that was able to recognize speech using 1,011 words with satisfactory accuracy. The system was possibly the first to use a finite state network making computation more efficient (Juang & Rabiner, 2005).

While ARPA encouraged speech recognition research in academic institutions, commercial companies IBM and AT&T Bell Laboratories took different approaches to speech recognition for their applications. IBM's motivation was driven by the need for a speaker-dependent system that involved a large recognition vocabulary. The goal was to develop a low-cost replacement for a human stenographer, who would take notes in short hand and manually key them in on a typewriter. The system, dubbed Tangora, required training by each individual user so it could understand the user's speech with greater accuracy. It utilized a language model that involved statistical probabilities for a given ordered sequence of words that would be correct. It was also known as an n-gram model, which defines word relations within a span of n words, and is very popular in large vocabulary ASRs. The end result is usually in the form of a transcription.

In contrast, AT&T Bell Laboratories was interested in developing a speaker-independent system. It wanted to be able to handle speech from various speakers from different regions that involved different accents. The primary goal was to create a system that would automate tasks such as voice dialing and routing calls to correct departments. Keyword spotting is employed so callers can use natural language. For example, "I'd like to charge it to my credit card", the system would recognize "credit card" and perform the appropriate task. The tradeoff of a speaker-independent system was an extremely limited vocabulary set that the system could recognize.

Speech recognition research in the 1980s and 1990s evolved from an emphasis on linear pattern recognition to an intensive statistical modeling methodology. Hidden Markov Models (HMM) were introduced and became a popular method to optimize speech recognition accuracy (Picone, 1990). It mitigates the problems associated with speakers of various educational levels, dialects, and background noises. Applying the information obtained from recent speech sequence and known vocabulary, it determines best possible speech matches for next unknown utterance which would eventually translate into a valid word. This technique was found to be effective when combined with a finite state network. Carnegie Mellon University (2012) developed the Sphinx system that successfully incorporated HMM into network search offered by its predecessor, the Harpy system. Another system, Hidden Markov Model Tool Kit (HTK), also made available by Cambridge University, provided software tools to facilitate speech recognition research around the world. Data sets, including speech vocabulary, concepts, and algorithms, used by systems to recognize speech expanded significantly in the two decades and continue to grow steadily today (Juang & Rabiner, 2005).

Some of the improvements may be attributed to the continuous introduction of more powerful computers and better vocabulary databases. However, accuracy is the major issue that dictates the success of automatic speech recognition. A common measurement to describe the accuracy of a recognizer is word error rate (WER). There are several factors that affect how well a recognizer can correctly identify patterns, including subject's speech characteristics, level of noise from environment, and engine vocabulary. Because natural language often requires a level of interpretation and judgment that a machine cannot perform, human involvement may be necessary to make corrections and improve accuracy. In addition, training the software to understand how to map user's speech behavior to text vocabulary often requires a significant time commitment (Juang & Rabiner, 2005).

There are two scenarios that permit speech recognition technology to maintain a WER low enough that the results are useful. A large vocabulary that covers a breadth of topics necessitates training by an individual speaker. To be speaker independent, an ASR system has to severely limit the size of the vocabulary. Unfortunately, the task of making web-based media more deaf-friendly requires both speaker independence, since it will need to accommodate speakers in all media, and, since the media are not limited to specific topics, it will need to recognize a large vocabulary. An ASR technology that translates speech into text for better deaf accessibility to the Web cannot have restrictions on either speaker or vocabulary, because Web media contain audio information for thousands of speakers on thousands of topics. Unfortunately, accomplishing the goal of a speaker-independent ASR capable of recognizing a large vocabulary has continued to be a herculean task.

Since 1996 National Institute of Standards and Technology (NIST) has been inviting researchers from companies and universities to participate in the Speaker Recognition

Evaluation (SRE) every 1–2 years (National Institute of Standards and Technology, 2011). The goal is to establish benchmarks and measure progress over time on systems that support large vocabulary without speaker training. Over 40 research sites from all over the world, including Massachusetts Institute of Technology, Carnegie Mellon University, and IBM, have participated and contributed to the trials (Martin & Greenberg, 2008). Despite collective and collaborative efforts, results are still far from accomplishing a consistent 2–4% WER that is considered within range of typical human error in transcription. According to NIST STT Benchmark Test History graph (see Figure 1), the best system could only maintain a WER of 10% while many state-of-art systems, covering other speech tasks, have much higher WERs. Typical speech recognition software usually provides ability for users to listen and make corrections.

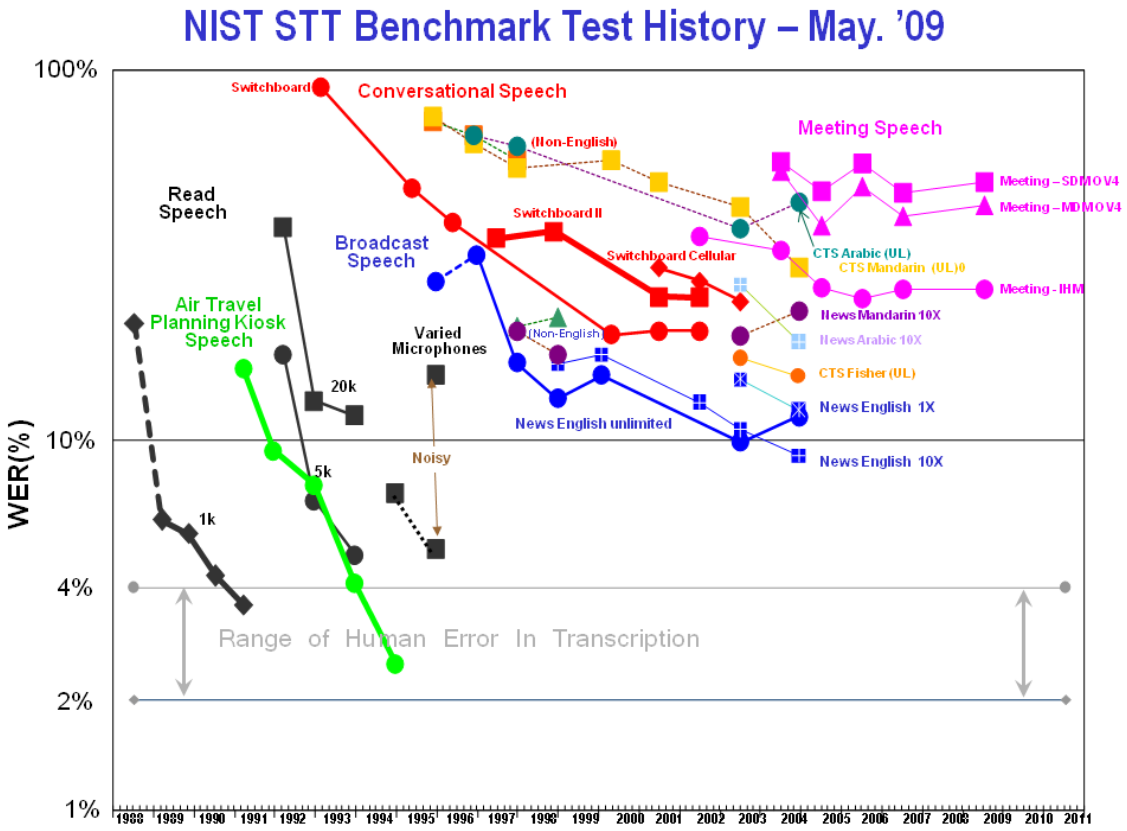


Figure 1: After decades of research the best systems from NIST trials continue to have high WERs (National Institute of Standards and Technology, 2011).

Despite advancements in ASR technologies, they are geared toward customers who have the benefit of being able to hear. While the software produces the transcriptions, the hearing users are able to catch errors while listening to themselves or the recording, and then are able to make corrections as needed. Unfortunately, deaf users do not have this benefit and will not know whether the transcribed text is accurate or incorrect, thus making it difficult, or nearly impossible, to trust the veracity of the transcription.

However, there is additional information that could be a valuable tool. All ASR software use probabilities, or confidence levels, to determine translations, but they are

discarded and users do not have access to them. This has the potential to benefit deaf users because it would include information on which passages are likely to be accurate and which other portions are likely to be incorrectly transcribed. This data would help add context to the translated text and empower the users to make decisions.

Despite not being able to hear, deaf users often cope with incomplete information particularly when lip reading. Because many words look very similar visually, such as "hat" and "at", and "bad" and "mad", deaf lip readers have to fill in the spaces with available information (Nitchie, 2006). Since only about 30% of spoken words are visually recognizable, deaf lip readers will need to guess the rest. Given their ability to fill in the gaps with the information available, they may be able to do the same with ASR transcriptions. If they are armed with knowledge of the confidence levels of each word, their ability to correctly interpret the translations may increase dramatically.

1.3 Speech Recognition Visualization

At times, when processing sound, speech recognition software may not be able to identify the words. This may be due to a speaker talking too rapidly, or using an atypical or unexpected word. In these cases, there may be multiple interpretations of the word being spoken. Deaf users do not have the option of reviewing the recorded speech and checking it against the recognized text for ambiguities, or errors. Typical commercial speech recognition software does not indicate that an ambiguity exists through visual means, such as annotated text or listing possible translation alternatives.

Despite enormous potential benefits to the deaf population, the available literature that focuses on ASR and deafness is scarce. Virtually every study that evaluates

visualization strategies of text created via ASR involves hearing users. However, it is useful to examine visualization of speech recognition for hearing users for approaches that could potentially be useful to deaf users. This section discusses possible visualization ideas and alternatives for ASR output. The following two studies involved hearing users.

Vertanen & Kristensson (2008) investigated possible benefits of employing an underlying visualization to emphasize low-confidence. The low-confidence implies that the recognizer's interpretation may have been done in error. This approach was created in the hope of lowering the cost of creating transcripts for spoken speech. An initial transcript was captured through ASR, and a human editor then read and corrected errors in the transcript while listening to the recorded speech. The goal of the approach was to help the editor to catch more errors in a time-effective manner. Visualization techniques were employed to indicate the speech recognition engine's confidence in the produced text (see Figure 2). Red underlines indicated words with low confidence, and the darkness of the underlines was proportional to the lowness of the confidence. This visualization helped users identify potential errors in only a limited number of cases. The users would catch errors more often only if low-confidence text were correctly flagged. On the other hand, if text was incorrectly identified and not underlined as low-confidence, chances were greater that the users would miss the problem. The authors concluded that it was possible that the users placed too much faith in recognizer's ability to present annotations accurately.

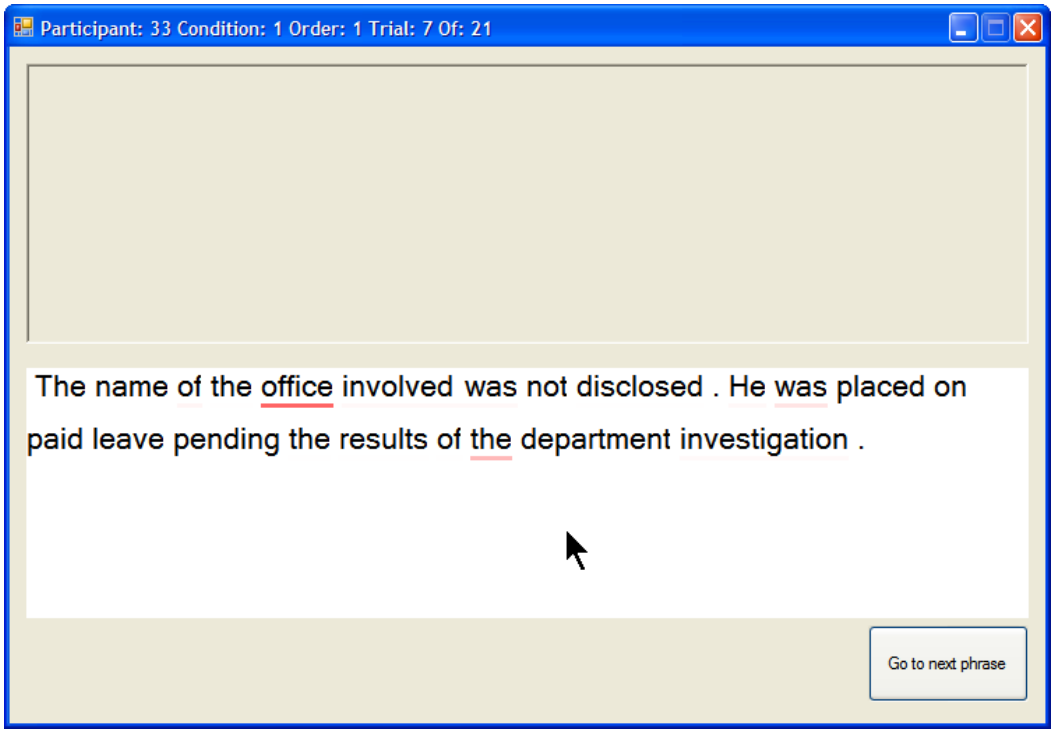


Figure 2: Shades of red underline is applied to words with low confidence. The word "office" has a lower confidence than the word "was" that appears in the second sentence.

A case study conducted by Collins, Penn, & Carpendale (2007) focused on uncertainty visualization through utilization of lattices to support decision-making which involved a multilingual chat application that used an automatic translation engine. The goal was to provide possible choices through visualization and empower users to choose a translation that makes most sense, or discard it altogether. Lattices were generated as representation of possible translations and included confidence levels through fill color and border thickness. Although the study involved spoken language translation, it has a close resemblance to automatic speech recognition. See Figure 3 for a screenshot of the application. The system was demonstrated at Computer Supported Cooperative Work (CSCW) in 2006. Although user testing was not performed, they collected informal user feedback. The participants expressed general interest in the visualization of uncertainty so

they could make appropriate decisions. For words for which the software could find no appropriate target word, the software substituted a picture instead. The users commented that this feature was useful for words that were not translatable.

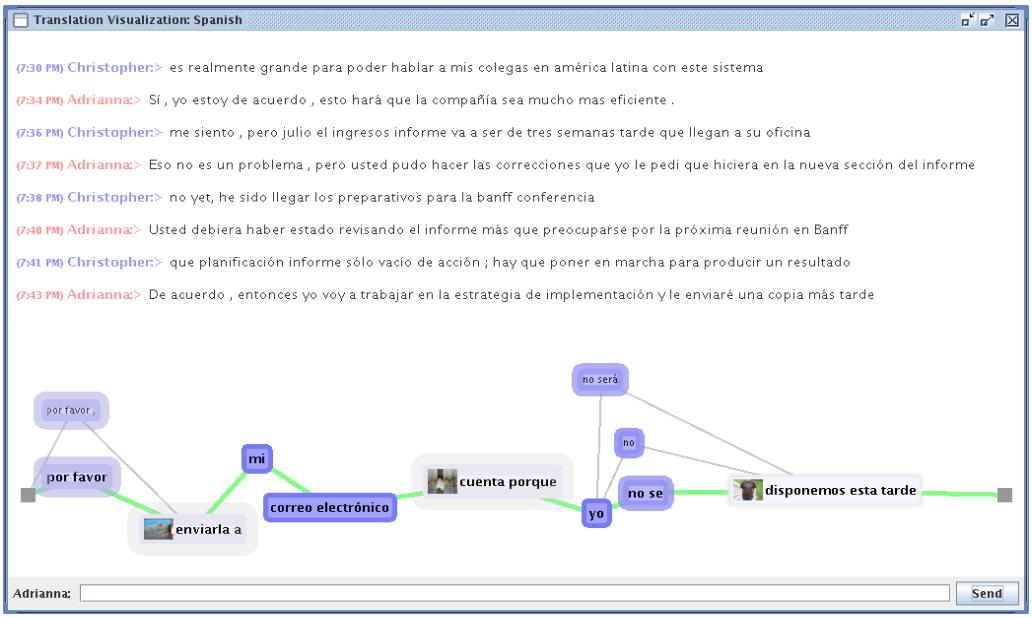


Figure 3: In this example, the source language is English and the target language is Spanish.

1.4 Utilizing Speech Recognition to Aid Comprehension

Wald (2006) explored the possibility of utilizing Automatic Speech Recognition (ASR) to aid classroom learning specifically for students with disabilities including deaf, hard-of-hearing, blind, and dyslexic. This is the only extant study that explores any aspect of ASR visualization for increased deaf accessibility. It also investigated the benefits of using ASR to enhance quality of learning and teaching for students without disabilities. It found that one of the problems with real-time speech-to-text synthesis was a lack of punctuation. Without punctuation, the ASR-created transcripts were difficult to read and

understand. A workaround was to add single and double spaces to the transcripts as visual cues of brief pauses and long pauses respectively.

The study also discussed *Liberated Learning*, an approach that involves lecturers investing additional time to train ASR with their voice, adding new vocabulary to system, and correcting errors so they don't occur again. After lectures, the teacher edits transcripts to remove errors and uploads the material online.

Wald's implementation of Liberated Learning used IBM ViaScribe with a ViaVoice engine. The software has several features including creating a standard file format (SMIL) to provide synchronized captioning, supporting spontaneous speech, and reading from online speech files (Bain, Basson, Faisman, & Kanevsky, 2005). Leitch & MacMillan (2003) reviewed the effectiveness of Liberated Learning Initiative with 44 students with physical, sensory, and cognitive disabilities. They found that students and lecturers believed the method was helpful with their learning and teaching experience as long as the accuracy rate was reasonable, or at least 85%. Many students developed coping strategies to deal with errors and most of them found transcriptions a useful supplement to the lectures. Several participants commented that they were able to "get the gist" of the lecture despite the errors. Having other parts that provided context was helpful in understanding lectures.

The accuracy rate of ASR poses a huge challenge to comprehension. It is often difficult to reach an accuracy rate of over 85% in higher education classroom environments from instructors' speech (Leitch & MacMillan, 2003). However, the rate is improved when the speech source is from an office environment. Rate of delivery is cited as a possible reason for this pattern. Accuracy can be improved through post-speech activities including editing and re-voicing by an individual. Errors are manually edited through inserting,

deleting, or amending. The editor is able to decide how to prioritize those that impact readability and comprehension the most. Another possible editing technique may involve the user listening to the original speaker and repeating to ASR in a quiet room which is known as re-voicing. In fact, this method has been used for live television subtitling in United Kingdom (Lambourne, Hewitt, Lyon, & Warren, 2004). It has also been utilized in classrooms and courtrooms in the United States (Francis & Stinson, 2010)

Another study, which was published at approximately the time as Wald's, investigated the benefits of using ASR transcripts to aid hearing users in the search and review of webcast archives (Munteanu, Baecker, Penn, Toms, & James, 2006). It conducted a 48 within-subjects design under four conditions: 1) perfect transcripts, 2) transcripts with 25% Word Error Rate (WER), 3) transcripts with 45% WER, and 4) no transcript. The participants were given 12 minutes to listen to a 38-minute webcast and complete the quiz. Because they didn't have time to listen to the entire webcast, they relied on given transcripts. The data showed that ASR accuracy linearly influences both user performance and experience. Transcripts with an error rate of 25%, or less, was better than no transcript at all, but those with 45% WER were not beneficial.

A similar audio browsing study was conducted by Vemuri, DeCamp, Bender, & Schmandt (2004). It applied time-compression techniques to audio files as constrasted with (Munteau et al., 2006) who imposed a time limit to complete a quiz. It explored the benefits of utilizing ASR transcripts to aid in the review of the time-compressed files. Searching through audio files at normal speed takes a prohibitive amount of time since average rate of an English speaker is only 180 words per minute (WPM). In constrast, reading speed could be achieved at about 400 WPM, i.e., twice as fast.

Vemuri prepared the time-compressed files by applying synchronized overlap-add fixed synthesis (SOLAFS). This technique increases the WPM of a speaker without impacting the sound pitch. As the rate was increased from 120 to 810 WPM, the listener comprehension decreased linearly.

IBM's ViaVoice ASR software was then used to generate transcripts as a supplement to the time-compressed audio files. The generated transcripts had a WER of 16% to 67% (mean=42%, sd=15%), which was comparable to other commerical ASRs. During the experiment, five different conditions were followed:

- C1: Perfect transcript
- C2: Transcript generated by ASR (using word brightness)
- C3: Transcript generated by ASR
- C4: Completely incorrect transcript
- C5: No transcript, audio only

The study identified C1 as being the best but costly option and is time-consuming and requires manual intervention. More cost-effective options C2 and C3, which were generated by ASR, were found to do nearly as well as C1 (see Figure 4). Transcript generated for C2 utilizing word brightness did not show any significant improvement over C3. Finally, as expected, performance under conditions C4 and C5 were the poorest, but interestingly, there was no statistical difference between them. A possible explanation may be that C4's transcript was so bad that the participants ended up ignoring it altogether. The researchers concluded that ASR transcripts improve comprehension when listening to time-compressed speech.

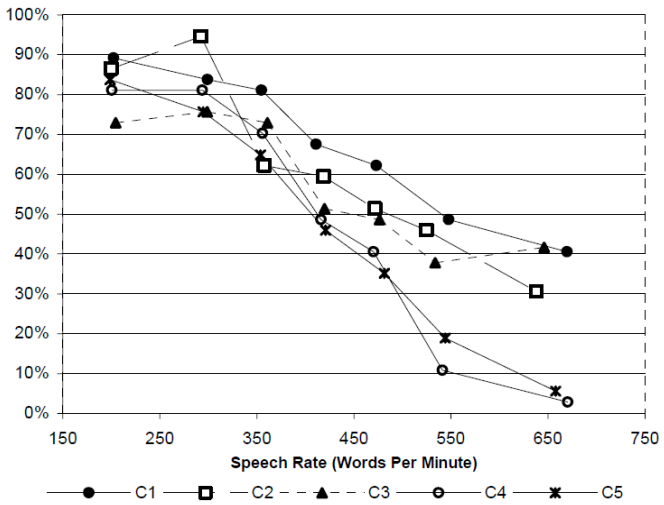


Figure 4: In this chart, C2 and C3 were shown to do nearly as well as C1 which used a perfect transcript.

1.5 Discussion

With exception of Wald's efforts, all of the studies placed emphasis on use of ASR transcripts to benefit hearing users. Wald wanted to improve classroom learning for students with disabilities including those with hearing loss.

In those studies using ASR for hearing populations, utilizing automated transcripts to retrieve time-compressed audio content and skimming through webcast archives facilitated greater comprehension. The visualization tool offered by Vertanen and Kristensson utilized shaded, red underlines to emphasize words having low-confidence. However, it is reliant on ASR being accurate; when ASR incorrectly recognizes words as having high confidence, its effectiveness deteriorates. Although the application created by Collins, Carpendale, and Penn (2007) did not involve using ASR, it visualized probabilities and made them available as lattice paths to offer decision-making to users. This approach

could be useful for an ASR tool that has the statistics, but needs an effective way to output the results.

An automatic speech-to-text tool designed specifically for deaf users has the potential to provide a better bridge to audio/video media. Although ASR has accuracy issues, several studies have shown that well-chosen visualizations have the potential to help users to glean additional information from error-laden texts.

2 Exploratory Study

A first step for developing better deaf accessibility for the Internet requires identifying an application to serve as a case study for applying and evaluating speech visualization technology. Since the current literature did not provide clear indication of a compelling choice, an exploratory study was necessary. The goal was to identify the areas of greatest need for enhanced deaf accessibility of the Web multimedia. The study proposal was evaluated and approved by DePaul Institutional Review Board (IRB) as noted by DePaul IRB #JR052311NUR. A copy of the application and approval can be found in Appendix B.

2.1 Participants

The exploratory study involved interviewing a total of 20 deaf and hard-of-hearing participants from various parts of United States. They were recruited primarily through email invitation (see Appendix C) and some others were contacted through social media such as Facebook. They were mostly contacts made previously at Deaf-related gatherings. A lowercase deaf refers to an audiological condition, and the uppercase Deaf refers to a particular group of deaf people who share a language—American Sign Language (ASL)—and a culture (Padden & Humphries, 1990).

Virtually all participants were profoundly deaf. Only one participant was hard-of-hearing. Fourteen interviewees stated that they were born deaf while five reported becoming deaf younger than 5 years old. Just three became deaf at the age of 5 years or older. Eleven of them were aged 30–39, while nine were 40 years or older. Thirteen were identified as male and seven female. The demographic questionnaire can be located in

Appendix D and the charts depicting demographic characteristics of participants are available in Appendix E.

2.2 Procedure

All interviews were conducted via video phone which is probably the most common means of distance communication in the Deaf community. The first step of the interview was to gain informed consent from the participant. The interview candidate was provided with an information sheet explaining the study and alerting that the interview may be recorded. The information sheet is in Appendix F. Any questions or concerns were discussed prior to the interview. After the initial procedure, each participant answered several basic demographic questions. During the interview, they were asked 13 open-ended questions regarding their experience in using multimedia over the Internet (see Appendix G).

2.3 Results

After the interviews were completed, the responses were transcribed from sign language to English. The responses were grouped by question and a card sort analysis (Fincher & Tenenberg, 2005) was conducted on the responses to each individual question to identify patterns of commonality in the responses. The following is a summary of user responses that correspond to each question.

1. When you read news articles on the Internet, do you ever watch the videos? Why or why not?

When asked whether they watch news-related videos on the Internet, all 20 interviewees agreed that news videos are not useful without captions. Some of them pointed out that they tend to read news articles as opposed to videos online. One interviewee emphasized that if there are no captions, it isn't worth watching. Another said, "It's not worth my time."

2. Have you watched YouTube?

All 20 users have visited YouTube at least once when using the Internet to watch videos. Out of 20, 15 watch it "sometimes". Five pointed out that videos are often not captioned. Six users preferred watching Deaf-oriented, or ASL-signed videos, and two mentioned teaching ASL using YouTube.

3. Have you ever found yourself needing the information on a video?

Everyone responded at least "sometimes" but 16 of them pointed to lack of captions. Five complained of being stuck because videos lacked captions, or were not accessible. Four users mentioned resorting to searching via Google, or another search engine, for texts related to the video.

4. Can you describe the type of video it was?

Fourteen respondents referred that the video they wanted was news-related and five mentioned CNN specifically. Knowledge about what is going on in the world was emphasized by four participants. Finally, four users mentioned interest in using training videos to increase knowledge and skills and keep up with current trends. Five users

complained of being referred to a page with video, but after clicking on the link they realized that it was a video without captions and thus the information was inaccessible.

5. What do you do to obtain the information or contents from videos?

All interviewees would resort to reading related text, or an article, when available. Half of them would use Google or search online for related text, article, and/or posts such as Facebook to learn more about contents from videos. Three users would attempt contact with source, or author, to request transcripts on video. Three participants would resort to asking an interpreter, or hearing person, to help translate selected videos.

6. How often are you frustrated about inaccessibility when you use the Web?

Sixteen users experienced frustration while four have either developed tolerance or given up altogether. Eight reported frustration every day, or often, and seven reported being frustrated sometimes. A participant succinctly described that the unavailability of captions stirred feelings of frustration to those experienced when the Internet was down.

7. Describe top three frustrations that you've experienced. What happened?

The top frustration was lack of captions on new video clips. Seventeen interviewees mentioned this. Five users pointed out non-captioned self-tutorials and e-learning videos as a problem. Five viewers complained about lack of captions on YouTube. Finally, three users expressed disgust when clicking on link only to be redirected to a video without captions

8. Are you familiar with Google automatic captions? If yes, please tell me about your experience.

Seventeen respondents had some experience while the other three didn't have any. Twelve pointed out that the captions had accuracy issues or too many errors, and four thought they were useless or too hard to follow. However, three users felt it was a good start and step in right direction. In addition, two said it was better than no captions at all. Two participants thought it might be useful for hearing people who spoke another language.

9. Have you used any other Automated Speech Recognition technology? If so, what was it? How did it work for you?

Fourteen respondents stated they have not used Automated Speech Recognition technology, five have used it, and one did not respond. Three users mentioned either Naturally Speaking by Dragon and iPhone's Siri feature, but they did not elaborate on how they used it.

10. Now I need to review a couple of items of terminology with you. Captioning is the process of displaying text on a television, video screen or other visual display. Captions typically show a transcription of the audio portion of a program as it occurs. A transcript is a document containing a complete written or printed version of content originally presented as a video or recording. Which approach do you prefer? What are the advantages and disadvantages of each?

Captions were preferred by 19 participants since the captions were usually better than transcripts. However, one lone participant said it is fine either way. Nine described reading transcripts as being harder to use and requires too much effort. Eleven pointed that captions made sense since they're always in sync with video.

11. Are there any situations where you prefer captions over transcripts and vice versa?

Seven interviewees favored captions in all situations while five preferred captions and mentioned that transcripts were only acceptable as a backup. Four pointed that transcripts could be more useful for reference, information verification, and research. Six would rather have captions since they are in sync with videos and are easier to follow.

12. We are currently investigating technologies that may improve accessibility. Which situations that you mentioned earlier do you feel this would benefit the most?

News-related videos were mentioned 15 times while four demanded all videos to be captioned. Four believed that all television shows should be accessible and three felt that investment, or financial clips, should be covered. Three emphasized any television shows or movies that are already captioned should also be captioned online.

13. Do you have any advice or suggestions in regards to improving accessibility on the Internet?

Five interviewees suggested additional government involvement such as legislation, lawsuits, and FCC, and three recommended educating people about the need since there may be a lack of awareness. Four demanded everything to be captioned. Three people suggested that any videos that have already been captioned should also be captioned online. Lastly, three thought speech recognition may be useful.

2.4 Results Analysis

Based on responses from twenty participants, the most common theme was the lack of accessibility for news online videos. Some interviewees were especially disappointed with well-known news outlets such as CNN, because they made no effort to make their videos accessible. The participants indicated general preference to have captions available as opposed to transcripts. The captions are designed to be synchronous with the videos, which makes it easier to follow. On the other hand, transcripts could be useful for rereading and reference work. However, it takes extra work to view the video and read a separate transcript back and forth. The general consensus was that captions shown on the video itself would make it easier to follow the video. Some participants did mention that having a transcript is better than nothing at all.

3 A Study to Evaluate Captioning Technology

The results of the interviews indicated a marked need for more accessibility to online news videos. This second study examined different presentation strategies for making online news videos more accessible to the Deaf community. The goal of this study was to answer the following research questions:

1. Although there is strong anecdotal evidence that captioning makes Internet media more accessible to the deaf and hard-of-hearing, there is no controlled study supporting this. Does the presence of captions make Internet media more accessible?
2. If captions for Internet media are created automatically via ASR, are they as effective as captions created by human transcribers? Does web media that has been automatically captioned in a cost effective manner provide the same level of accessibility as web media that has been captioned manually and is cost prohibitive?
3. Does knowing that the captions are automatically generated have an effect on a person's confidence in the captioning? Does it have an effect on a person's ability to access the content?
4. Does having a visual indication of the estimated accuracy of the automatic captions affect a person's ability to access the content?
5. Does having the visual indication make automatic captions more acceptable than automatic captioning having no visual indication?

The study was reviewed and approved by the DePaul University Institutional Review Board (BS031313CDM). The application materials and approval letters are in Appendix H.

3.1 Participants

A total of 95 deaf and hard-of-hearing people participated in the study. Since the test was entirely anonymous, there was no way to know if any of the 95 participants also participated in the initial exploratory test. These participants were recruited through Deaf mailing lists and previously established contacts in the Deaf community. Additional participants were discovered through word of mouth and forwarding of the solicitation email. They were all 18 years and older and had at least some college education.

3.2 Stimuli

The stimuli are the captioned videos and there are four caption styles:

1. Captions created through ASR using Microsoft's Speech Platform (Microsoft Corporation, 2012). It employed a visualization technique indicating the confidence level of the recognized text. Words with a higher confidence level were displayed in a more prominent color. Words with a lower confident level were displayed in a less prominent color. The color choices were indicative of the confidence level. The WER for this story was 20% (see Table 2).
2. Captions created through ASR, but without the visualization technique. As Table 2 shows, the WER for this stimulus was 12%t.
3. No captions
4. Manual captions created from the original text. The text appearing as captions was 100% accurate.

Participants viewed all four different caption styles, so they could compare among the four alternatives. Due to a potentially large transfer of learning effect, it was not

possible to use the same news story for all four caption styles. For this reason, the four news stories were controlled for content level and delivery, as discussed in the next paragraphs.

The process of creating the stimuli is a multi-step process, including

1. Identifying the stories
2. Recording the stories
3. Preparing the recordings for captioning
4. Adding captions to the recordings

3.2.1 Identifying the stories

One of the problems in using actual news stories as test stimuli is that viewers may have seen the story previously and thus may have prior knowledge of its content. To control for this possibility, this study utilized "pseudo news" stories by selecting material from standardized reading tests that were designed for students in the eighth grade. The material needed to be believable as a news story, but came from a reading test that had previously validated for level of difficulty. Four reading passages chosen from the standardized tests were converted into the news stories. The four passages identified came from the public schools of North Carolina Reading Comprehension Test (Public Schools of North Carolina, 2013) and the reading Florida Comprehensive Assessment Test (Florida, 2013). Test questions from the original standardized reading tests served as the basis for the performance metrics in this study. A detailed discussion of the measurements will follow in the "Independent variable and measures" section.

As an additional assessment of the difficulty of the content, this study applied the Flesch Reading Ease Formula (Flesch, 1948) and the Flesch Kincaid Grade Level Test

(Kincaid, Fishburne, Robers, & Chissom, 1975) to each of the pseudo news stories. The goal was to identify stories that are all at the same level of difficulty. Table 1 lists the four stories, their levels of difficulty, and their word counts.

Story	Flesch Reading Ease	Flesch-Kincaid Grade Level Test	Word Count
Bear sightings	61.8	9	389
Eagles	54.6	9.5	686
Mushrooms	55.5	9.6	620
Antarctica	69.5	9	508

Table 1: Four stories and their metrics

Very few validated reading tests are publicly available. Of all possible available reading passages, these four matched as closely as possible for level of difficulty and word count. For the Flesch Reading Ease metric, higher scores indicate easier reading passages. The maximal score is 206.8 and scores can go below -145.00. Typically scores of 60–70 indicate a passage easily readable by 13–15 years olds. The word counts are slightly longer than the median length of popular YouTube news videos (Journalism.org, 2013).

3.2.2 Recording the stories

Controlling for the delivery of the news stories included the choice of speaker, the speaker environment, and nonverbal considerations. Consistency was the key. This required using a single speaker to read the stories in a fixed environment. To control for nonverbal communication, the news reader used a neutral vocal tone and did not move his arms while reading from a simulated teleprompter (CuePrompter, 2012). The environment consisted of a neutral background, a chair for the seated news reader and a table in front to the news reader. The table was not visible in the test footage. The lighting was consistent

with the lighting found in modest news studios. See Figure 5 for a screenshot of the news reader and the environment.

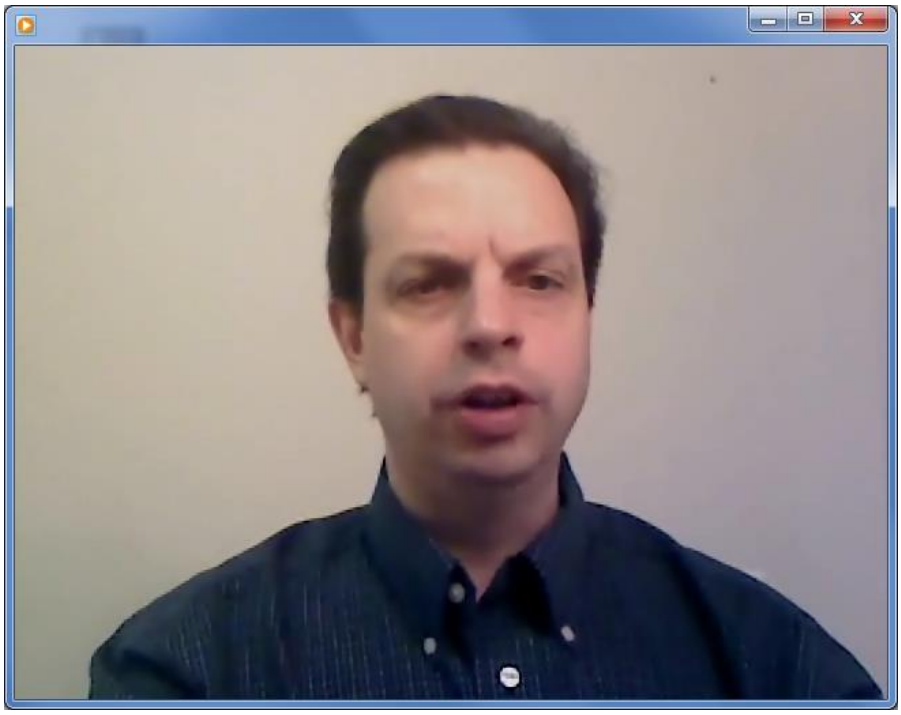


Figure 5: News reader with neutral background

3.2.3 Preparing the recordings for captioning

The recordings were edited and sized for use as Internet media. For the videos that were captioned via ASR, an icon was added to the upper left corner of the video. No additional graphics were used as they would add information external to the captions and could potentially influence user comprehension.

3.2.4 Applying captions to the videos

A description of the process of applying captions to videos is included here for completeness, but the development of new captioning technology is outside the scope of this study.

Automatic captions were created by removing the sound track from the video and using Microsoft Speech to create a SubStation Alpha captioning file (MultimediaWiki, 2013). SubStation Alpha format supports a selection of font colors. For the automatic captioning with visualization, a white font color indicated a 100% confidence and darker font colors indicated a lower confidence level from the speech recognition engine. See Figure 6 for a screenshot showing the visualization. To indicate that the captions were automatically generated, an icon was added to the upper left corner of the video. This is also visible in Figure 6.

Figure 6: Screenshot of video with visualized captions and ASR icon

To create the manual captions for the fourth caption style, the sound track was also processed using Microsoft Speech, and then manually corrected. The open source software Virtual Dub (VirtualDub, 2013) applied the captions to videos. None of the videos had a sound track. All but one of the videos had captions. Table 2 summarizes the process taken.

WER	Story	Action
20	Bear sightings	Captions generated by ASR. Visualization of confidence levels
12	Eagles	Captions generated by ASR. No visualizations.
100	Mushrooms	No captions
0	Antarctica	Perfect, manual captions

Table 2: Summary of stories and their WER

Table 2 also gives the word error rate (WER) for each video. The video with manual captions has a WER of 0% whereas the video with no captions has a WER of 100%. The

WER for the captions generated by ASR were measured using NIST's *sclite* (Information Access Division, 2013). According to Munteanu et al., (2006), these WER are below the level (25%) where the captured speech loses its utility.

3.3 Procedure

These tests were performed online. The online survey software was created from scratch in order to record usage data about the video controls. The media transport data from participants needed to be recorded. The software was developed using Visual Studio 2010 with C# using Microsoft .NET 4.0 Framework as the driving force. The YouTube SDK was used to control the video via the YouTube Javascript Player API (Google, 2013).

The application flow can be described as follows:

1. Cover page - This is where the user enters access code. This is passed in via URL get variable from the link sent via email.
2. Information sheet (Appendix I)
3. Warning page to not use back or refresh buttons
4. Pre-test questionnaire
5. For each story out of four:
    a. Video
    b. Content questions
    c. Preference questions
6. Post-test questionnaire
7. Thank you page with field to enter email address for gift card redemption

Figure 7 contains a screenshot of the interface when a participant is answering questions about a video. The video appears on the left, and media transport buttons appear underneath it. Appendix J contains screenshots of all steps in the process.

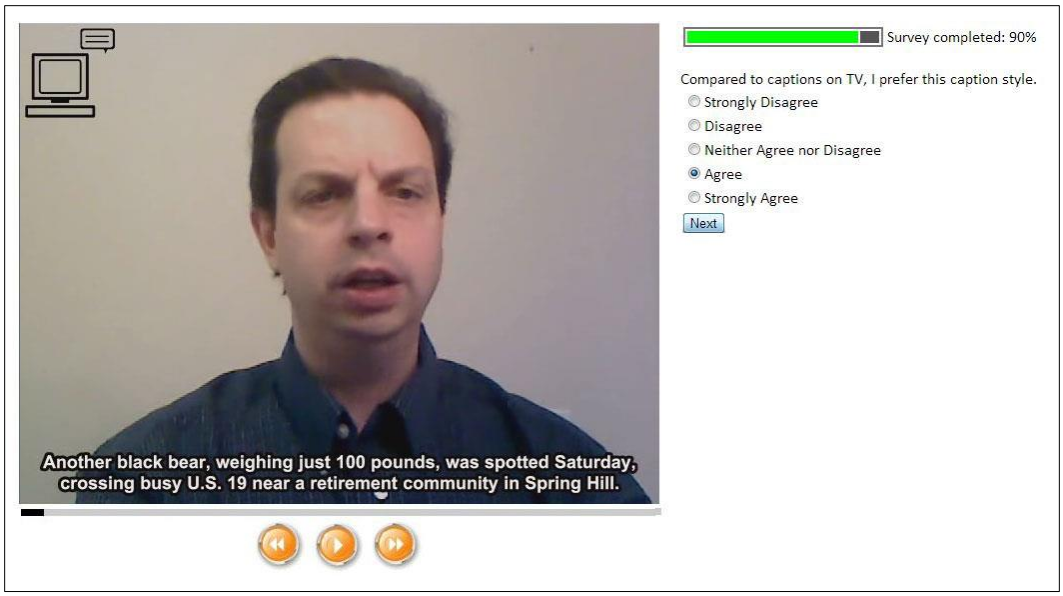


Figure 7: Main component of the survey application

After navigating to the test site, the participant read the information sheet in lieu of an informed consent, and if they wished to participate, clicked a button to continue. The participant then filled out a questionnaire that confirmed their eligibility to participate in the test. This questionnaire asked them to self-identify as deaf, hard-of-hearing, or hearing, queried the extent to which they use captions when watching television, and asked them to indicate their level of completed education. See Appendix K for the entire set of questions.

Participants viewed a series of four news videos that consisted of three captioned versions and one without. Both transcripts and questions of each news story can be found in Appendix L and Appendix M, respectively. At the end of each video, the participant answered questions, which were extracted from the reading tests, about the video's content. While answering the questions, the participant could replay the video as often as desired. After answering the content-related questions, they rated their viewing experience. See Appendix N for this list of preference questions. After the four videos, participants

answered a post-test questionnaire, found in Appendix O, which asked the participant to rate and compare all four captioning styles. . The thank you and gift card message displayed immediately following completion of survey can be found in Appendix P. A solicitation email, found in Appendix Q, was created and sent to help recruit participants.

3.4 Order of presentation

The captioning styles of 1) automatic captions with visualizations, 2) automatic captions without visualizations and 3) no captions were presented in random order. Lastly, participants viewed the video with the manual (perfect) captions. The randomization of captioning styles helps counter any transfer of learning effect as the participant goes through each video. The last video is the one with the manual captions. This video was deliberately placed last to take advantage of any transfer of learning effect and to serve as an upper bound for performance. This will help determine the differences between watching captions with inaccuracies and watching captions that are perfect. However, the participants were not aware that the last video contained manual captions, as contrasted to captions generated through ASR.

3.5 Independent variable and measures

The independent variable was captioning style. The dependent variables can be divided into

- Performance measures, including
    - Accuracy in answering the content-related questions
    - Time required to complete the content-related questions
    - Number of times the video is replayed

- Preference measures, including
    - Perceived understandability
    - User confidence in the accuracy of the captions
    - Perceived ease of use
    - Comparison of the four videos
        - Understandability
        - Visibility of ASR indicator
        - Desirability of color-coding to indicate accuracy

4 Results and Analysis

A total of 95 deaf and hard-of-hearing participants took the online survey between June 18 and July 9, 2013. They were recruited through email, social media, and word of mouth. The custom software recorded responses from the participants into a comma-separated values file.

4.1 Research Questions (review)

1. Although there is strong anecdotal evidence that captioning makes Internet media more accessible to the deaf and hard-of-hearing, there is no controlled study supporting this. Does the presence of captions make Internet media more accessible?
2. If captions for Internet media are created automatically via ASR, are they as effective as captions created by human transcribers? Is web media that has been automatically captioned provide the same level of accessibility as web media that has been captioned manually?
3. Does knowing that the captions are automatically generated have an effect on a person's confidence in the captioning? Does it have an effect on a person's ability to access the content?
4. Does having a visual indication of the estimated accuracy of the automatic captions affect a person's ability to access the content?
5. Does having the visual indication make automatic captions more acceptable than automatic captioning having no visual indication?

For questions 1 and 2, this study will examine three performance measures:

- Accuracy in answering the content-related questions
- Time required to complete the content-related questions
- Number of times the video is replayed

For questions 3–5, this study will examine the following preference measures

- Ease of reading
- Ease of understanding
- Confidence of captioning accuracy
- Preference of caption style
- Confidence in caption highlighting

The expected outcome for question 1 was “yes,” based on the strong support for captions by participants interviewed in the first study. Since there were improvements in ASR in recent years, the expected outcome for question 2 was “yes." Since the current perception is that ASR has improved, although not yet perfect, the expected conclusion for question 3 was “yes.” For question 4, the expected answer was “yes” since the goal of the color coding was to provide additional information about the accuracy of the automatically generated captions. This way the viewers would see the color and decide for themselves on the reliability of the words in the caption. Finally, for question 5, the expected answer was “yes," since the color coding would not hide any possible errors, and since people could judge the accuracy of the automatically generated captions for themselves, they would be more likely to trust it.

4.2 Participants

A total of 78 participants identified themselves as being Deaf while 17 aligned with the hard-of-hearing category. Thirty-eight wore hearing aids, 16 had cochlear implants, five had both, and 36 used neither. Fifty-four subjects, or a majority, were in the 18–29 age group, 23 belonged to 30–39 age group, 12 identified with the 40–49 age group, and six were 50 years and older. All participants, except for one, stated they always use captions while the lone participant admitted he uses captions only sometimes. As for educational background, 86 participants had at least 4-year college degree, eight had some college, and a lone participant finished high school/GED. For additional details on the demographic background, please visit Appendix R.

4.3 Performance measures

The formula for calculating the performance measures is given in Table 3. Each was calculated for every captioning style. For the "Time Required" measure, the per-question time was computed to compensate for the fact that each story did not have the same number of questions.

Measure	**Formula**
Accuracy	Percent of correctly-answered questions about the content of the story
Time Required	Average amount of time to answer a single content question
Number of Replays	Number of times the "Rewind" button was pressed

Table 3: Calculation of performance measures

The presentation of three of the styles (auto captions with highlighting, auto captions without highlighting, and no captions) was fully randomized to counterbalance any transfer of learning effect. The last captioning style, perfect manual captioning, should, in

theory, yield the best results because its captions were 100% correct as contrasted with the other styles where the captions had errors, or were missing altogether. It was deliberately presented last to capture any transfer of learning, and serve as an upper bound for the performance measures. This yielded six presentation orders as seen in Table 4. In addition, a review of stories, their associated story ID, and caption style can be found in Table 5.

Presentation order
1234
1324
2134
2314
3124
3214

Table 4: Presentation Orders. 1: automatic captions with highlighting, 2: automatic captions without highlighting, 3: no captions, 4: perfect manual captions

Story ID	Content	Caption Style
1	Bear sightings	Captions generated by ASR. Visualization of confidence levels.
2	Eagles	Captions generated by ASR. No visualizations.
3	Mushrooms	No captions.
4	Antarctica	Perfect, manual captions.

Table 5: Summary of videos and their caption styles

4.3.1 Was presentation order a factor?

A combination of Microsoft Office Analysis ToolPak (Microsoft Corporation, 2013) and McGraw-Hill MegaStat Excel (McGraw-Hill Glencoe, 2013) tools were used to conduct analysis on the study results. Using presentation order as the independent variable and accuracy as the dependent variable, a one-way analysis of variance (ANOVA) calculation

showed no significant difference among presentation orders. The single-factor ANOVA test was selected to determine whether there were significant differences between means. The analysis concluded that the order of presentation was not significant as seen in Table 6. Similar results were found with time required for all four stories and, finally, no significant difference was found for number of replays on all four stories. Visit Appendix S to review the complete calculations.

Captioning Style	**Story 1**	**Story 2**	**Story 3**	**Story 4**
Accuracy	F(5,89) = 0.963 p = 0.445	F(5,89) = 0.830 p = 0.532	F(5, 89) = 0.458 p = 0.807	F(5,89) = 1.085, p = 0.374
Time Required	F(5,89) = 0.675 p = .643	F(5,89) = 0.590 p = 0.708	F(5,89) = 0.704 p = 0.622	F(5,89) = 0.418 p = 0.835
Replays	F(5,89) = 0.693 p = 0.630	F(5,89) = 0.810 p = 0.546	F(5,89) = 2.147 p = 0.067	F(5,89) = 1.334 p = 0.257

Table 6: Analysis of variance, with order of presentation as independent variable

4.3.2 Did caption style influence accuracy?

The next statistic considers captioning style as the independent variable and accuracy as the dependent variable. Single-factor analysis was executed on average correct responses of each captioning style. Stories 1 and 2 were found to have statistically significant higher scores than Story 3. See Table 7.

Style	*Mean*	*n*	*Std. Dev*
Story 1	0.7105	95	0.26376
Story 2	0.6877	95	0.23349
Story 3	0.5579	95	0.28818
Story 4	0.6158	95	0.21481

Table 7: Accuracy as a function of captioning style.

Evidence from the ANOVA statistic demonstrated significant differences in accuracy scores between caption style conditions [F(3,376) = 7.27, p = .0001]. This was followed up with a post hoc Tukey analysis, as seen in Table 8.

		Story3	Story4	Story2	Story1
		0.5579	0.6158	0.6877	0.7105
Story3	0.5579				
Story4	0.6158	1.59			
Story2	0.6877	3.56	1.97		
Story1	0.7105	4.18	2.59	0.62	

Table 8: Post hoc analysis, listing p-values for pairwise tests.

Post hoc analysis employing Tukey simultaneous comparison t-values showed that the following pairs reflected significant differences:

- Story 1 vs. Story 3 had a score of 3.56
- Story 2 vs. Story 3 had a score of 4.18.

These scores were above the critical value for experiment-wise error rate of 3.18 (p = 0.01). Four other pairs did not show significant differences since they fell below the critical value:

- Story 3 vs. Story 4 (1.59)
- Story 2 vs. Story 4 (1.97)
- Story 1 vs. Story 4 (2.59)
- Story 1 vs. Story 2 (0.62)

Visit Appendix T to view the full analysis.

4.3.3 Did caption style influence time spent to answer questions and number of rewinds?

The review of average time required to complete each question and number of rewinds showed no significant differences between captioning styles. Mean and standard deviation scores for time required and number of rewinds can be found in Table 9 and Table 10 respectively. Note that the standard deviation is unusually high for Story 3 (non-captioned video) for time required. A possible explanation would be that the participants spent more time trying to figure out the answers because the information was lacking.

Style	*Mean*	*n*	*Std. Dev*
Story 1	22.2099	95	25.18688
Story 2	21.4410	95	22.81805
Story 3	24.9364	95	82.90679
Story 4	22.2041	95	27.40773

Table 9: Time required as a function of captioning style.

Style	*Mean*	*n*	*Std. Dev*
Story 1	0.1	95	0.52
Story 2	0.2	95	1.60
Story 3	0.1	95	0.54
Story 4	0.2	95	0.82

Table 10: Number of rewinds as a function of captioning style.

Single-factor ANOVA analysis showed no significant differences were found with time required [$F(3,376) = 0.102$, $p = 0.959$] and number of replays [$F(3,376) = 0.31$, $p = 0.816$]. The complete calculations can also be found in Appendix T.

4.4 Preference measures

The study involved two sets of preference measures. The first set was employed immediately after the user completed the performance measure that followed each video. It focused on the user's initial reactions with the captioning style. The second set of measures was taken during post-test after the participant had viewed all four captioning styles. This gave participants the opportunity to mentally review the four videos, to reflect upon all of the captioning styles, and to mentally compare all of them.

4.4.1 Responses immediately following each video

Immediately following the viewing of video, participants evaluated

- Ease of reading (Question 1)
- Ease of understanding (Questions 2 and 4)
- Confidence of captioning accuracy (Question 3)
- Preference of caption style (Questions 5 and 6)

Since Likert scales are non-parametric, a Kruskal-Wallis analysis was done on the responses using MegaStat (McGraw-Hill Glencoe, 2013). During the analysis the responses were automatically converted to rankings. Table 11 lists the rankings for the question 1, "The captioning was easy to read."

Story	*Ranking*
Story 2	3.00
Story 1	2.50
Story 4	2.00
Story 3	0.00

Table 11: Story ranking, based on responses to the question, "The captioning was easy to read."

Since the CSV file was missing data from one of the participants, the lone response had to be dropped and n set to 94 instead of 95.

Since the scores recognized significant differences among four caption styles at p = 0, a Mann-Whitney calculation was performed through an online calculator provided by Social Science Statistics (Social Science Statistics, 2013). Since there are four groups, we had six possible pairs so a multiplier of six was applied to each original p-value in which is known as Bonferroni adjustment. The calculation identified all pairs except for Story 1 vs. 4 as having statistically significant differences with the Bonferroni adjustment applied. Since there was no difference between Stories 1 and 4, it could imply that color-coded automatic captions, even with some errors, are as easy to read as the perfect manual captions. See Table 12 to review the Mann-Whitney results.

Story	*1*	*2*	*3*
2	z=-3.2341 p=0.00124 p'=0.00744		
3	z=8.8875 p=0 p'=0	z=10.5884 p=0 P'=0	
4	z=0.8712 p=0.19215 p'=1.15290	z=3.8225 p=0.00007 p'=0.00042	z=-8.2415 p=0 p'=0

Table 12: Mann-Whitney results for "The captioning was easy to read."

The question, "The captioning made it easy to understand the story," had similar median scores on all 95 responses with significant differences at p=0 confirmed by Kruskal-Wallis analysis (see Table 13).

Story	*Ranking*
Story 1	3.00
Story 2	3.00
Story 4	2.00
Story 3	0.00

Table 13: Story ranking, based on responses to the question, "The captioning made it easy to understand the story."

During the pair-wise analysis all but Story 1 vs. Story 4 were found to have significant differences. Pairs 1 vs. 2 and 2 vs. 4 had slightly less significance with Bonferroni adjustment at 0.03048 and 0.01734 respectively, but remained at $p < 0.05$ level.

The question, "From the captions, I feel that I fully understood the story," had a median score of 2.00 (neutral) on three of the four stories. Only the story one without

captions had a median score of 0.00 (strongly disagree) with significant differences identified by the Kruskal-Wallis analysis. See Table 14 for the story ranking schedule.

Story	*Ranking*
Story 1	2.00
Story 2	2.00
Story 4	2.00
Story 3	0.00

Table 14: Story ranking, based on responses to the question, "From the captions, I feel that I fully understood the story."

The Mann-Whitney post hoc confirmed that the three stories with captions had significant differences when compared to the non-captioned story. To review the results for both questions, see Table 15 and Table 16.

Captioning Style	*1*	*2*	*3*
2	z=-2.5712 p=0.00508 p'=0.03048		
3	z=8.5051 p=0 p'=0	z=10.0723 p=0 P'=0	
4	z=0.2335 p=0.40905 p'=2.4543	z=2.7585 p=0.00289 p'=0.01734	z=-8.1726 p=0 p'=0

Table 15: Mann-Whitney results for "The captioning made it easy to understand the story."

Captioning Style	*1*	*2*	*3*
2	z=-0.8192 p=0.20611 p'=1.23666		
3	z=7.889 p=0 p'=0	z=8.7439 p=0 P'=0	
4	z=-0.839 p=0.20045 p'=1.2027	z=-0.0435 p=0.48405 p'=2.9043	z=-8.1383 p=0 p'=0

Table 16: Mann-Whitney results for "From the captions, I feel that I fully understood the story."

Question 3, "I have confidence in the accuracy of the captioning," had the same median score for all three captioned stories at 2.00 (neutral) while Story 3 with no captions had 0.00 (strong disagree). See Table 17 for summary.

Story	*Ranking*
Story 1	2.00
Story 2	2.00
Story 4	2.00
Story 3	0.00

Table 17: Story ranking, based on responses to the question, "I have confidence in the accuracy of the captioning."

Because a Kruskal-Wallis analysis detected significant differences, a Mann-Whitney analysis was conducted. It found that the three captioned stories compared against the non-captioned Story 3 had statistically significant differences with Bonferroni adjustment, while all other pairs did not have any. It is extremely unlikely that the differences happened by some random chance. This brings the ASR videos (stories 1 and 2) in the same neutral category as Story 4 (with perfect, manual captions). For additional details see Table 18.

Captioning Style	*1*	*2*	*3*
2	z=-1.5198 p=0.06426 p'=0.38556		
3	z=7.3085 p=0 p'=0	z=8.1265 p=0 P'=0	
4	z=-0.0449 p=0.48405 p'=2.9043	z=1.3852 p=0.08226 p'=0.49356	z=-6.3706 p=0 p'=0

Table 18: Mann-Whitney results for "I have confidence in the accuracy of the captioning."

For caption style, it seems that participants preferred Story 2, the ASR version without color coding. It had a median score of 3.00 (neutral) while the other two captioned stories had 1.00 (disagree) and non-captioned had 0.00 (strongly disagree). See Table 19.

Story	*Ranking*
Story 2	3.00
Story 1 Story 4	1.00 1.00
Story 3	0.00

Table 19: Story ranking, based on responses to the question, "I like this style of captioning."

Kruskal-Wallis analysis confirmed significant differences and Mann-Whitney identified all except 1 vs. 4 as having significant differences. Similarly, "Compared to captions on TV, I prefer this caption style," had the following median scores from story 1 through 4: 1.00, 2.00, 0.00, and 1.00 with statistical significant differences recognized by the Kruskal-Wallis statistic. See Table 20 for the summary.

Story	*Ranking*
Story 2	2.00
Story 1	1.00
Story 4	1.00
Story 3	0.00

Table 20: Story ranking, based on responses to the question, "Compared to captions on TV, I preferred this style of captioning."

Mann-Whitney identified the same pairs from the last question as having differences. See Table 21 and Table 22 for additional details.

Captioning Style	*1*	*2*	*3*
2	z=-5.1252 p=0 p'=0		
3	z=-6.1107 p=0 p'=0	z=9.7649 p=0 P'=0	
4	z=0.0567 p=0.47608 p'=2.85648	z=4.7453 p=0 p'=0	z=-5.6503 p=0 p'=0

Table 21: Mann-Whitney results for "I like this style of captioning."

Captioning Style	*1*	*2*	*3*
2	z=-4.3046 p=0 p'=0		
3	z=4.4062 p=0 p'=0	z=7.7676 p=0 P'=0	
4	z=-0.9683 p=0.16602 p'=0.99612	z=3.0197 p=0.00126 p'=0.00756	z=-4.8746 p=0 p'=0

Table 22: Mann-Whitney results for "Compared to captions on TV, I prefer this caption style."

Also, after viewing the captioning style that involved highlighting (Story 1), participants responded to the statement, "The color coding was helpful to understanding the story." A strong majority (68 subjects) either disagreed or strongly disagreed with that statement. The visualization technique may have been distracting or unsatisfactory to the participants, so perhaps an additional study or revisit of the color coding may be in order. See Figure 8 for the histogram and Appendix U for complete calculations from the Kruskal-Wallis/Mann-Whitney analysis on captioning preference responses.

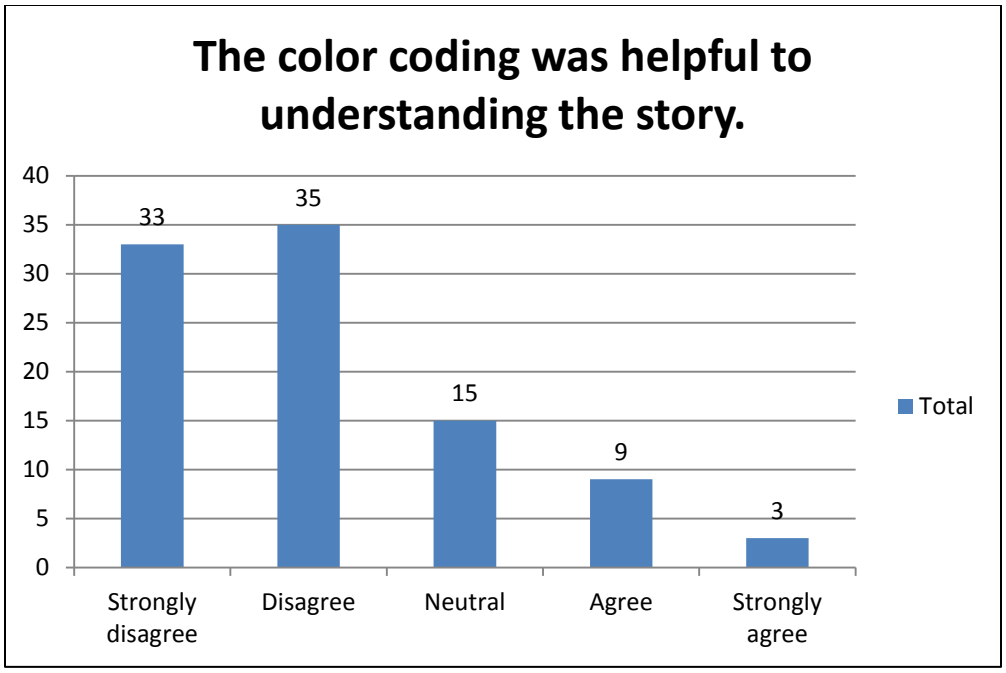


Figure 8: Histogram depicting the responses to whether color coding was helpful.

4.4.2 Responses after viewing all four captioning styles

When asked which video was the easiest to understand, none of the participants chose Story 3, the one without captions. Their selections came in the following (from high to low): Story 2 (44), Story 4 (32), and Story 1 (19). They also had an opportunity to select which video was the hardest to understand. An overwhelming majority, 80 of 95 participants chose the one without captions. Similarly, 85 users preferred videos with automatic captions while only 10 preferred to watch a video without captions.

Fifty-one participants recognized the logo indicating that the video was in ASR mode while 44 did not see it. When asked whether there is a better way to indicate that captions are created with ASR, 62 said "no," but 33 participants said "yes". When asked for suggestions on how to convey to users that ASR is being used for captioning, six participants suggested some kind of notification at the beginning of video so the user is alerted.

Additionally, five participants recommended using some text indicating that ASR is being employed to caption the video. Two users thought using parentheses just before each caption may be effective since it would be a continuous reminder that they are automatic captions. Someone also suggested using a different color altogether to distinguish automatic captions from manual captions.

If there was an ASR video containing errors, a strong majority (73) wanted color coding to depict potential errors while 22 did not feel it was necessary. Finally if there was a choice between watching a video with automatic captions and reading a transcript, 51 preferred automatic captions while 44 preferred the transcript. All the histograms from the post-test responses are provided in Appendix V.

4.5 Analysis

The results from the study provide solid evidence that captioning makes online videos more accessible to the deaf and hard-of-hearing. The participants were able to accurately answer the questions about videos that were automatically captioned significantly better than the non-captioned video. Interestingly, the accuracy for the video with perfect captions was lower than the accuracy for videos with ASR-generated captions. There are two possible explanations for this. Several participants complained that the final video were hard to read due to lack of punctuation and capitalization which interfered with clarity that the ASR videos had. Another possibility would be the order of presentation. The goal of placing the manually-captioned video last in the presentation order was to give it the benefit of any transfer of learning that may have occurred. However, the participants may have been fatigued from the lengthy survey and could have paid less attention to the video and perhaps hurried through the final set of questions.

In the post-test wrap-up, 80 of 95 participants felt that the video without captions was the hardest to understand. It was assumed that virtually every participant would state that the non-captioned video was the hardest to understand, so this was not expected. A possible explanation would be that there may have been some users who simply gave up on the non-captioned video and disregarded it altogether due to frustration. Also, some may have assumed the video was not captioned by mistake and instead focused on the captioned videos while answering questions. Furthermore, two of the four performance questions and available multiple-choice answers associated with the non-captioned video, the one about mushrooms, were answerable through common knowledge and that may have influenced the results:

1. What is the warning the author gives about mushrooms?
    a. Only royalty can eat them.
    b. Some varieties can be poisonous.
    c. They should only be eaten in pizzas.
    d. They must be grown in dark, damp places.
2. How are mushrooms different from green plants?
    a. They are very colorful.
    b. They are a part of people's diet.
    c. They appear around "fairy rings".
    d. They don't need the sun for growth.

Interestingly, when prompted for which story that was the easiest to understand, not a single participant selected the non-captioned video. In addition, 85 of 95 users would opt for a video with automatic captions over the one without. Based on the evidence presented,

the answer to question 1, “Does the presence of captions make Internet media more accessible?” is a resounding "yes."

These findings lend insight into question 2, “Are captions generated by ASR as effective as captions that were manually created by humans?” There were no significant differences in accuracy between ASR stories and the manually captioned story. Interestingly, the responses to ASR videos were scored slightly better than the perfect version. This was despite the fact that it was placed as the final video in presentation order and enjoyed any transfer-of-learning that may have taken place while the users watched the previous videos. Finally, the responses in preference and post-test responses exhibit positive remarks on automatic captions.

However, a confounding factor may have stemmed from the fact that the manually captioned video lacked punctuation and capitalization. So the answer to question 2 is a qualified "yes." Any additional study in the future should include more carefully formatted videos that align manual captions more closely to the quality presented as seen in automatic captions.

The responses to the preference question, “I have confidence in the accuracy of the captioning,” had a median score of 2.00 (neutral) on all captioned videos while the non-captioned video had a score of 0.00 (strongly disagree). The responses demonstrated that the participants had confidence in automatic captions. The automatic captioned videos shared same confidence level as manually captioned video. A majority (51 of 95) of participants were able to recognize the ASR logo placed on the upper-left video. About a third of participants (33) felt that there was a better way to indicate the video was showing automatic captions. Two popular suggestions mentioned having ASR notification at the

beginning of video so the viewer is aware and using text instead of a logo that emphasizes that the captions are generated by ASR. Research question 3, “Does automatic captions have an effect on person’s confidence in the captioning?" is a cautious "yes."

The fourth question considered whether having a visual indication of the estimated accuracy of the automatic captions affect a person’s ability to access the content. Although there was a significant difference between the accuracy in answering questions in the story without captions and the story with highlighted captions, there was no significant difference in accuracy among any of the captioning styles. Similarly, there was no significant difference in the amount of time taken to answer questions or in the number of replays of the video. So the answer to this question is “not in this study.”

The fifth and final question was whether having a visual indication of estimated accuracy of the automatic captions affects the viewer’s ability to access the content. There were mixed results. After the participants viewed the story about bears (color-coded ASR video) they responded to the statement, "The color coding was helpful to understand the story." Sixty-eight either disagreed or strongly disagreed with the statement. However, during post-test, when they were asked whether they wanted color coding to identify potential caption errors, 73 of 95 participants felt it would be useful while 22 did not see it as being necessary. This could mean that the color-coded video available to the study was not helpful, or suitable to their expectations. Further study will be necessary to help identify possible color-coded captioning options that would be useful to viewers.

5 Conclusion and Future Work

To summarize the research questions and their results:

1) Do captions make Internet media more accessible than no captions? YES
2) Are automated captions as effective as manual captions? MAYBE. Unclear due to the confounding factor.
3) Does knowing the captions are automatically generated have an effect on a person's confidence? CAUTIOUS YES
4) Does a visual indication of estimated accuracy improve performance compared to no visual indication? NO
5) Do users prefer a visual indication of estimated accuracy when the captions are generated via ASR? MAYBE. The participants responded positively when asked about this option. However, there was no significant difference in the preference of either visualization, or plain captions, when comparing the videos.

There are two contributions that stemmed from this work. The first may be more of a political, rather than academic, contribution, but Deaf advocacy groups need evidence that captioning is in the best interest of the Deaf and hard-of-hearing community. No previously-conducted controlled experiment has investigated this. Thus, one of the outcomes of this study is that it provides additional evidence that captioning makes Internet media more accessible to the Deaf community and will be useful to Deaf advocacy groups.

There may be an explanation to the weak support seen for color-coding the WER. The WER on the automatic captioned videos may have been low enough that the extra help provided by the visualization techniques may not have been necessary. Additionally, the

reader confidence in ASR-related videos was surprisingly rated higher than the manually captioned video. Additional studies should be conducted to determine whether color-coding and other visualization technologies could provide additional edge to help users understand ASR videos with higher WER.

There has been little movement in captioning Internet media. According to the media industry, the major reason for not captioning news articles is that the cost is prohibitive. Results from the study offer strong evidence that employing an alternative method would improve accessibility at a fraction of the cost of manual captioning. A news outlet could provide a web server that could automatically caption videos when requested by a user. Although the captioning would not be perfect, at least some of the information would become accessible, which is an improvement over the current situation. Since this study shows that there is benefit to adding less-than-perfect captions to videos, it may give additional evidence to Deaf advocacy groups to continue to lobby for change in the current FCC regulations.

Despite issues with speech recognition accuracy, the study provided strong evidence that deaf viewers can benefit from the videos with automatic captions. There are several possible ideas to help improve accuracy and train the speech recognizer to become more reliable. A central database and protocol of speaker speech profiles could be established so it is more effective and easier to train the speech recognizer engines. This could significantly help reduce redundancy and expand domain coverage. Another idea would be to employ Twitter or text-based Internet media to build speech recognizer's vocabulary and enhance matching confidence based on relevance. Finally, a community could be involved with maintaining and contributing to a central database of caption files. Speech recognition could perform the majority of captioning work, and then a user would

clean up the text and correct errors. The benefits of utilizing speech recognition to improve Internet accessibility for deaf users are endless. It would help narrow the accessibility gap that deaf Internet users experience daily and lead to leveling of the playing field. No one should be denied access to the abundance of information that the Internet has to offer. "Knowledge is power," is a well-known quote coined by Francis Bacon in 1597 in the *Meditationes Sacrae* (Bartlett, 1919) and it resonates well with the motivation behind this work.

6 References

Bain, K., Basson, S., Faisman, A., & Kanevsky, D. (2005). Accessibility, transcription, and access everywhere. *IBM Systems Journal*, 589-603.

Bartlett, J. (1919). *Familiar Quotations, 10th ed.* Retrieved August 20, 2013, from Bartleby.com: http://www.bartleby.com/100/139.39.html

Bartolome, A. I., & Cabrera, y. G. (2005). New Trends in Audiovisual Translation: The latest Challenging Modes. *Miscelanea: A Journal of English and American Studies*, 89-104.

Caption First. (2010). Retrieved October 22, 2010, from Caption First: http://www.captionfirst.com/

Carnegie Mellon University. (2012). *CMU Sphinx*. Retrieved 09 21, 2012, from http://cmusphinx.sourceforge.net/

Collins, C., Penn, G., & Carpendale, S. (2007). Visualization of Uncertainty in Lattices to Support Decision-Making. *Eurographics/ IEEE-VGTC Symposium on Visualization.* The Eurographics Association.

CuePrompter. (2012). *CuePrompter.com - The Online Teleprompter*. Retrieved September 21, 2012, from http://cueprompter.com/

Custom Captions. (2010). Retrieved October 20, 2010, from Custom Captions: http://www.customcaptions.com/#PriceComparisonChart

Downey, G. J. (2008). *Closed Captioning: Subtitling, Stenography, and the Digital Convergence of Text with Television.* Baltimore: The Johns Hopkins University Press.

Ellcessor, E. (2011). Captions On, Off, on TV, Online : Accessibility and Search Engine Optimization in Online Closed Captioning. *Television New Media, 13*(4), 329-352.

Fincher, S., & Tenenberg, J. (2005). Making sense of card sorting data. *Expert Systems, 22*(3), 89-93.

Flesch, R. (1948). A new Readability Yardstick. *Journal of Applied Psychology*, 221-233.

Florida, U. o. (2013). *Reading Florida Comprehensive Assessment Test (FCAT)*. Retrieved June 30, 2013, from FCAT Express: http://fcit.usf.edu/express/

Francis, P., & Stinson, M. (2010, November 20). *The C-Print Speech-to-Text System for Communication Access and Learning*. Retrieved November 20, 2010, from http://www.csun.edu/cod/conf/2003/proceedings/157.htm

Google. (2013). *YouTube JavaScript Player API Reference*. Retrieved July 15, 2013, from Google Developers: https://developers.google.com/youtube/js_api_reference

Hyatt, G. W. (2010, January 20). *Accessibility Hack #1: Captioning Online Video*. Retrieved from DotGov: http://dotgov.com/2010/01/accessibility-hack-1-captioning-online-video/

Information Access Division. (2013). *Information Access Division*. Retrieved July 15, 2013, from Information Technology Laboratory - Information Access Division: http://www.itl.nist.gov/iad/mig/tools/

Journalism.org. (2013). *YouTube & News Video Length.* Retrieved July 15, 2013, from Journalism.org: http://www.journalism.org/analysis_report/video_length

Juang, B. H., & Rabiner, L. R. (2005). Automatic Speech Recognition - A Brief History of the Technology Development. In *Elsevier Encyclopedia of Language and Linguistics, 2nd Edition.*

Kincaid, J. P., Fishburne, R. P., Robers, R. L., & Chissom, B. S. (1975). *Derivation of New Readabiity Formulas (Automated Reliability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel.* Naval Air Station. Memphis: Research Branch Report.

Lambourne, A., Hewitt, J., Lyon, C., & Warren, S. (2004). Speech-based real-time subtitling services. *International Journal of Speech Technology*, 269-279.

Leitch, D., & MacMillan, T. (2003, July). Retrieved from Liberated Learning: http://www.liberatedlearning.com/resources/pdf/RC_2003_Year_IV_Research_Report.pdf

Lowensohn, J. (2009, April 16). *YouTube launches CaptionTube: A caption editor*. Retrieved from http://news.cnet.com/8301-17939_109-10221213-2.html?tag=mncol;title

Martin, A., & Greenberg, C. (2008, June 17-18). 2008 NIST Speaker Recognition Evaluation Slides from Workshop Presentation. Montreal, Quebec, Canada.

McGraw-Hill Glencoe. (2013, August 1). *MegaStat*. Retrieved August 1, 2013, from McGraw-Hill Glencoe Online Learning Center: http://glencoe.mcgraw-hill.com/sites/0010126585/student_view0/megastat.html

Microsoft Corporation. (2012). *Microsoft Tellme*. Retrieved September 21, 2012, from http://www.microsoft.com/en-us/tellme/

Microsoft Corporation. (2013, July 15). *Analysis ToolPak*. Retrieved July 15, 2013, from Microsoft Office: http://office.microsoft.com/en-us/excel-help/load-the-analysis-toolpak-HP010021569.aspx

MultimediaWiki. (2013). *SubStation Alpha*. Retrieved July 15, 2013, from MultimediaWiki: http://wiki.multimedia.cx/index.php?title=SubStation_Alpha

Munteanu, C., Baecker, R., Penn, G., Toms, E., & James, D. (2006). The Effect of Speech Recognition Accuracy Rates on the Usefulness and Usability of Webcast Archives. *CHI Proceedings - Visualization and Search*, (pp. 493 - 502). Montréal.

National Association of the Deaf. (2010, Fall). *NADmag*. Silver Spring, MD: National Association of the Deaf.

National Institute of Standards and Technology. (2011). *Multimodal Information Group - Speaker Recognition*. Retrieved March 2, 2011, from National Institute of Standards and Technology: http://nist.gov/itl/iad/mig/sre.cfm

Nitchie, B. E. (2006). *Lip-Reading: Principles and Practice: A Hand-Book for Teachers and for Self Instruction.* London: Adamant Media Corporation.

Padden, C. A., & Humphries, T. L. (1990). *Deaf in America: Voices from a Culture.* Boston: Harvard University Press.

Picone, J. (1990, July). Continuous Speech Recognition Using Hidden Markov Models. *IEEE ASSP Magazine*, pp. 26-41.

Public Schools of North Carolina. (2013). *End of Grade Tests - Grade 8 - Reading Comprehension*. Retrieved June 30, 2013, from http://www.ncpublicschools.org/docs/accountability/testing/releasedforms/grade8readingreleased.pdf

Social Science Statistics. (2013, July 31). *Mann-Whitney U-value Calculator*. Retrieved July 31, 2013, from Social Science Statistics: http://www.socscistatistics.com/tests/mannwhitney/

Sutter, J. D. (2010, February 9). *An engineer's quest to caption the Web*. Retrieved February 21, 2011, from CNN.com: http://articles.cnn.com/2010-02-09/tech/deaf.internet.captions_1_captioning-deaf-people-sign-language?_s=PM:TECH

U.S. Congress. (2010). *S. 3304: Twenty-First Century Communications and Video Accessibility Act of 2010*. Retrieved from http://www.govtrack.us/congress/bill.xpd?bill=s111-3304

Vemuri, S., DeCamp, P., Bender, W., & Schmandt, C. (2004). Improving Speech Playback Using Time-Compression and Speech Recognition. *CHI Proceedings*, (pp. 295 - 302). Vienna.

Vertanen, K., & Kristensson, P. O. (2008). On the Benefits of Confidence Visualization in Speech Recognition. *CHI Proceedings*, (pp. 1497 - 1500). Florence.

VirtualDub. (2013). *VirtualDub*. Retrieved July 15, 2013, from VirtualDub: http://www.virtualdub.org/

Wald, M. (2006). An exploration of the potential of Automatic Speech Recognition to assist and enable receptive communication in higher education. *Research in Learning Technology*, 9-20.

Appendix A – Selected Glossary and Abbreviations

ASL: American Sign Language

ASR: Automatic Speech Recognition

CART: Communication Access Real-time Translation

CC: Closed Captioning

deaf: deaf (with a lowercase "d") refers to the audiological condition of not hearing.

Deaf: Deaf (with a capital "D") refers to a particular group of deaf people who share a common language, the American Sign Language.

NCI: National Captioning Institute

SDH: Subtitles for the deaf and hard-of-hearing

WER: Word Error Rate

WPM: Word per Minute

Appendix B – IRB Documentation for the Interviews (Study 1)

DEPAUL
UNIVERSITY

Institutional Review Board
1 East Jackson Boulevard
Chicago, Illinois 60604-2201
312-362-7593
Fax: 312-362-7574

Research Involving Human Subjects

NOTICE OF INSTITUTIONAL REVIEW BOARD ACTION

To: Brent Shiver, graduate student, College of Computing and Digital Media
Rosalee Wolfe, Ph.D., faculty advisor, College of Computing and Digital Media

Date: June 14, 2011

Re: Research Protocol #JR052311NUR
"Identifying best practices for deaf accessibility of web-based multimedia"

Please review the following important information about the review of your proposed research activity.

Review Details
☒ Original Review
☐ Amendment
☐ Unanticipated Problem Report

☒ Exempt Review, under 45 CFR 46.101(b)

Your research project meets the criteria for an exemption as benign study that will investigate the areas of greatest need on the Internet to improve deaf people's accessibility to web multimedia.

Category of Review: 2

(2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:
(i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

Approval Details
☐ Approved ☒ Approved (Previous contingencies have been resolved.)

Approval Date: June 14, 2011

Number of approved participants: 20
Funding Agency: Unfunded

Reminders
- Under DePaul's current institutional policy governing human research, research projects that meet the criteria for an exemption determination receive administrative review. Once projects are determined to be exempt, the researcher is free to begin the work and is not required to submit an annual update (continuing review). As your project has been determined to be exempt, your primary obligation moving forward is to resubmit your research materials for review and classification/approval, before changes are implemented in the research, if you propose

substantive changes to the project. Substantive changes would include changes in the design or focus of the research project, revisions to the consent/information sheet for participants, addition of new measures or instruments, and any change to the research that might alter the exemption status (either add additional exemption categories or make the research no longer eligible for an exemption determination).

- **Once the project is complete, you should submit a closure report to the IRB.**

The Office of Research Protections would like to thank you for your efforts and cooperation and wishes you the best of luck on your research. If you have any questions, please contact me by telephone at (312) 362-7497 or by email at kwalsh31@depaul.edu.

For the Board,

Kari Walsh
Assistant Director, Office of Research Protections
Academic Affairs, DePaul University
1 East Jackson Blvd.
Chicago, IL 60604

Office Location: 14 E. Jackson Blvd., Suite 1030

Exempt Application for Review of Human Subjects Research

Principal Investigator

Name:		**Title:**			
Dept.:		**Phone:**		**Fax:**	
Email:					

Study Personnel Information

Name:		**Title:**			
Dept.:		**Phone:**		**Fax:**	
Email:					
Study Status:	☐ co-principal investigator ☐ contact person at collaborating institution		☒ faculty sponsor ☐ other, please specify:		

Project Information

Title of Research Activity:			
Research Type:	☒ Dissertation ☐ Master's Thesis ☐ Bachelor's Paper	☐ Unfunded faculty/staff research ☐ Funded faculty/staff research ☐ Other, please specify:	
Proposed Starting Date:		**Proposed Ending Date:**	
Will the study take place at or in collaboration with another institution?		☐ **YES** ☒ **NO**	
If yes, please provide the name of the institution, and provide a copy of the most recent IRB approval memo. (If the collaborator has no IRB, you may include a letter of collaboration in lieu of the approval memo.)			

Specific Objectives

(Character Limit=1500)

Please provide the following specific information about your study:

- purpose/aim of the project, number of participants
- type of data to be collected and method of collection.

(Character Limit=1500)

Please Answer All the Following Questions Regarding this Research Activity:

1. Where are the participants of this research activity or the archival, identifiable data located?

(Character Limit=1000)

2. If you will have direct contact or intervention with the human participants, please indicate/describe the following:

- **how many participants you expect the project to include;**
- **your method of recruitment; and**
- **whether vulnerable populations will be included (e.g. children, prisoners, pregnant women).**

(Character Limit=1000)

3. If your research activity will utilize archival data, please indicate what type of data will be involved (e.g. medical records, survey responses).

(Character Limit=1000)

4. If you will utilize archival data, will you do any of the following:

- **Retrieve the data from a privately-available source?** ☐ **YES** ☐ **NO**
- **View the data at the collaborator's facility & extract only deidentified/uncoded data?** ☐ **YES** ☐ **NO**
- **Receive data from a collaborator with identifiers/codes?** ☐ **YES** ☐ **NO**

5. With respect to the data you described in #3 and #4 above,

- **Does the data already exist?** ☐ **YES** ☐ **NO**
- **Are they being collected for the purpose of this study?** ☐ **YES** ☐ **NO**
- **Or a combination of (a) and (b)?** ☐ **YES** ☐ **NO**

If "Yes" to (c), please describe below:

(Character Limit=700)

6. An IRB must review and approve the use of existing identifiable data and those that are coded and may be linked in any way to an individual. The use of data that are anonymous (i.e., are not coded or linked in any way to an individual) may be considered for exemption from IRB review and approval.

- **Are the data you expect to collect, receive, and send anonymous?** ☒
YES ☐ **NO**
- **If so, do you intend to make the data identifiable at any point by combining it with other data or through another means?** ☐
YES ☒ **NO**

7. Documentation of Human Subjects Training

Background: All investigators and research assistants working with human research participants or analyzing data must complete human subjects training before data collection begins. Under new federal regulations, documentation of training must be placed on file with the Research Protections Office. The IRB recommends the following online training module: http://www.citiprogram.org.

Please include below names and training status for all research personnel who will be involved in recruiting, consenting, or collecting data from participants and/or those personnel who will analyze study data.

Type name of principle investigator
☒ Documentation attached
☐ Training/documentation pending
☐ Documentation on file for project: Type project title here

Type name of co-investigator or research assistant
☐ Documentation attached
☐ Training/documentation pending
☐ Documentation on file for project: Type project title here

***If there are more co-investigators, please attach further sheets with their name and status of their human subjects training documentation.**

Please attach separate sheet if there is anything else you wish to add or any answer you wish to explain further.

Investigator's Assurance

I certify that the information provided in this application is complete and correct. I understand that as Principal Investigator, I have ultimate responsibility for the protection of the rights and

welfare of human participants, conduct of the study and the ethical performance of the project. I agree to comply with all IRB policies and procedures, as well as with all applicable federal, state and local laws regarding the protection of human participants in research, including, but not limited to, the following:

- The project will be performed by qualified personnel according to the DPU IRB certified protocol,
- No changes will be made in the protocol or consent form until approved by the DPU IRB,
- Legally effective informed consent will be obtained from human participants if applicable, and

Adverse events will be reported to the DPU IRB in a timely manner.**I further certify that the proposed research is not currently underway (except for those protocols of research previously approved and currently seeking renewal) and will not begin until approval has been obtained.**

Principal Investigator's Signature:_____ Date:_____

FACULTY SPONSOR'S ASSURANCE FOR STUDENT OR GUEST INVESTIGATORS

By my signature as sponsor on this research application, I certify that the student or guest investigator is knowledgeable about the regulations and policies governing research with human participants and has sufficient training and experience to conduct this particular study in accord with the approved protocol. In addition,

- I agree to meet with the investigator on a regular basis to monitor study progress,
- Should problems arise during the course of the study, I agree to be available, personally, to supervise the investigator in solving them,
- I insure that the investigator will promptly report significant or untoward adverse effects to the DPU IRB in a timely manner,

If I will be unavailable, as when on sabbatical leave or vacation, I will arrange for an alternate faculty sponsor to assume responsibility during my absence and I will advise the DPU IRB by letter of such arrangements. I further certify that the proposed research is not currently underway and will not begin until approval has been obtained.

Faculty Sponsor's Signature:_____ Date:_____

*The faculty sponsor must be a member of the DPU faculty. The faculty member is considered the responsible party for legal and ethical performance of the project.

Please submit the following supporting materials along with this form, as applicable to your project. Please also indicate which of these materials have been included by checking the appropriate boxes:

☒Evidence of human subjects training for all study personnel(link for training available at http://research.depaul.edu)
☒Exempt info sheet (recommended; to be used in place of consent)
☒ Surveys, questionnaires, interview questions/guides

Incomplete application packets or applications that have had questions deleted may result in review delays.

Appendix C – Emailed Invitation (Study 1)

Email invitation

I am looking for deaf professionals who use the Internet on a daily basis to participate in a study to identify ways to make the Internet more deaf friendly. The interview will take about 30 minutes of your time. First a background questionnaire will need to be completed. It will ask about your deafness and Internet usage. Then I will ask questions about your experience with Internet and its accessibility. Email me if you are interested.

Thanks in advance,

Brent Shiver
DePaul University
bshiver@cdm.depaul.edu

DePaul IRB #JR052311NUR

Appendix D – Demographic Questionnaire (Study 1)

Background Questionnaire

Age [ ] 18-29 [ ] 30-39 [ ] 40+

Gender [ ] male [ ] female

Occupation

Degree of hearing loss

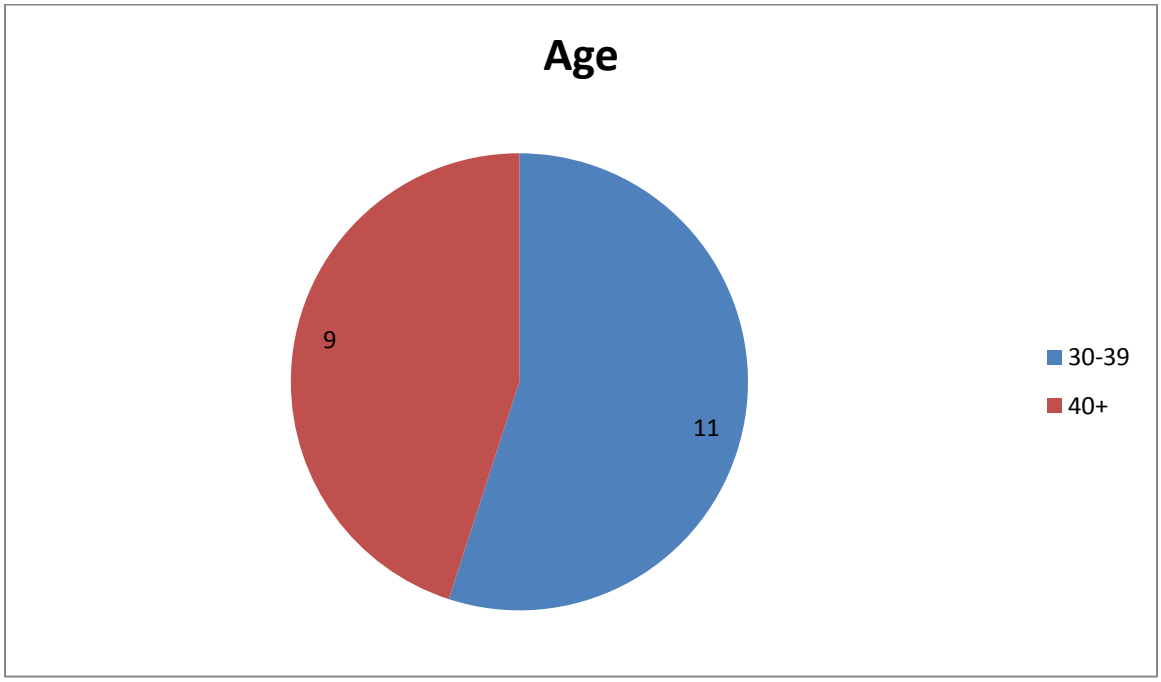
Since when?

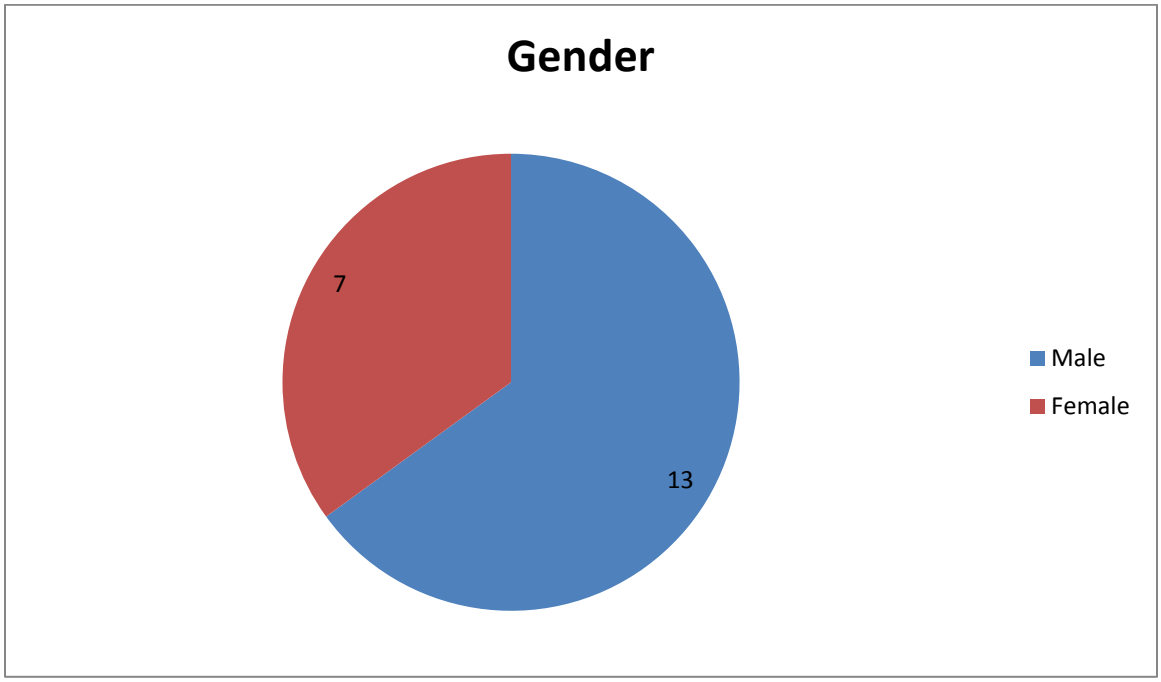
How often do you use the Internet to

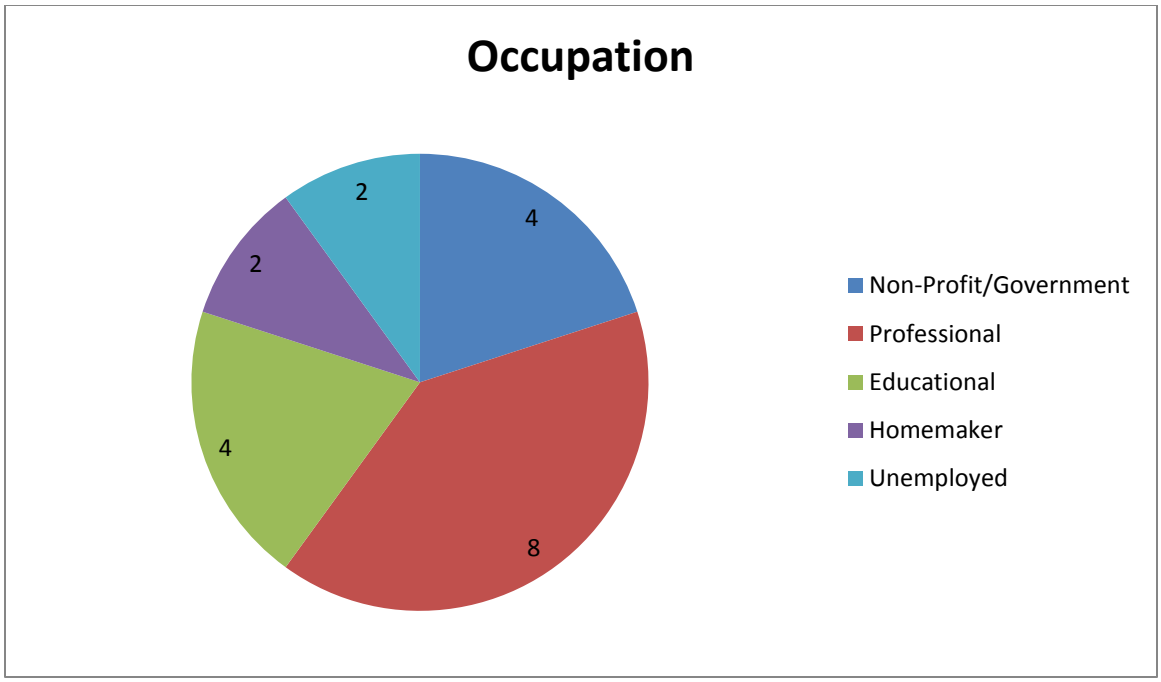
	Rarely	Once a week	Once a day	Many times a day	How many hours?
Read email					
Read news articles					
Online auction					
Online shopping					
YouTube videos					
Social media (Facebook, LinkedIn, Twitter, etc)					
Training/remote class					

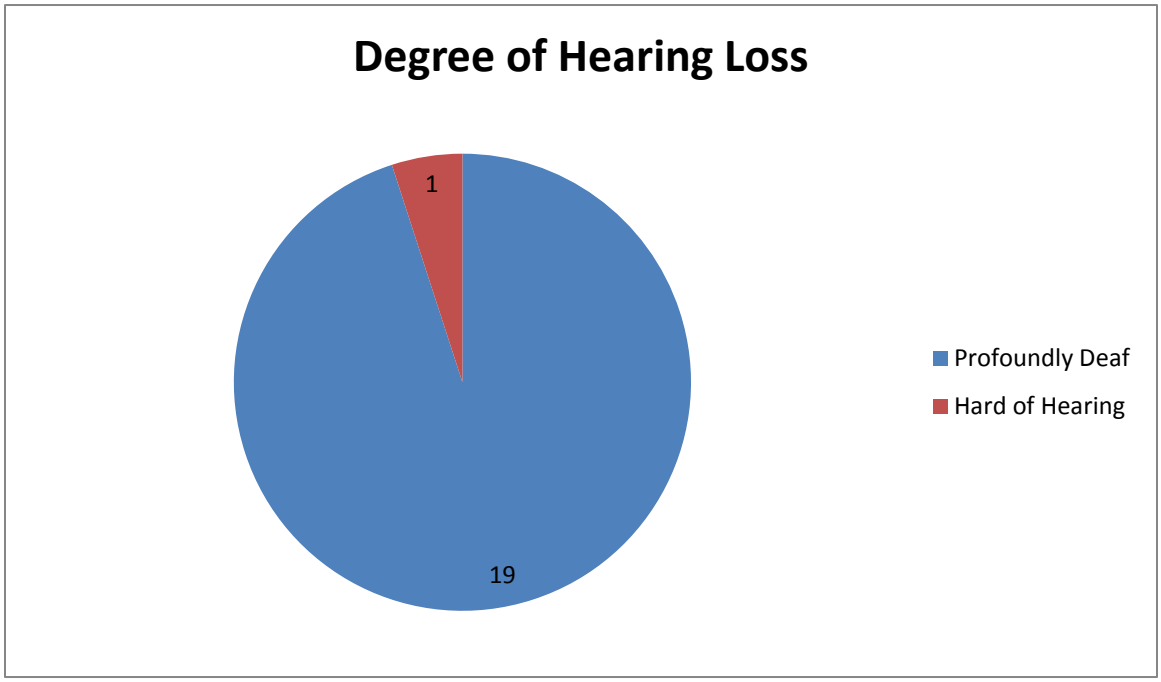
Are there any other Internet activities that you do every day? If so, what are they?

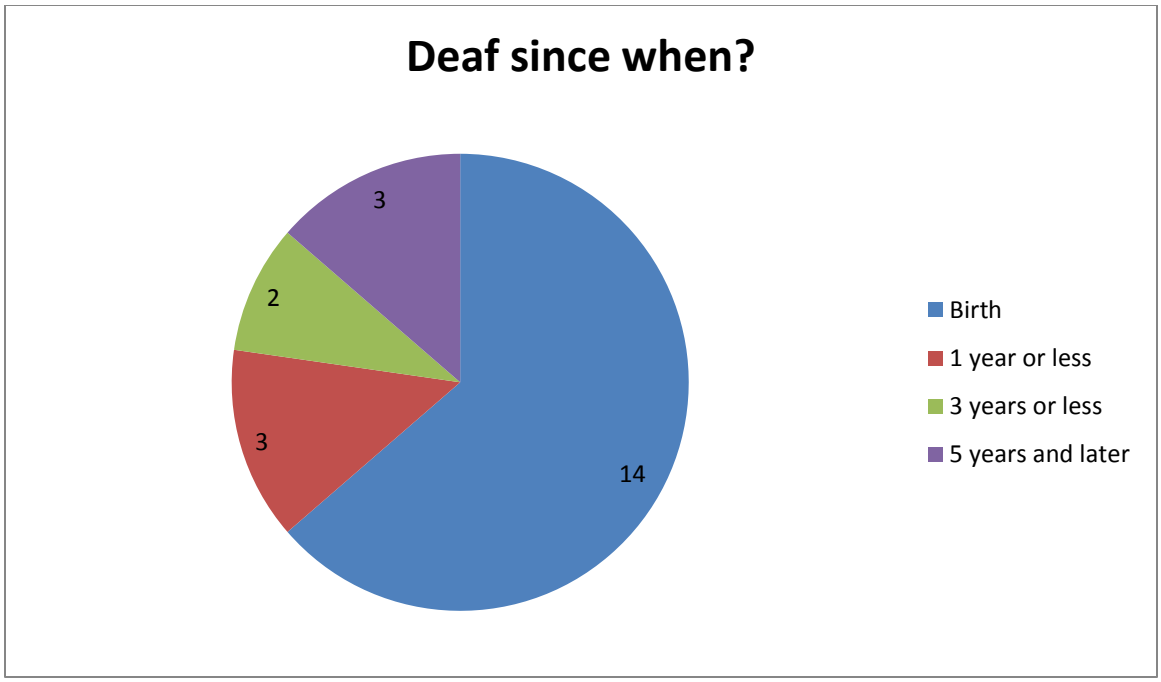
Appendix E – Demographic Background Charts (Study 1)

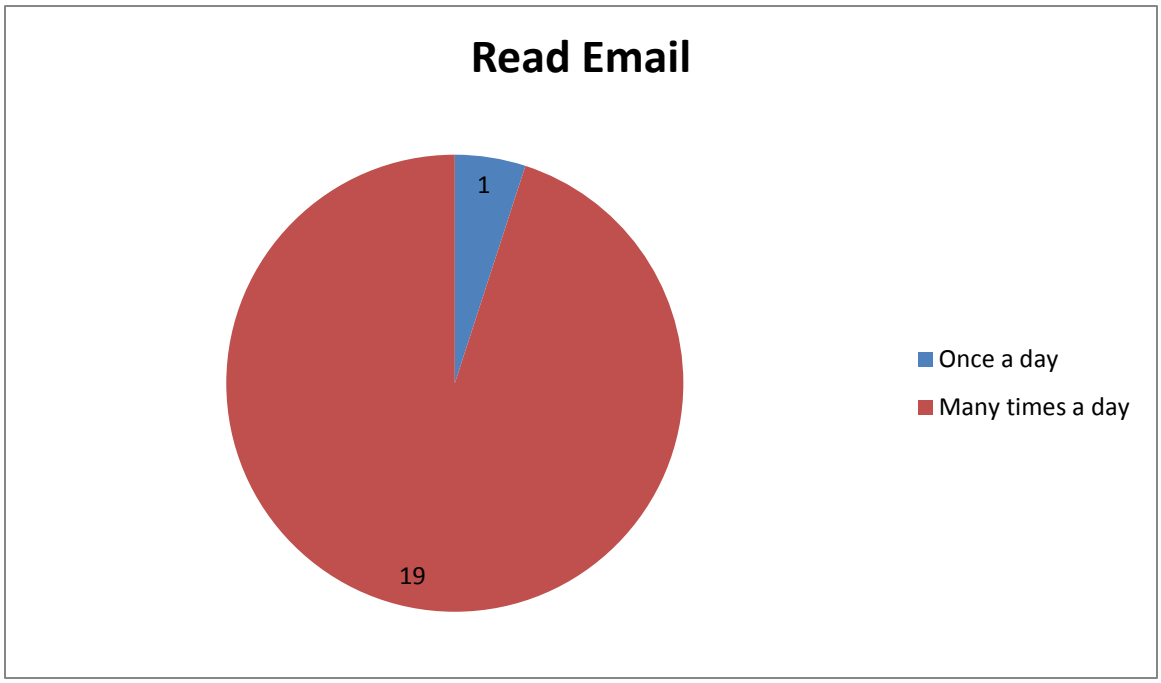


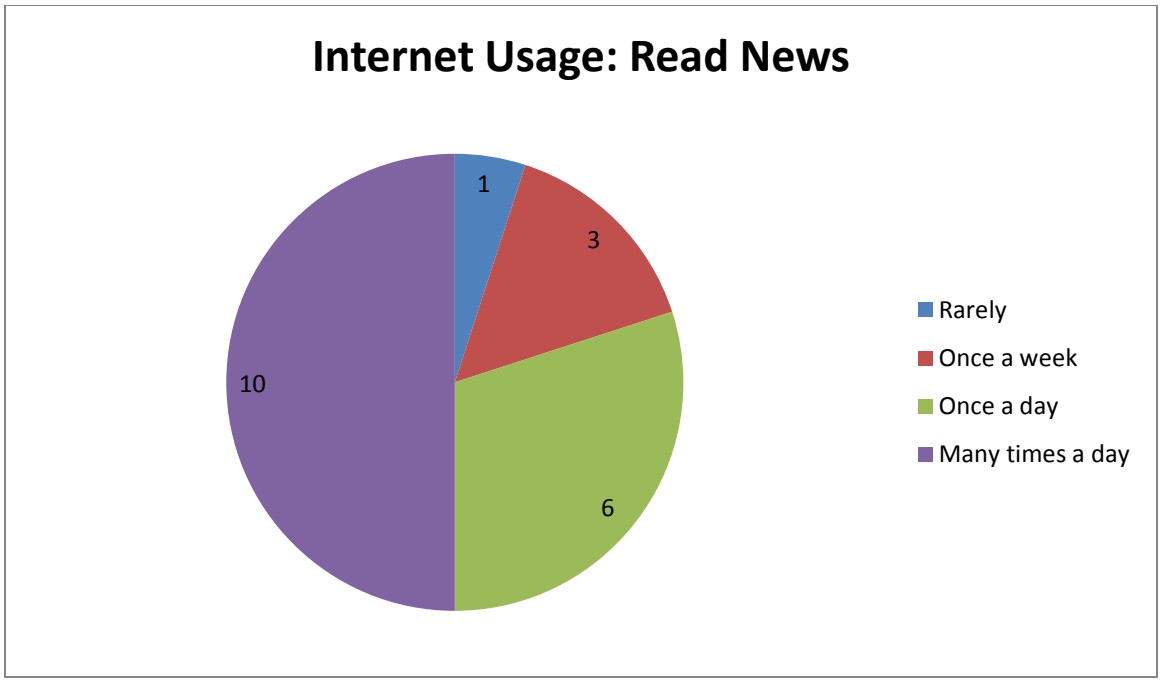


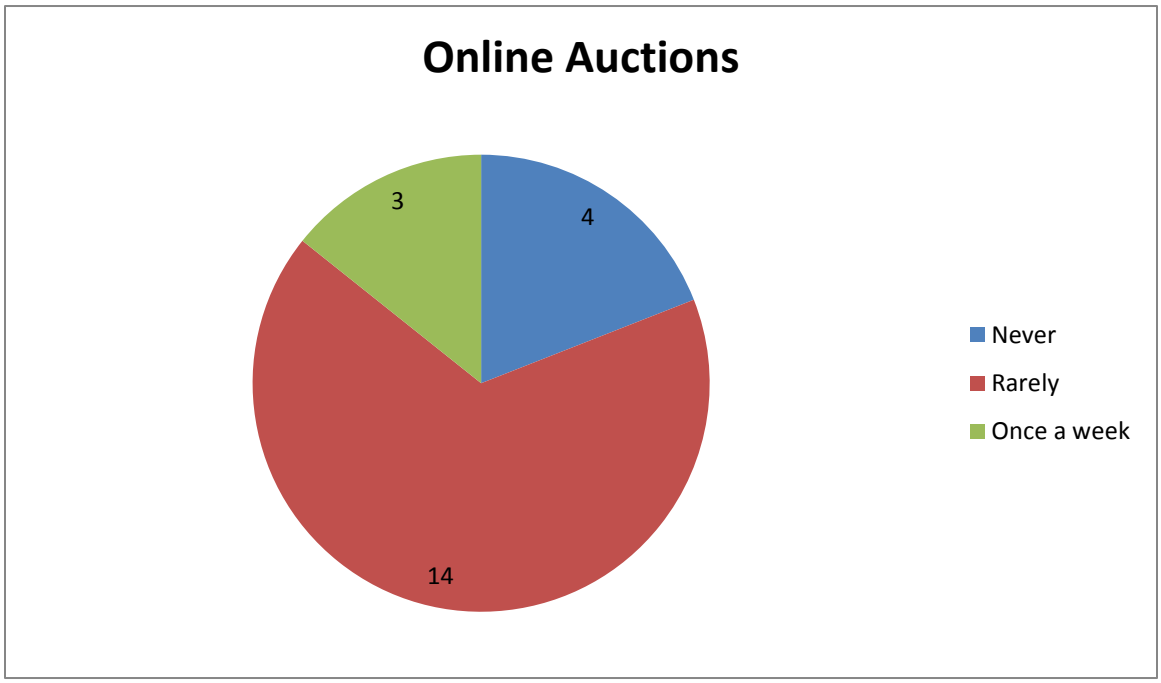

Occupation
4
8
4
2
2
Non-Profit/Government
Professional
Educational
Homemaker
Unemployed

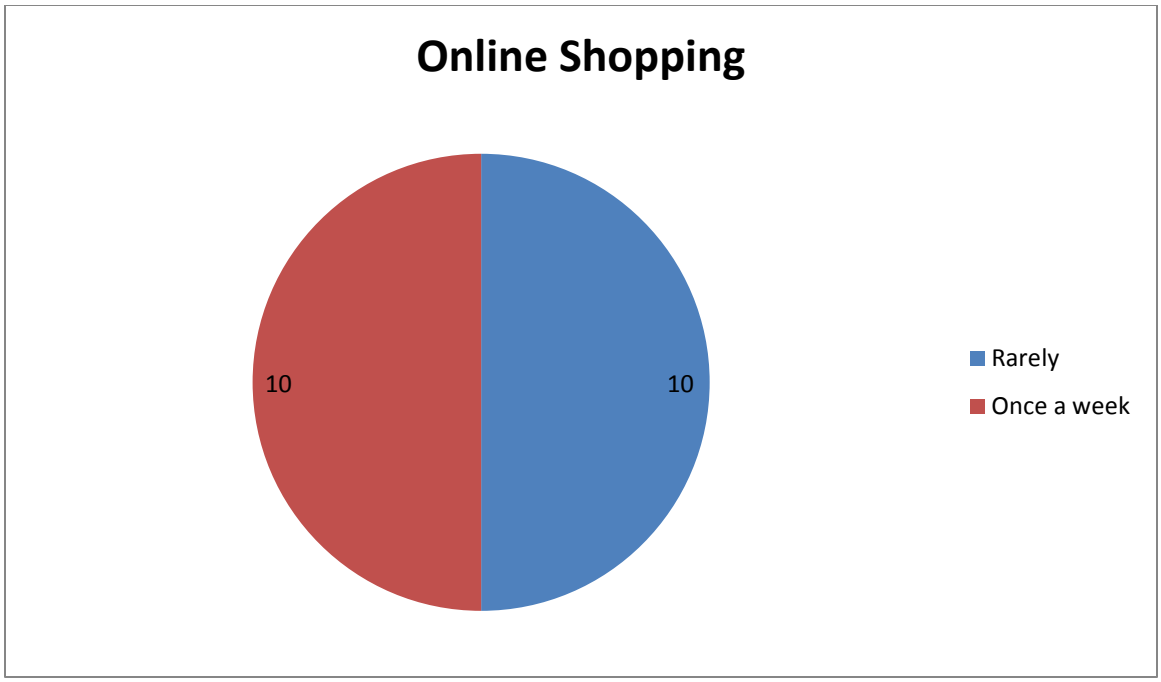

Degree of Hearing Loss
19
1
Profoundly Deaf
Hard of Hearing

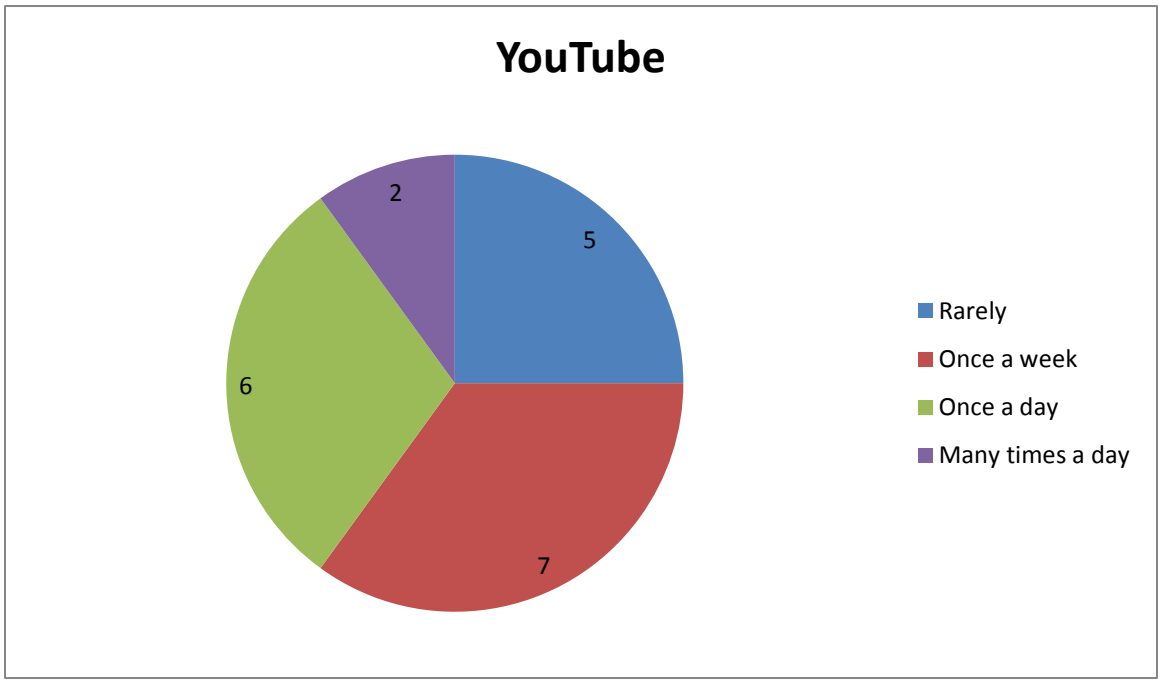

Deaf since when?
3
2
3
14
Birth
1 year or less
3 years or less
5 years and later

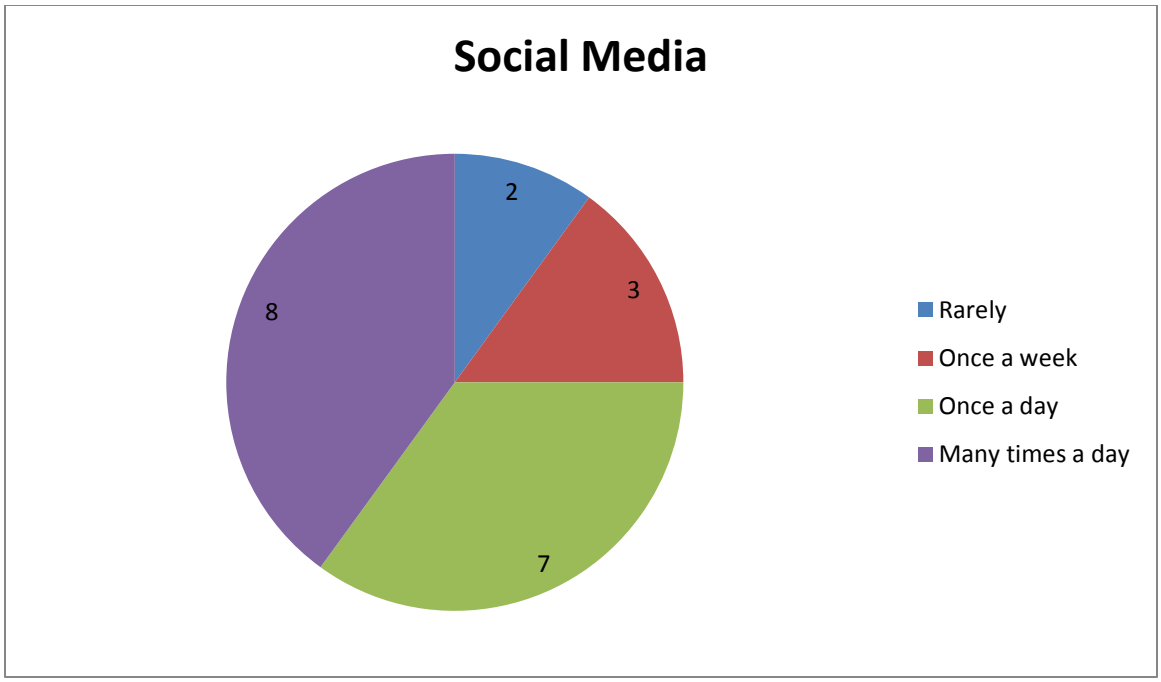

Read Email
1
19
Once a day
Many times a day

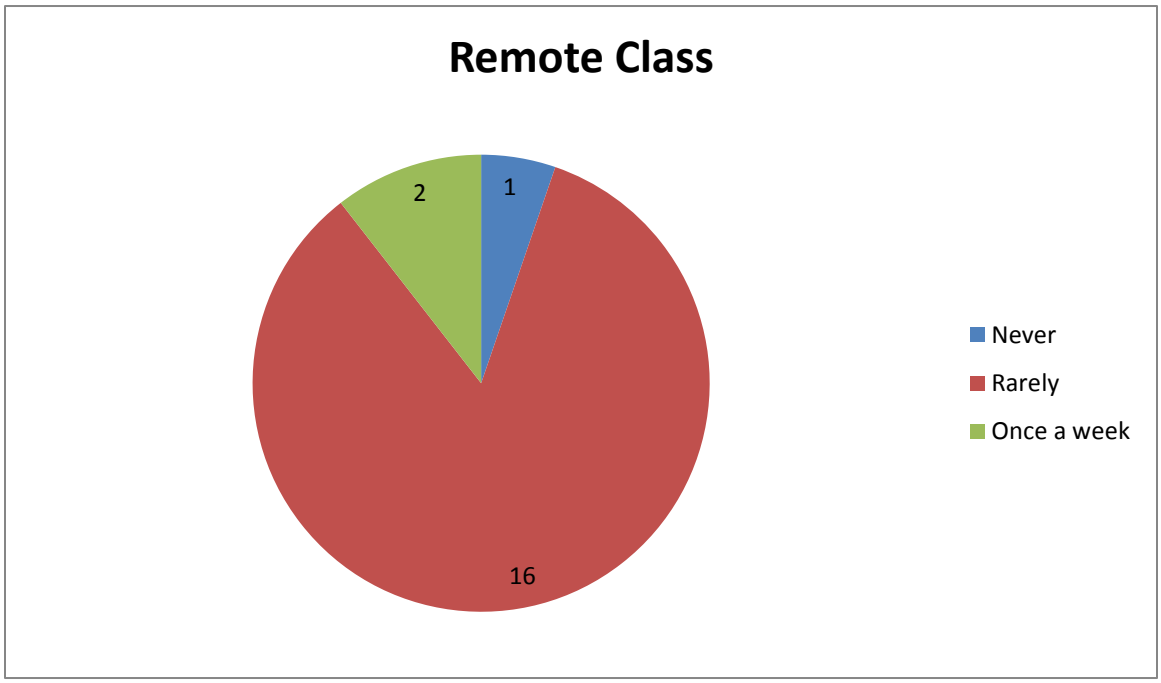

Internet Usage: Read News
1
3
6
10
Rarely
Once a week
Once a day
Many times a day


Online Auctions
4
14
3
Never
Rarely
Once a week


Online Shopping
10
10
Rarely
Once a week


YouTube
2
5
6
7
Rarely
Once a week
Once a day
Many times a day


Social Media
2
3
7
8
Rarely
Once a week
Once a day
Many times a day


Remote Class
1
2
16
Never
Rarely
Once a week

Appendix F – Information Sheet (Study 1)

INFORMATION SHEET FOR PARTICIPATION IN RESEARCH STUDY

Identifying best practices for Deaf accessibility of Web-based multimedia

You are being asked to participate in a research study being conducted by Brent Shiver at DePaul University. We are asking you to participate because we are trying to identify areas of need for better Deaf accessibility of web-based multimedia and learn about current coping mechanisms for overcoming accessibility barriers.

This study will take about 30 minutes of your time. If you agree to be in this study, you will be asked to fill a background questionnaire and complete the interview. The questionnaire will ask about your deafness and Internet usage. The main interview will include questions about your experience with Internet and its accessibility.

You can choose not to participate. There will be no negative consequences if you decide not to participate or change your mind later. If you change your mind, all you need to do is leave. You can do this without any negative consequences.

There are no benefits and there are no foreseeable risks in participating in the study. I may record the session but it will be retained only long enough to collect aggregate data and then destroyed.

If you have questions about this study, please contact Brent at 312-957-8950 or via email bshiver@cs.depaul.edu. You can also contact his advisor, Dr. Wolfe, by phone at 312.362.6248 or through email wolfe@cs.depaul.edu. If you have questions about your rights as a research subject, you may contact Susan Loess-Perez, DePaul University's Director of Research Protections at 312-362-7593 or by email at sloesspe@depaul.edu.

You may keep this information for your records.

Appendix G – Interview Questions (Study 1)

This is just to remind you that this session will be recorded. The recording will only be retained long enough for me to check my notes and collect aggregate data, and then the recording will be destroyed. If you agree to this, then let us begin.

1. When you read news articles on the Internet, do you ever watch the videos? Why or why not?

2. Have you watched YouTube?

3. Have you ever found yourself needing the information on a video?

4. Can you describe the type of video it was? What were the circumstances? Why was the information necessary?

5. What do you do to obtain the information or contents from videos?

6. How often are you frustrated about inaccessibility when you use the Web?

7. Describe top three frustrations that you've experienced. What happened?

8. Are you familiar with Google automatic captions? If yes, please tell me about your experience.

9. Have you used any other Automated Speech Recognition technology?
    - If so, what was it?
    - How did it work for you? Please tell me what worked and what didn't.

Now I need to review a couple of items of terminology with you. **Captioning** is the process of displaying text on a television, video screen or other visual display. Captions typically show a transcription of the audio portion of a program as it occurs. A **transcript** is a document containing a complete written or printed version of content originally presented as a video or recording.

10. Which approach do you prefer? Advantages and disadvantages of each?

11. Are there any situations where you prefer captions over transcripts and vice versa?

12. We are currently investigating technologies that may improve accessibility. Which situations that you mentioned earlier do you feel this would benefit the most?

13. Do you have any advice or suggestions in regards to improving accessibility on the Internet?

Thank you for your time.

Appendix H – IRB Documentation for the Online Survey (Study 2)

Form: Application for Claim of Exemption Version 12/17/2012	DePaul University Office of Research Services Institutional Review Board 1 East Jackson Blvd. Chicago, Illinois 60604-2201 Email: orp@depaul.edu Phone: (312) 362-7593 Web: http://research.depaul.edu/

Form completion or revision date: May 10, 2013

Which exemption category (ies) applies to your research:

- ☐ (1) Research conducted in established or commonly accepted educational settings, involving normal educational practices, such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.
- ☒ (2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:
(i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.
- ☐ (3) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior that is not exempt under paragraph (b)(2) of this section, if:
(i) the human subjects are elected or appointed public officials or candidates for public office; or (ii) federal statute(s) require(s) without exception that the confidentiality of the personally identifiable information will be maintained throughout the research and thereafter.
- ☐ (4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.
- ☐ (5) Research and demonstration projects which are conducted by or subject to the approval of department or agency heads, and which are designed to study, evaluate, or otherwise examine:
(i) Public benefit or service programs; (ii) procedures for obtaining benefits or services under those programs; (iii) possible changes in or alternatives to those programs or procedures; or (iv) possible changes in methods or levels of payment for benefits or services under those programs
- ☐ (6) Taste and food quality evaluation and consumer acceptance studies, (i) if wholesome foods without additives are consumed or (ii) if a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or environmental contaminant at or below the level found to be safe, by the Food and Drug Administration or approved by the Environmental Protection Agency or the Food Safety and Inspection Service of the U.S. Department of Agriculture.

Principal Investigator:

Name (Last): Shiver		(First): Brent	Degree: Ph.D.
Status: Grad Student	Title: Student	Dept: School of Computing	College/School: Computing and Digi
Phone: 312-957-8950	DePaul Email: bshiver@cs.depaul.edu		Alternate Email: bshiver@gmail.com
Human Subjects Training	☒ Completed	☐ Pending	

Faculty Sponsor: (If you are a student, in a staff position functioning as a training position, or a fellow, you must have a faculty sponsor when you are the Principal Investigator.)

Name (Last): Wolfe		(First): Rosalee	Degree: Ph.D.
Status: Faculty	Title: Professor	Dept: School of Computing	College/School: Computing and Digi

Phone: (312)362-6248	DePaul Email: wolfe@cs.depaul.edu	Alternate Email:
Human Subjects Training	☒ Completed	☐ Pending

* Additional Key research personnel should be listed on *Form: Co-Investigators and Key Research Personnel*

I. Project Information:

1. Title: Improving Deaf Accessibility to Web-based Multimedia

2. A. Proposed starting date: 3/15/2013 B. Proposed end date: 3/15/2014

3. Research Type: ☐ Faculty/Staff Research ☒ Student Research (Check one below)

- ☐ Undergraduate project or paper
- ☐ Master's Thesis or Graduate paper
- ☒ Dissertation
- ☐ Other: Specify:

II. Funding:

1. Is this research funded by an internal or external source?

☐ No ☒ Yes or Pending If Yes or Pending, provide information about the funding source in #2. below.)

If no, and there are research expenses, such as payment to subjects, explain how the expenses will be covered.

I have support from DePaul CDM PhD Summer Funding and Competitive Grants program to cover the the research expenses for this study.

2. Type of Funding

Extramural:

☐ Federal ☐ Foundation ☐ State

☐ Subcontract from non-DePaul Agency or Orgnization ☐ Other: Specify:

***If the research is federally funded or pending federal funding, attach a copy of the funding proposal submitted to the funding agency.**

Intramural:

☐ University Research Council ☐ Departmental ☒ Other: Specify: DePaul Summer Funding and Competitive Grants

3. Complete this section for all sources of anticipated funding, including when funding is pending. If you have more than one source of funding attach an additional sheet providing the information for each funding source.

A. Name of Funding Source:

B. Name of PI for funding:

C. Funding Agency Grant/Contract number, if available:

D. Grant/Contract/Project Title, if it differs from IRB title:

III. Conflict of Interest (COI):

Federal guidelines emphasize the importance of assuring there are no potential conflicts of interest in research projects that could affect the rights and welfare of human subjects. All investigators involved in the design, conduct, or reporting of research are required to disclose real, apparent, or potential significant financial conflicts of interest that could impact the conduct of the research or the integrity of the research data. For the purposes of Federal PHS policy and DePaul Policy, "Investigator" is defined as any person responsible for the design, conduct, or reporting of the research. The term investigator includes the Principal Investigator, Faculty Sponsor, Co-investigators, and other key research personnel. When determining whether there is a Significant Financial Interest in the research for these personnel, their spouses and dependent children should also be considered.

The following are the current DePaul policies governing conflicts of interest:

1. General Conflict of Interest Policy: http://policies.depaul.edu/policy/policy.aspx?pid=23
2. Conflict of Interest in Externally Sponsored Projects: http://policies.depaul.edu/policy/policy.aspx?pid=253
3. Conflict of Interest in Public Health Service (PHS) Funded Research: http://policies.depaul.edu/policy/policy.aspx?pid=302

The three DePaul policies governing conflict of interest differ slightly in their reporting requirements. Please review the three policies before answering the following questions.

Please note: Significant Financial Conflicts of Interest require review by the Conflict of Interest Committee before IRB review and approval. Final IRB approval cannot be granted until all conflicts of interest are appropriately managed. The IRB may require disclosure of the conflict to subjects in the consent document or information sheet/process as part of the management plan.

1. Does the Principal Investigator, co-investigator, any of the research personnel, or any of their family members have a managerial role in any entity associated with the research, or otherwise have a significant financial relationship or any other relationship with the funding source or company, or have a financial stake in a product associated with this research that may be viewed as affecting the protection of human subjects involved in the project, the scientific objectivity of the research or the integrity of the research or the research data?

☒ No ☐ Yes If yes, attach a COI Statement of Explanation that includes a list of names of the investigators with conflicts, a brief description of each conflict, and the plan for managing the conflicts or the management plan approved by the Conflict of Interest Committee.

IV. Performance Sites:

A Performance site is a location where the research is conducted, where data are gathered from subjects or records, where subjects are recruited for the research, and/or where subjects provide consent for research participation. These sites are considered performance sites whether or not the research activities are funded or unfunded.

1. Are there any non-DePaul performance sites? ☒ No ☐ Yes If yes:

a) List the performance sites and explain their role in the research.

b) Copies of the collaborating institution's IRB approval memos attached. ☐ Yes ☐ No ☐ Pending

c) If the collaborator has no IRB, copies of letters of collaboration or support in lieu of the approval memo are attached.

☐ Yes ☐ No ☐ Pending

V. Research Objectives (purpose, aims, or goals):

1. Explain the research purpose, aims, or goals.

Most web-based media do not have closed captions, and thus makes the media inaccessible to the Deaf community. The goal of this study is to evaluate the efficacy and acceptability of using automatic speech recognition to generate closed captioning.

VI. Research Procedures and Target Populations:

1. Number of subjects to be recruited and enrolled (or cases/records accessed, or specimens collected):

	Total: 110	Female: N/A	Male: N/A

[Note: If there are multiple subgroups creating the total, detail the subgroups in the research summary in #5 below. For example, if you will recruit and enroll students, parents, and teachers, describe the anticipated number of each in your research summary.]

2. Age range(s) for subjects (if more than one group, list each separately):

18-65

3. Describe the inclusion and exclusion criteria for selecting subjects for the research. If criteria differ for different groups (e.g., age groups), state criteria for each group separately.

Inclusion Criteria:	Deaf, Hard-of-Hearing, and Hearing adults with some college education who might use captioning
Exclusion Criteria:	Adults who have not attended college

4. Indicate any vulnerable populations that may be targeted for inclusion in the research. Research involving prisoners as research subjects may not be approved under an exemption determination. See the form instructions for other specific exclusions of vulnerable populations from exemption categories.

☐ Minors (under the age of consent in the jurisdiction where subjects are recruited; 18 in Illinois)
☐ Pregnant Women
☐ Decisionally Impaired Persons
☐ Economically or Educationally Disadvantaged Persons
☐ DePaul Students
  ☐ Psychology Subject Pool
☐ DePaul Employees
☐ K-12 Students
☐ Other

Description:

5. Provide a summary of the research that includes:
   - the research methods and procedures,
   - the frequency and duration of research procedures,
   - the type of data to be collected,
   - the method of data collection,
   - the location of the subjects during the research or while data are collected from or about them, or the location of records if data will be collected from records.

Research Methods
The plan of work calls to evaluate different presentation strategies for making online news videos more accessible. The research will send an email solicitation (see appendix F) to a mailing list comprised of members of the Deaf community. The email has a short set of screening questions and an invitation to contact the researcher for a URL (link) if they qualify to take the test. There will also be a poster (see appendix H) posted outside CDM 822, where there is a lot of foot traffic, to recruit hearing participants.

The test is entirely Web based. After reading the information sheet (appendix A) and clicking to confirm, the participant will fill out a pre-test questionnaire (appendix B) so we can gather demographic details. In the test, a participant will view four different caption styles. After viewing a video, a participant will answer questions about the video. See Appendix C for a transcript of each of the videos and the set of questions that the participant will answer. In the test itself the order of videos will be randomized. After answering the questions pertaining to the fourth video, the participant will complete a post-test questionnaire (see Appendix D). At the conclusion, participants have the option of having a gift card sent to them (Appendix E).

Frequency and Duration of Research Procedures
It is anticipated that each participant will need up to 60 minutes to complete the test. Each participant only takes the test one time. No personally-identifiable information will be solicited.

Type of data being collected
In addition to the demographic information from the pre-test questionnaire, the type of data will include participant responses to closed-ended questions regarding the videos. There are open-ended questions in the post-test questionnaire.

Method of data collection
The data will be gathered via a web application. Users will login to a web site, view the captioned videos and answer the questions online.

Location of test participants
The geographic location of the participation will be unknown, since they will be using a web site to take the test.

The responses to the test questions will be stored on a password-protected web server.

6. Indicate which of the following will be used to recruit, identify or initially contact the subjects. Attach any recruitment scripts, emails, flyers, online postings, etc., for IRB review.

☒ Recruitment Email	☒ Recruitment flyer/poster
☐ Recruitment Letter	☐ Internet (specify websites below in question #7)
☐ Ad (print)	☐ Ad (radio-provide script, then tape)
☐ Ad (TV - provide script, then video)	☐ Social media (specify below in question #7)
☐ Mailed letter	☐ Telephone script for response to ad
☐ Verbal script/announcement	☐ Brochure
☐ Screening script to confirm eligibility	☐ No recruitment materials
☐ Other, Specify:	

7. Describe how each of the recruitment items/methods indicated above will be used to recruit, identify or initially contact the subjects and identify which type of recruitment item/method will be used for differing target populations (if applicable). Keep in mind that recruitment methods should protect subject privacy.

Email invitations (Appendix F) will be sent to a mailing list of Deaf people. This is a list of alumni that I have known since my time at Gallaudet and I am the custodian of this information. There is no need to obtain a list that I already have. The invitation will include a brief screening survey and instructions how to contact me if they qualify. They must answer 'yes' to all the 4 questions asked in the email. I will reply with an URL with embedded code that will access the test page. All references to the code and the code itself are discarded once the participant begins the test. There will also be a poster (see appendix H) posted outside CDM 822, where there is a lot of foot traffic, to recruit hearing participants.

8. Describe the informational process for informing subjects and obtaining their voluntary agreement, including how and when the information will be provided to the subjects and by whom.

When the participant follows the survey link, they will see the information sheet for the study (see appendix A). At the bottom of the page will be two buttons "Agree and Continue" or "I do not agree". If the person agrees, then they will have the option of having a copy of the information sheet sent to them. The email address acquired for this purpose will not be stored in the data file containing the participant's responses.

VII. Confidentiality:

1. Will any data or information be collected using audio or video recordings of subjects? ☐ Yes ☒ No

If yes, explain if and how subjects will be identified in and on the taped materials, what measures will be used to maintain the confidentiality of the taped data, how long the taped materials will be kept, and if the tapes will be destroyed, when and how that will occur.

2. Will the data or information being collected be recorded such that the subjects can be identified directly or through indirect identifiers (e.g., codes, demographics, Social Security Number, IP address, record number) linked to the subjects? ☐ No ☒ Yes

If yes, describe how the confidentiality of the data will be maintained (i.e. how the data will be recorded and stored, who will have access, and how long the data will be maintained in an identifiable format).

After the test, the participant will be asked if they want to claim a gift card. They will need to supply their email address in order to claim it. To maintain confidentiality, the email addresses will be stored in a file entirely separate from the data recorded from the test. After the gift cards are distributed via email, the file will be destroyed.

3. Will identifiable information be made available to anyone else other than the PI? ☒ No ☐ Yes

If yes, explain who and why they will be given access to identifiable data.

VIII. Payment, Compensation, Reimbursement:

1. Will subjects be given gifts, payment, compensation, or reimbursement? ☐ No ☒ Yes

If yes, describe the amount and method of payment (e.g., cash, gift card, check, raffle):

$15 gift card that may be funded by CDM summer competitive grants (Appendix E). Per http://policies.depaul.edu/policy/policy.aspx?pid=297 I will seek approval from the Office of Accounts Payable. Because the amount is under

$75, there is no need to acquire additional personal information.

IX. Use of Existing Data

1. Will existing (archival) identifiable data, documents, records, or biological samples be used? ☒ No ☐ Yes

***Existing means that the data exists now, at the time you are proposing the study.**

If yes:

a) What is the source of the existing materials or information?

b) If the source was a previously approved and conducted IRB research protocol, provide the IRB number.

c) Indicate the type or source of data that will be collected and used for the research (e.g., medical records, survey responses).

d) Will the data be collected or retrieved directly from the record source? ☐ No ☐ Yes

e) Will the data be recorded in the research records with:

- ☐ Direct Identifiers (e.g., name, initials, social security number, medical record number)*
- ☐ Indirect Identifiers (e.g., assigned a code which can be used by the investigators to identify individuals)*
- ☐ No Identifiers (the researcher cannnot identify the individual from the information provided)

***Research does not meet the criteria for an exemption determination under category 4.**

X. HIPAA

1. Does the research involve the use, access, or disclosure of protected health information (PHI) as defined in the HIPAA Privacy Rule? ☒ No ☐ Yes

If yes, provide a detailed explanation of the type of PHI that will be used, accessed or disclosed, and identify the source of the PHI.

XI Assurances

Principal Investigator's (PI) Assurance:
I certify that the information provided in this application is complete and accurate. I understand that as Principal Investigator, I have the ultimate responsibility for the protection of the rights and welfare of human subjects enrolled in the research. I assure that I will conduct the study ethically and in compliance with all Federal regulation, state and local laws, and DePaul IRB policies and procedures. I assure the following:

- I have completed the required human subjects training program as outlined in current DePaul IRB policy.
- The project will be performed by qualified and trained personnel in accordance with the DePaul IRB approved protocol.
- No changes will be made to the protocol or approved protocol documents without prospective IRB approval.
- Subjects will be provided full information about the study before they begin participation using the IRB approved informational process and without undue influence or coercion.
- Any unanticipated problems that occur during the conduct of the research will be reported to the IRB in a timely manner according to policy.
- I will submit a Final Closure Report once the study is completed or before I leave DePaul University.

I further certify that the proposed research is not currently underway and I will not begin the project until final approval has been obtained.

Principal Investigator's Signature: Brent Shiver
Digitally signed by Brent Shiver
DN: cn=Brent Shiver, o, ou, email=bshiver@gmail.com, c=US
Date: 2013.02.17 22:07:32 -06'00'

Date: Feb 27, 2013

Faculty Sponsor's Assurance for Student, staff position functioning as a training position or fellow, such as a research fellow:

By my signature as faculty sponsor on this research application, I certify that the student, staff member in a training position, or fellow is knowledgeable about the regulations and policies governing research with human subjects and has sufficient training to conduct this particular study in accordance with the approved protocol. In my role as faculty sponsor, I assure the following:

- I have completed the required human subjects training program as outlined in current DePaul IRB policy.
- I have read the research application and supporting materials and approved them for submission to the IRB.
- As the faculty mentor for this student on this project, I agree to meet with the student, staff member in training, or fellow investigator on a regular basis to monitor and assist them with conducting the research.
- I agree to be available, personally, to supervise and assist the investigator in the event problems arise during the conduct of the study.
- I ensure that the investigator will promptly report any unanticipated problems that occur during the conduct of the research in a timely manner and in accordance with IRB policy and procedures.
- I will ensure that a Final Study Closure Report will be submitted to the IRB when the research is completed or before the investigator leaves the university.
- If I am unavailable for an extended period of time, such as when on sabbatical or leave, I will arrange for an alternate faculty sponsor to assume my responsibilities during my absence and the DePaul IRB will be informed of the change via an amendment.

I further certify that the proposed research is not currently underway and will not begin until IRB approval has been obtained.

Faculty Sponsor's Signature: Rosalee Wolfe
Digitally signed by Rosalee Wolfe
DN: cn=Rosalee Wolfe, o=DePaul University, ou=School of Computing, email=wolfe@cs.depaul.edu, c=US
Date: 2013.02.18 09:34:15 -06'00'

Date: February 28, 2013

*The faculty sponsor must be a member of the DePaul faculty (full time, part time, or adjunct). The faculty sponsor is considered the party ultimately responsible for the legal and ethical performance of the project and ensuring that all duties of the Principal Investigator are fulfilled.

DePaul IRB Amendment Application

(for Previously Approved Research)

Step I: Project Information

Principal Investigator: Brent Shiver Phone: 312-957-8950 E-mail Address: bshiver@cs.depaul.edu

Protocol Number (According to Last Approval Memo): BS031313CDM Date of Last Approval: 4/19/2013
Project Title: Improving Deaf Accessibility to Web-Based Multimedia

What level of review did your original project receive? ☒ Exempt ☐ Expedited ☐ Full Board
(see the original approval memo or notification email)

Step II: Type of Amendment

Check all of the appropriate boxes and attach revised forms and/or additional information, when necessary:
1. ☒ Revisions to Assent/Consent Form(s), HIPAA Authorization(s) or Information Sheet(s)
2. ☐ Revisions to Assent/Consent Procedures
3. ☒ Revisions to Inclusion/Exclusion Criteria
4. ☒ Revisions to Methodology or Data Collection Procedures
5. ☒ Change/Addition of Recruitment Materials
6. ☒ Addition of funding source or change in status of funding. Please specify: The grant application from the CDM PhD summer program has been approved.
7. ☐ Other--Please specify:
8. ☐ Change/Addition of Research Staff (Explain the role of the new person and/or any related changes to the research, as applicable.)
9. ☐ Change/Addition of Collaborators/ Sites (Explain the role of the new site, the change, and provide documentation of collaboration, if applicable.)

Step III: Summary of Changes

Summarize the proposed amendment(s) and provide the rationale for the change(s). Provide a brief discussion of the impact of the change on research methodology, if applicable.

- Changed email address of PI from bshiver@gmail.com to bshiver@cs.depaul.edu. The new email address is more appropriate for the study.
- Changed estimated time of study from 90 minutes to 60 minutes. It has been demonstrated in field tests that the study typically takes about 60 minutes instead of 90 minutes.
- Additional 10 participants, all hearing. Doctoral committee preferred to add 10 hearing participants as a baseline for the study.
- Other changes have to do with clarity issues on pre-test and post-test questionnaires

Step IV: Risk-Benefit Assessment

A. Are the risks to participants affected (increased or decreased) by the amendment? ☐ Yes ☒ No

B. If yes, describe how the amendment will affect the risk-benefit ratio.

Step V: Informed Assent/Consent and HIPAA Research Authorization

(You may skip this section if your proposed amendment itself is a change to the assent/consent/HIPAA form and/or process.)

A. Could the amendment affect the informed assent/consent process? ☐ Yes ☒ No ☐ N/A

B. Could the amendment affect the informed assent/consent forms or HIPAA form? ☐ Yes ☒ No ☐ N/A

C. If you answered yes to any of the items in this Step above, please describe the amendment's effect, including whether it is necessary to inform those who have already consented and, if so, how you would do so.

****Submit 1 electronic copy to sloesspe@depaul.edu and 1 signed, hard copy by mail or fax.****
If you have questions or concerns, please contact Susan Loess-Perez at sloesspe@depaul.edu.

Step VI: Signatures

Principal Investigator: _____ Print Name: _____ Date:_____

Co-Investigator (if appropriate): _____ Print Name: _____ Date:_____

Faculty Sponsor (if appropriate): _____ Print Name: _____ Date:_____

****Submit 1 electronic copy to sloesspe@depaul.edu and 1 signed, hard copy by mail or fax.****
If you have questions or concerns, please contact Susan Loess-Perez at sloesspe@depaul.edu.

DEPAUL
UNIVERSITY

Office of Research Services
Institutional Review Board
1 East Jackson Boulevard
Chicago, Illinois 60604-2201
312-362-7593
Fax: 312-362-7574

Research Involving Human Subjects

NOTICE OF INSTITUTIONAL REVIEW BOARD ACTION

To: Brent Shiver, Graduate Student, College of Computing and Digital Media

Date: May 23, 2013

Re: Research Protocol # BS031313CDM-R1
"Improving Deaf Accessibility to Web-Based Multimedia"

Please review the following important information about the review of your proposed research activity.

Review Details
This submission is an Amendment. Amendment R1 involves 1) changing the email address of PI; 2) changing the estimated tome of the study from 90 minutes to 60 minutes; 3) the addition of 10 participants; and 4) clarifying pre-test and post-test questionnaires.

Your research project meets the criteria for Exempt review under 45 CFR 46.101 under the following category:

> *(2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:*
> *(i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.*

Approval Details
Your research was reviewed and approved on May 23, 2013.

> **Number of approved participants:** 110 Total
> ***You should not exceed this total number of subjects without prospectively submitting an amendment to the IRB requesting an increase in subject number.***
>
> **Funding Source:** 1) DePaul CDM PhD Summer Funding and Competitive Grants
>
> **Approved Performance sites:** 1) DePaul University

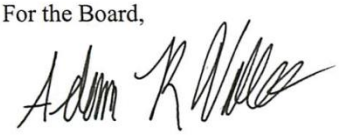
Reminders
- Under DePaul's current institutional policy governing human research, research projects that meet the criteria for an exemption determination may receive administrative review by the Office of Research

Services Research Protections staff. Once projects are determined to be exempt, the researcher is free to begin the work and is not required to submit an annual update (continuing review). As your project has been determined to be exempt, your primary obligation moving forward is to resubmit your research materials for review and classification/approval when making changes to the research, but before the changes are implemented in the research. **All changes to the research must be reviewed and approved by the IRB or Office of Research Services staff.** Changes requiring approval include, but are not limited to, changes in the design or focus of the research project, revisions to the information sheet for participants, addition of new measures or instruments, increasing the subject number, and any change to the research that might alter the exemption status (either add additional exemption categories or make the research no longer eligible for an exemption determination).

- **Once the project is complete, you should submit a final closure report to the IRB.**

The Office of Research Services would like to thank you for your efforts and cooperation and wishes you the best of luck on your research. If you have any questions, please contact me by telephone at (312) 362-6168 or by email at avander1@depaul.edu.

For the Board,



Adam R. Vanderloo, JD
Research Protections Coordinator
Office of Research Services

Cc: Rosalee Wolfe, Ph.D., Faculty Sponsor, College of Computing and Digital Media

Appendix I – Information Sheet (Study 2)

INFORMATION SHEET FOR PARTICIPATION IN RESEARCH STUDY

Improving Deaf Accessibility to Web-based Multimedia

Principal Investigator: Brent Shiver, graduate student, School of Computing
Institution: DePaul University, USA
Faculty Advisor: Dr. Rosalee Wolfe, Ph.D., School of Computing, DePaul University
Research Team: no additional personnel

We are conducting a research study because we are trying to learn more about effectiveness of automatic speech recognition to improve online accessibility, specifically news videos. We are asking you to be in the research since you are either hearing, hard-of-hearing or deaf, you might use closed captions and you have attended college. If you agree to be in this study, you will be asked to use the Web to complete pre-test and post-test questionnaires, view four videos, and answer questions about the videos.

This study will take up to 60 minutes of your time. Your survey responses will be anonymous.

Your participation is voluntary, which means you can choose not to participate. There will be no negative consequences if you decide not to participate or change your mind later after you begin the study. You can withdraw your participation at any time prior to submitting your survey. If you change your mind later while answering the survey, you may simply exit the survey

If you pass the pre-screen questionnaire and complete the survey you will be given a $15 gift card. After concluding the survey, you will be taken to a separate page where you will enter your email address so that you can get compensated for being in the study. The email address you give will be used solely for sending your gift card, and will not be used for any other purposes. Further, it will not be stored permanently or linked to your survey responses.

You must be age 18 or older to be in this study. This study is not approved for the enrollment of people under the age of 18.

If you have questions, concerns, or complaints about this study or you want to get additional information or provide input about this research, please contact Brent Shiver at 312.957.8950/bshiver@cs.depaul.edu or faculty sponsor Dr. Rosalee Wolfe at 312.362.6248/wolfe@cs.depaul.edu.

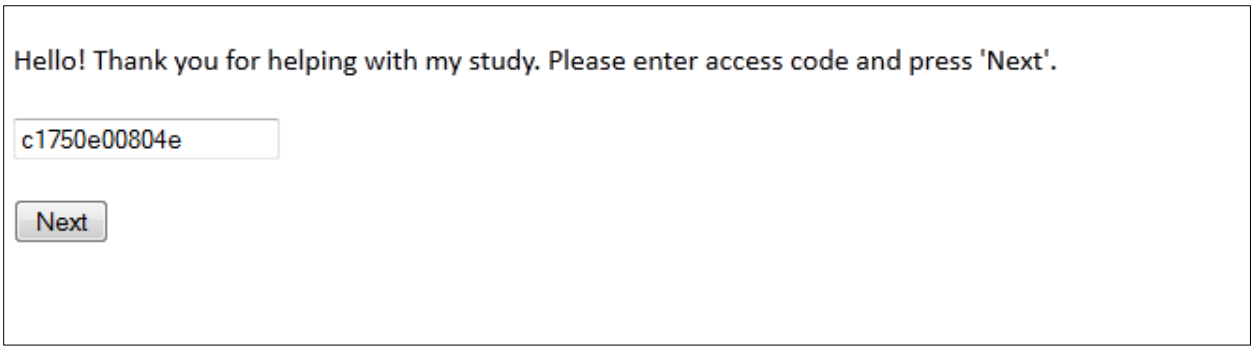
If you have questions about your rights as a research subject you may contact Susan Loess-Perez, DePaul University's Director of Research Compliance, Office of Research Protections in the Office of Research Services at 312-362-7593 or by email at sloesspe@depaul.edu. You may also contact DePaul's Office of Research Protections if:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.

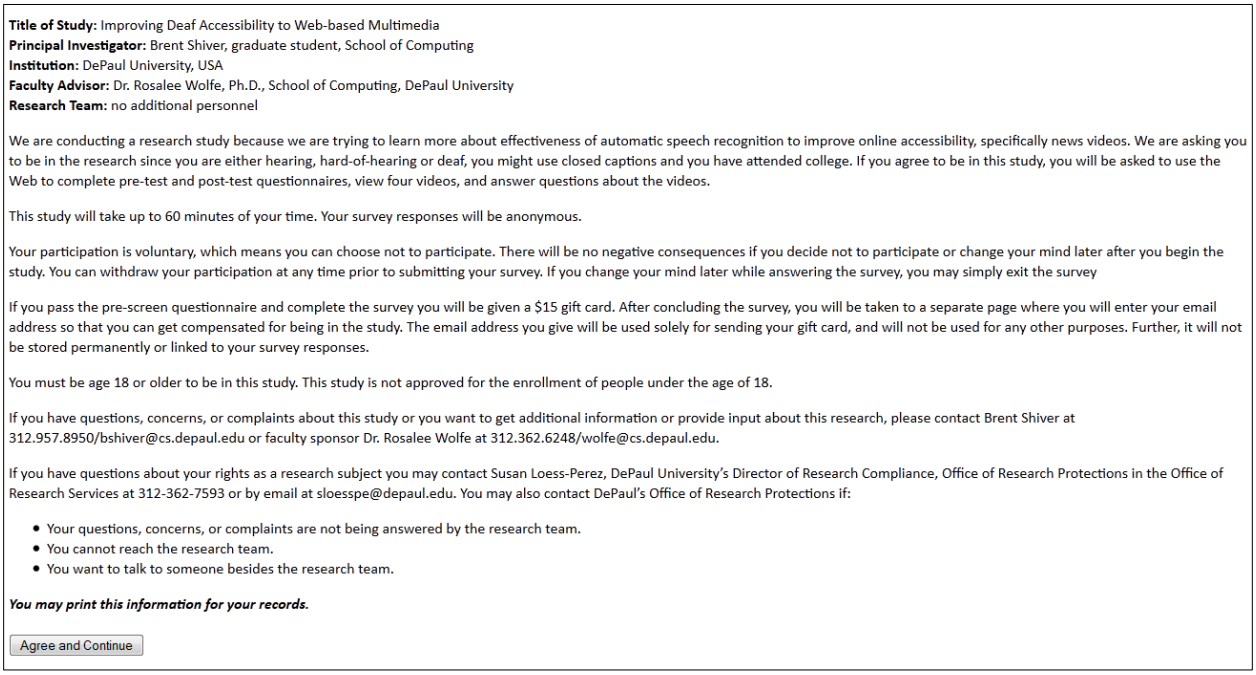
You may print this information for your records.

Appendix J – Application Flow Screenshots

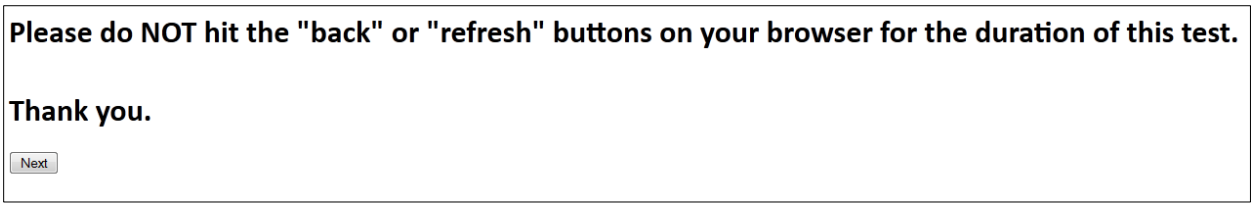
1. Cover sheet



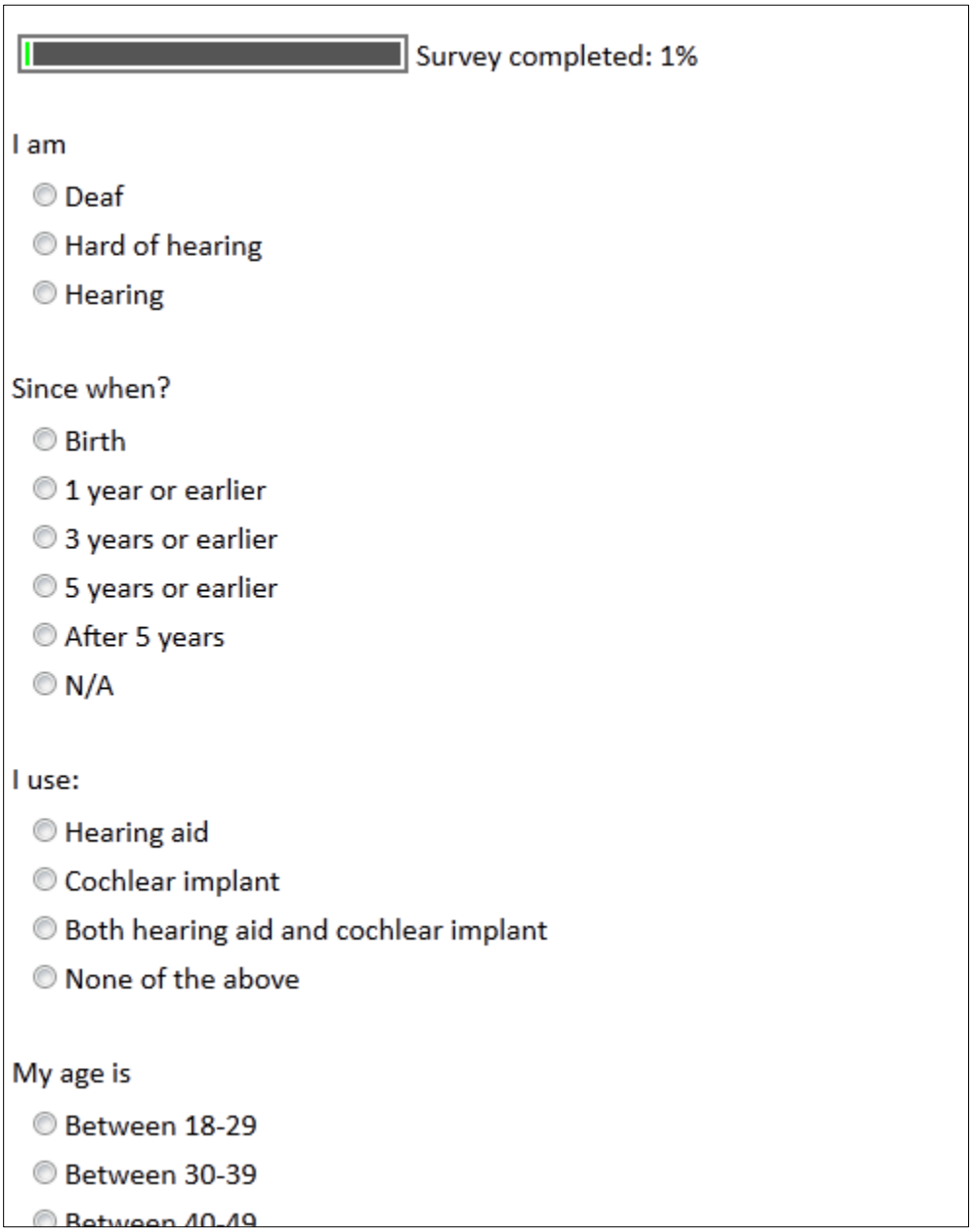
2. Informed consent



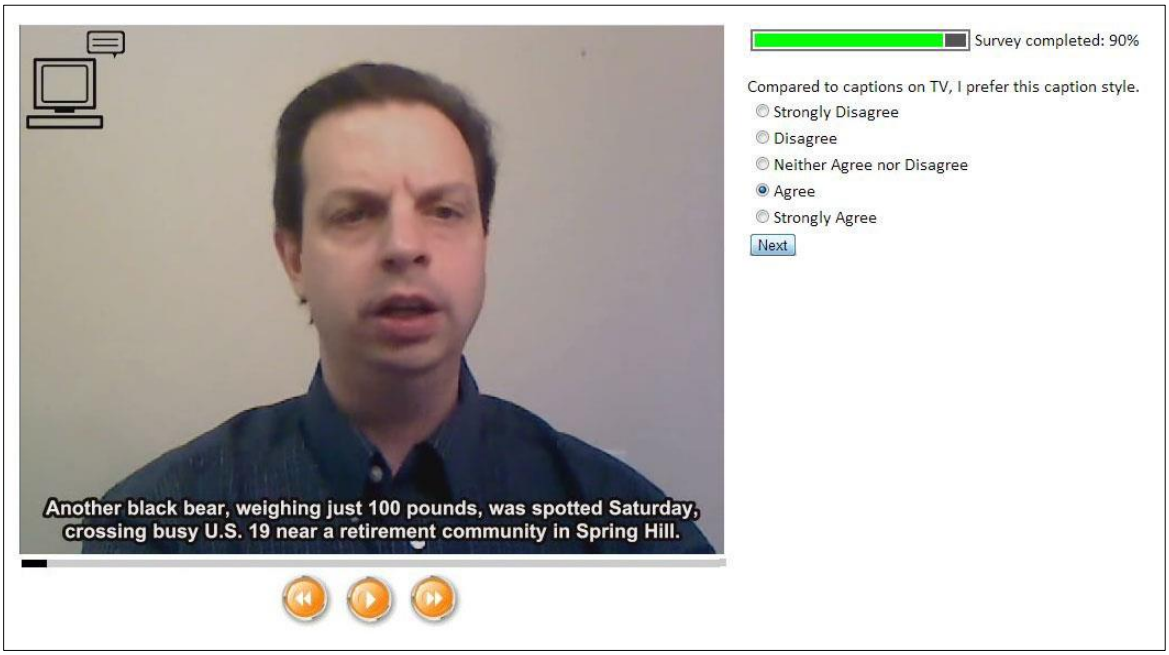
3. Browser warning message



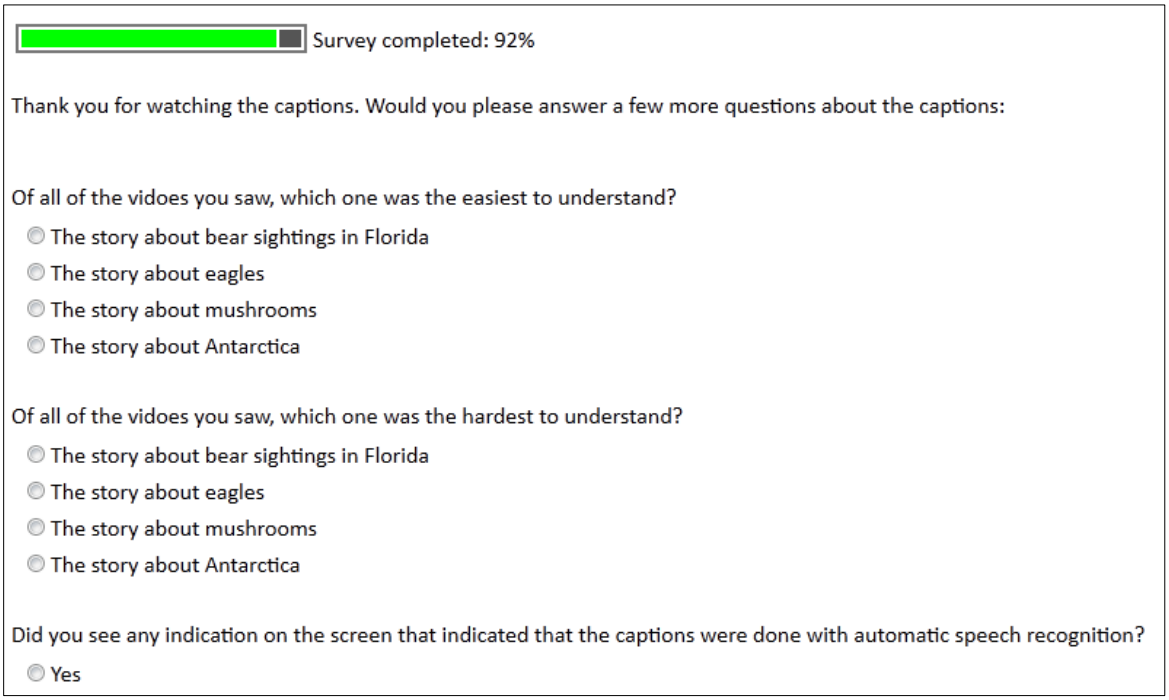
4. Pretest questionnaire



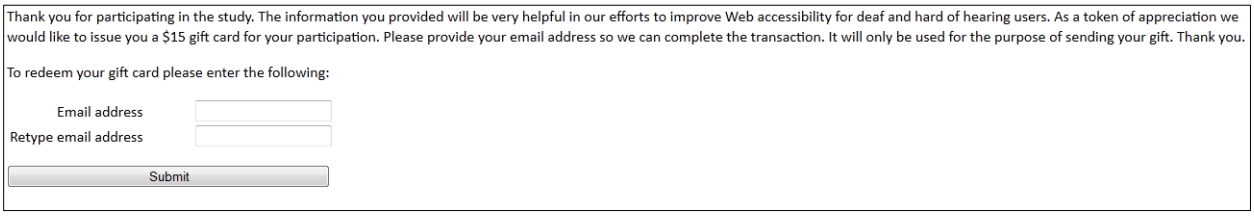
5. Main video and questions page



6. Post-test questionnaire



7. Thank you page



Appendix K – Pre-test Questionnaire (Study 2)

Age [ ] 18-29 [ ] 30-39 [ ] 40-49 [ ] 50+

Gender [ ] male [ ] female

Occupation [ ] non-profit/government [ ] educational [ ] professional [ ] homemaker [ ] unemployed [ ] student [ ] other

Degree of hearing loss [ ] profoundly deaf [ ] hard of hearing [ ] hearing

Since when? [ ] birth [ ] 1 year or earlier [ ] 3 years or earlier [ ] 5 years or earlier [ ] after 5 years [ ] N/A

Do you use:

[ ] hearing aid [ ] cochlear implant [ ] both hearing aid and cochlear implant [ ] none of the above

Highest education level completed

[ ] less than high school [ ] high school/GED [ ] some college [ ] 4-year college degree

Do you watch news videos online?

[ ] yes [ ] no

When you watch television or movies do you watch captions?

[ ] Always, whenever available [ ] Sometimes [ ] Rarely [ ] Never

How often do you use the Internet to

	Rarely	Once a week	Once a day	Many times a day	How many hours?
Read email					
Read news articles					
Online auction					
Online shopping					
YouTube videos					
Social media (Facebook, LinkedIn, Twitter, etc)					
Training/remote class					

Appendix L – Video Transcripts (Study 2)

The following is a transcript of a bear sightings story that the participant viewed as a captioned video.

BROOKSVILLE – It's that time of year again, when young bears are on their own for the first time. And that means more sightings where bears usually aren't supposed to be.

This weekend, a 258-pound black bear took a brief tour of Brooksville, with a harried host of police and firefighters chasing.

Another black bear, weighing just 100 pounds, was spotted Saturday, crossing busy U.S. 19 near a retirement community in Spring Hill. Several other bear reports have emanated from Citrus County.

"It is real difficult for the younger bears to find their own territory around here," said Niki Everitt, bear hot line coordinator for the Gulf Coast Conservancy.

Brooksville's bear first was spotted late Thursday crossing State Road 50, headed toward Tom Varn Park. Police and firefighters tracked the bear through the park and the Brooksville Quarry golf course.

The bear then meandered down the middle of Broad Street before being surrounded near Luigi's Pizza. An official with the Florida Game and Fresh Water Fish Commission came with a tranquilizer spear, planning to stick the bear by hand.

"The guy saw the size of the bear and figured that wasn't a great idea," said Capt. Frank Phillips of the Brooksville Fire Department.

Emergency officials waited until a tranquilizer gun was brought from Land O'Lakes. The bear then was fitted with a transmitter collar and got a free ride to the Chassahowitzka Wildlife Management Area.

It is not unusual for bears to trundle into residential areas this time of year, experts say. Mothers give their male offspring the boot once the youngsters are 2 years old. With residential areas growing, the young bears keep finding smaller and smaller areas in which they can establish their own territory.

"They're trying to find a territory of their own, where they won't get beat up," said Lt. Rip Stalvey, a game commission spokesman.

Everitt said people should not be too concerned about the recent bear sightings, since "we have never had a bear attack in Florida."

Black bears primarily eat the tender parts of palmettos and Sabal palms, as well as acorns and berries. Recent weeks of drought likely have reduced their food supply.

"If we don't get some relief soon," Everitt said, "we're probably going to see a lot more of it.

End transcript

The following is a transcript of a story about eagles that the participant viewed as a captioned video.

No other bird is quite as famous. American eagles have been around so long, they're a U.S. fixture. Just do a little bird watching on your quarters and dollar bills. Check out your local post office. Look at the top of flagpoles. Is the eagle an important part of American culture? Without a doubt!

To appreciate the current popularity of these birds, you need historical perspective. American eagles now occupy a commanding perch as the premier American symbol, but for hundreds of years they have had their ups and downs.

Centuries before the American Revolution, eagles were displayed on ancient monuments, statues, and coins. Then in the American colonies, the eagle was used on a 1700 brass token in New York and a 1776 copper penny in Massachusetts.

For a moment in American history, though, eagles almost didn't get the seal of approval. On the sweltering afternoon of July 4, 1776, committees began to design the official seal of the United States, the mark that is imprinted in wax on all our country's official treaties. A final version was not adopted until June 20, 1782 – six years later. Obviously, this was not an easy task.

The committees finally agreed on a design that includes the distinctive American eagle, bald-headed and beautiful, as the large central figure, bearing a shield, clutching an olive branch in its right talon, and thirteen arrows in its left. Just in time, too. The seal was needed in a few months for signing the peace treaty with Great Britain.

How did those planners finally determine that American eagles would be the very best symbol? One theory is that they looked back at the many eagles connected with ancient majesty, power, and military victory. But Ben Franklin had a different opinion. He wanted the U.S. to adopt the turkey as its emblem. Franklin argued that, unlike eagles, the turkey is native to America. He also pointed out that the Thanksgiving gobbler is quite fierce if attacked, but eagles are known to steal other birds' prey.

While the committees were working on the seal, General George Washington was making a fashion statement. During his Revolutionary War service as commander of the American troops, he chose to wear a coat adorned with eagle-design brass buttons. Everywhere he went, patriotic crowds waved items bearing the picture of an eagle. His taste in uniforms probably helped the committee members decide to include the eagle.

Once the eagle design was selected, eagles were the bird of honor. You could see them on weathervanes, silver cups, buttons, money, ships, door knockers, over the doors of buildings, and woven into cloth. Just about everything for sale displayed an eagle. By the turn of the century (from the 1700's into the 1800s), Paul Revere was handing out business cards engraved with an American eagle, its wings outstretched, bearing the shield, the arrows, and the olive branch.

American eagles continue to be popular today. Look in the phone book for businesses with the word "eagle" in their names, or scan store shelves for eagle brand names. This bird is so well known around the world that the U. S. host symbol for the 1984 Summer Olympics in Los Angeles was a cartoon-style American eagle sporting a red, white, and blue top hat. If that isn't enough name and image recognition, how about eagles in outer space? In 1969 when the Lunar Lander touched down on the moon for the first time, its name was announced worldwide with these words: "The Eagle has landed."

Recently American bald eagles were news again. After many years of concern, they've been taken off the endangered species list in the continental United States. What helped these magnificent birds regain a talon-hold? Laws that ban the use of the pesticide DDT, for one thing. In addition, there are laws that protect eagles from hunters, prey, and habitat threats. Thanks to the efforts of many Americans, the number of eagle families has increased greatly, according to wildlife officials, from only 417 nesting pairs in 1963 to 5,748 pairs in 1998! Far from extinct, the American eagle is flying high!

End transcript

The following is a transcript of a story about mushrooms that the participant viewed as a non-captioned video.

Most of us are familiar with one or two kinds of mushrooms, usually white or brown varieties that find their way onto pizzas. Actually, more than 3,000 types grow around the world in a wide variety of flavors and sizes. Some are less than an inch high, and others are more than 15 inches tall. Some have unusual names like Portobello and Black Trumpet, and they are listed on sophisticated menus in fancy restaurants. But many centuries ago, long before pizzas and fancy restaurants existed, people were eating mushrooms.

Ancient hieroglyphics from more than 4,600 years ago tell us Egyptians called mushrooms "the magic food." They believed eating them resulted in immortality, and only pharaohs were given this privilege so that they could live forever. Of course, this meant Egyptian royalty enjoyed all the delicious mushrooms since no commoner could touch them! Other ancient civilizations in places such as Russia and Mexico thought mushrooms had ingredients that could produce superhuman strength and even help locate lost objects.

Centuries ago, people still associated magic with mushrooms. Sometimes they observed unusual places in a meadow, like a patch of bright green grass or a spot of bare soil. Then

they imagined these places were the result of footprints left by fairies dancing at night. When mushrooms appeared near the edge of these "fairy rings," people liked to think of them as seats where the tired fairies could rest. But today we have a more scientific approach to the mushroom.

All of the many species of mushrooms are classified as fungi. They are plant-like organisms that usually grow in damp, dark places like caves or forest floors, but they can also grow in grassy areas. Fungi work with other plants and animals called decomposers to keep the soil fertile for plant growth. Like many other plants, mushrooms serve as a source of food for insects and small animals. Mushrooms differ from green plants because they lack chlorophyll and do not require sunshine to grow.

As the demand for mushrooms increased over the centuries, people established mushroom farms to plant and grow the fungi in special environments. Some farms were in caves, some underground, and some in special buildings. In the 1600s, for example, France developed the formal cultivation of mushrooms in special caves near Paris. Until the 1940s, most mushroom farms were in the Far East, especially China and Japan. Then during World War II, many American soldiers tasted the delicious varieties of mushrooms and learned about mushroom farming. After the war, they took this knowledge back to the United States, which soon became one of the world's major mushroom producers.

Health and safety are always concerns when growing any crop. One of the complications with mushrooms is that they can be poisonous or nonpoisonous. Common nontoxic varieties such as table and field mushrooms are safe to eat and can be purchased in grocery stores. These mushrooms are praised by health experts because they are fat-free, cholesterol-free, and low in calories. They are rich in B-vitamins, potassium, phosphorus, and iron. Chefs use them in dishes ranging from soups to gourmet sauces, and some mushrooms even have medicinal benefits. The silver-ear mushroom, for example, can be used to lower blood pressure.

Over the years, edible mushrooms have proven to be extremely popular in the marketplace. Today the USA is the world leader in supplying mushrooms, and other major contributors include France, China, Canada, Great Britain, and Italy. In 1986, 470 million pounds of mushrooms were produced throughout the world, and by 1999 this figure had almost doubled. Production was up to 860 million pounds and the market value was $867 million. At this rate, it appears safe to say mushrooms are here to stay!

End transcript

The following is a transcript of a story about Antarctica that the participant viewed as a captioned video.

Learning How to be a "Happy (Cold) Camper" by Josh Williams

Ross Ice Shelf, Antarctica, Jan 19, 1999

Anyone who travels to what's known as "field camps" in Antarctica is supposed to take a survival course. The regular "Snowcraft" course, which is generally called "Happy campers school," includes camping out for the night in the snow.

This year, most of the journalists the National Science Foundation brought to Antarctica were late getting here because weather and other problems delayed their flights from Christchurch. They were told they could do a half-day lecture course on survival and pass up the camping trip.

But Bob Boyd of Knight-Ridder and I decided to do the full school. We arrived here on time and neither of us wanted to pass up the chance to experience even a little taste of what Antarctica's first visitors lived through. I'm glad I didn't miss being a happy campus even though it took most of Jan 19 and 20.

Our instructor was Bill McCormick, who's head of search and rescue at McMurdo and works as a mountain guide, including on Mt. McKinley in Alaska, when he's not here.

Every vehicle that leaves McMurdo, whether it's an airplane or a snowmobile, has to carry survival equipment, which includes a backpacking stove and fuel, enough sleeping bags for everyone in the vehicle, tents, food and other equipment. The idea of happy campers school is to make sure anyone who might need to use this equipment knows how.

After a morning of lectures, McCormick took the 11 of us in the course to the area on the Ross Ice Shelf used for the school. He told us that the idea was to learn how to survive. "We don't want this to be some kind of character-building thing. We aren't going to see if you crack."

In the snow, we learned how to build a snow trench – just what it sounds like, a trench in the snow that will shelter you from the wind – how to pitch the mountain tents found in survival kits and also the larger Scott tents used at field camps, and – most fun of all – how to build a snow mound shelter. For that, we piled up the bags with our sleeping bags and other equipment in them, covered them with a tarp and piled about two feet of snow on top of the pile. After packing down the snow, one person dug into the mound to drag out the bags, leaving a cozy hollow that is warmer than outside in the wind.

We also built a wall of snow blocks as a windbreak for a mountain tents and the "kitchen" area with four backpacking stoves.

Temperatures fell only into the 20s (above zero F), but the wind pushed the wind shill down to around 10 degrees while we were there. We know we were lucky. Earlier in the season temperatures might be 50 degrees colder for the school. I was comfortable in the sleeping bag in the mountain tent and even the dehydrated food wasn't bad.

End transcript

Appendix M – Content Questions (Study 2)

Questions about the bears story:

1. What was the purpose of this article?
    a. to motivate people to plant palms for young bears
    b. to inform people about emergency animal procedures
    c. to explain how Florida's police and firefighters rescue animals
    d. to explain why bear sightings occur in some areas of Florida
2. Bear sightings can be expected:
    a. in the winter
    b. in the summer
    c. near lakes or rivers
    d. near pizza restaurants
3. According to the article, bears use palmettos for:
    a. sleeping
    b. nourishment
    c. protection
    d. recreation
4. Why did the official decide the tranquilizer spear wasn't a good idea?
    a. The spear was dangerously sharp.
    b. The spear was too powerful.
    c. The bear was too large.
    d. The bear was too quick.

Questions about the eagles story:

1. Where have pictures of eagles been found since before the 1700s?
    a. on coins
    b. on stamps
    c. on the U.S. seal
    d. on business cards
2. Why did Benjamin Franklin want the turkey chosen instead of the eagle?
    a. Eagles are associated with war.
    b. Eagles are an endangered species.
    c. Eagles take the food of other birds.
    d. Eagles look unattractive in pictures.
3. The information in this article could best be used for a student research project on:
    a. American wars
    b. American presidents
    c. birds of distinction

    d. birds of prey
4. What would be a good title for this story?
    a. Money and Stamps
    b. Symbol of America
    c. An Endangered Species
    d. Eagles in Colonial America
5. What evidence supports the points made about the American eagle?
    a. documented evidence from research about eagles
    b. fictionalized anecdotes about eagles shown on stamps
    c. incidents revealed by people who train American eagles
    d. descriptive reports of American situations involving eagles
6. The author's purpose in writing this article was
    a. to encourage people to protect the bald eagle
    b. to explain how eagles came to be used in outer space
    c. to tell the history of the bald eagle in the United States
    d. to give examples of how eagles were used in the thirteen colonies

Questions about the mushrooms story:

1. The author's purpose in writing this article was to:
    a. tell why pharaohs ate mushrooms
    b. persuade people to buy more mushrooms
    c. explain the history and uses of mushrooms
    d. provide information for staying healthy with mushrooms
2. What is the warning the author gives about mushrooms?
    a. Only royalty can eat them.
    b. Some varieties can be poisonous.
    c. They should only be eaten in pizzas.
    d. They must be grown in dark, damp places.
3. How are mushrooms different from green plants?
    a. They are very colorful.
    b. They are a part of people's diet.
    c. They appear around "fairy rings".
    d. They don't need the sun for growth.
4. With which of the following statements would the speaker most likely agree?
    a. Mushrooms are profitable and fun to grow.
    b. Identifying poisonous mushrooms is a relatively simple process.
    c. The mushroom market has grown dramatically since World War II.
    d. Ancient civilizations were able to grow remarkably large mushroom crops.

Questions about the story about Antarctica:

1. Which *best* describes the difference between the half-day lecture course on survival and the Snowcraft course?
    a. The lecture course offered a reward, and the Snowcraft course offered a punishment.
    b. The lecture course was an introduction and the Snowcraft course was a conclusion.
    c. The lecture course provided guidelines, and the Snowcraft course offered practice.
    d. The lecture course instructed beginners, and the Snowcraft course instructed advanced students.
2. What do students learn in the regular Snowcraft course?
    a. How to improve one's character
    b. How to survive
    c. How to locate injured explorers
    d. How to work as a team
3. Why must every vehicle leaving McMurdo carry survival equipment?
    a. Traveling in Antarctica still involves risk.
    b. Camping out is part of the fun of the trip.
    c. Vehicles often break down in the snow.
    d. The equipment is user-friendly for beginners.
4. Which is the first step in building a snow mound shelter?
    a. gathering mounds of snow
    b. piling up the bags of equipment
    c. setting up a mountain tent
    d. digging out a hollow space
5. In the conclusion, what was the *most likely* the speaker's reason for mentioning the wind chill?
    a. To emphasize the actual weather conditions
    b. To show his acceptance of discomfort
    c. To discourage others from camping in Antarctica
    d. To admit that the journalists who avoided the trip were wise.
6. What kind of weather conditions did the speaker experience on his Antarctic camping trip?
    a. Painfully cold
    b. Comparatively mild
    c. Average for the season
    d. More severe than expected

Appendix N – Preference Questions (Study 2)

Please answer the following questions about the captioning style.

1. The captioning was easy to read.
   [ ] Strongly disagree [ ] Disagree [ ] Neutral [ ] Agree [ ] Strongly agree

2. The captioning made it easy to understand the story.
   [ ] Strongly disagree [ ] Disagree [ ] Neutral [ ] Agree [ ] Strongly agree

3. I have confidence in the accuracy of the captioning.
   [ ] Strongly disagree [ ] Disagree [ ] Neutral [ ] Agree [ ] Strongly agree

4. I like this style of captioning.
   [ ] Strongly disagree [ ] Disagree [ ] Neutral [ ] Agree [ ] Strongly agree

5. From the captions, I feel that I fully understood the story.
   [ ] Strongly disagree [ ] Disagree [ ] Neutral [ ] Agree [ ] Strongly agree

6. Compared to captions on TV, I preferred this style of captioning.
   [ ] Strongly disagree [ ] Disagree [ ] Neutral [ ] Agree [ ] Strongly agree

7. The color coding was helpful to understanding the story.
   [ ] Strongly disagree [ ] Disagree [ ] Neutral [ ] Agree [ ] Strongly agree

Appendix O – Post-test Questionnaire (Study 2)

1. Of all of the videos you saw, which one was the easiest to understand?
    a. The story about bear sightings in Florida
    b. The story about eagles
    c. The story about mushrooms
    d. The story about Antarctica

2. Of all of the videos you saw, which one was the hardest to understand?
    a. The story about bear sightings in Florida
    b. The story about eagles
    c. The story about mushrooms
    d. The story about Antarctica

3. Did you see any indication on the screen that indicated that the captions were done with automatic speech recognition?
    a. Yes
    b. No

4. Is there a better way to indicate that captions are created with automatic speech recognition
    a. Yes
    b. No

5. If yes, please tell us a better way: _____

6. If you had a choice of watching a video with no captions or watching a video with automatic captions, which would you choose?
    a. No captions
    b. Automatic captions

7. If you had a choice of watching a video with no captions or watching a video that had captions with errors, which would you choose?
    a. No captions
    b. Captions with errors

8. If you had a video with automatic caption containing errors, would you want color coding to indicate possible errors?
    a. Yes
    b. No

9. If you had a choice between watching a video with automatic captions, or reading a transcript, which would you prefer?
    a. Captioned video
    b. Transcript

Appendix P – Gift Card Message (Study 2)

Thank you for participating in the study. The information you provided will be very helpful in our efforts to improve Web accessibility for deaf and hard of hearing users. As a token of appreciation we would like to issue you a $15 gift card for your participation. Please provide your email address so we can complete the transaction. It will only be used for the purpose of sending your gift. Thank you.

Appendix Q – Email Solicitation (Study 2)

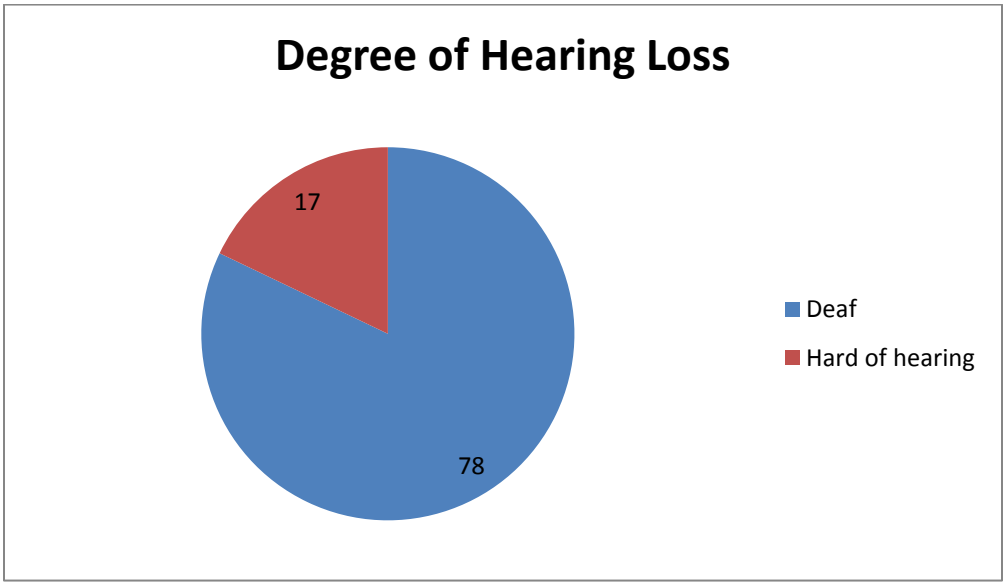
- Are you Deaf or hard-of-hearing?
- Have you attended college?
- Do you use captions when you watch television?
- Are you 18 years or older?

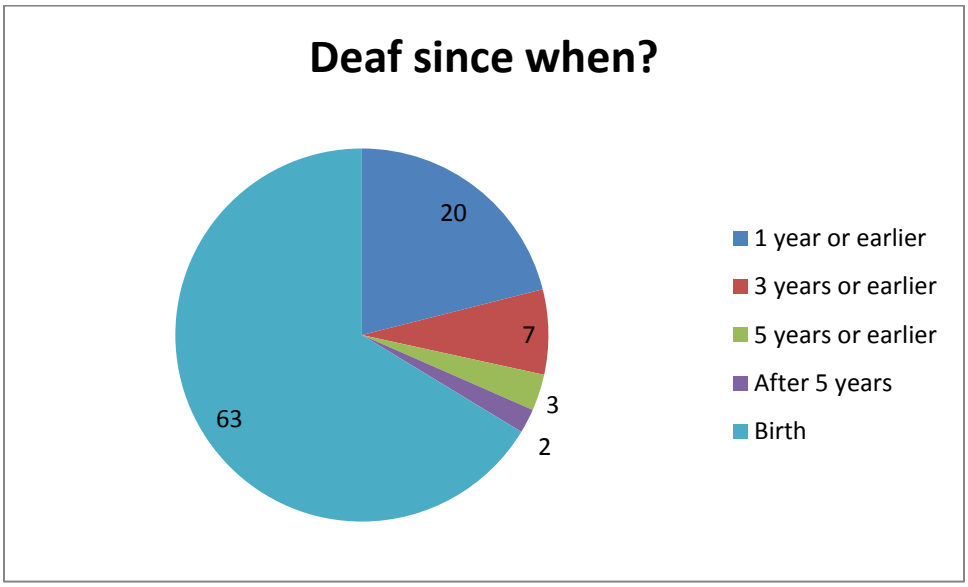
I am looking for deaf professionals who use the internet on a daily basis to participate in a study to identify ways to make online videos more accessible. The survey will take up to 60 minutes of your time. First a background questionnaire will need to be completed. It will ask about your deafness, your education, employment, and Internet usage. Then you will view a total of four videos and answer close-ended questions. Finally, we will have a brief post-test questionnaire. For your participation you will be given a $15 gift card. The test site will remain open until 100 people have taken the test. Email me if you are interested and I will send you further instructions.

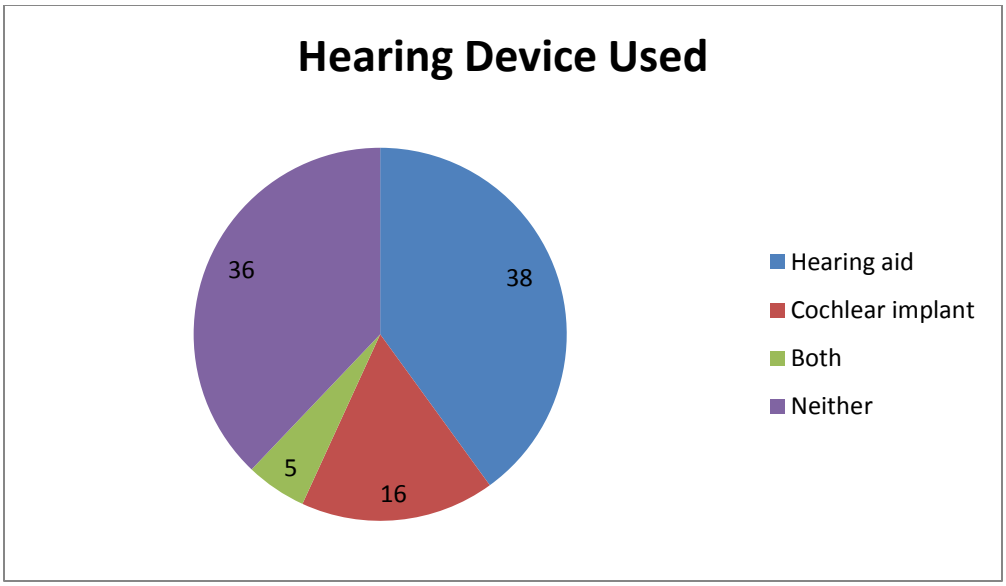
Thanks in advance,

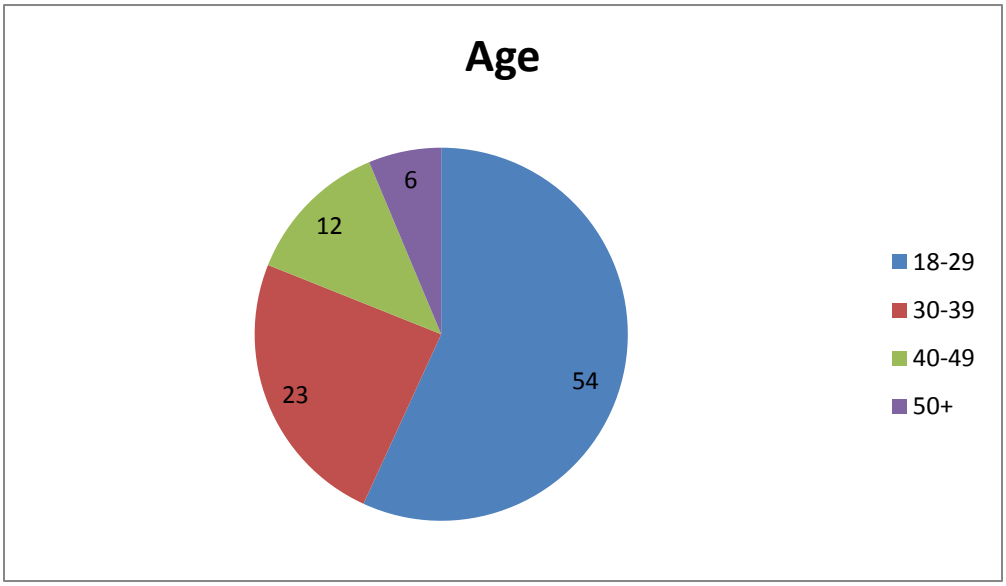
Brent Shiver
DePaul University
bshiver@cs.depaul.edu
DePaul IRB BS031313CDM

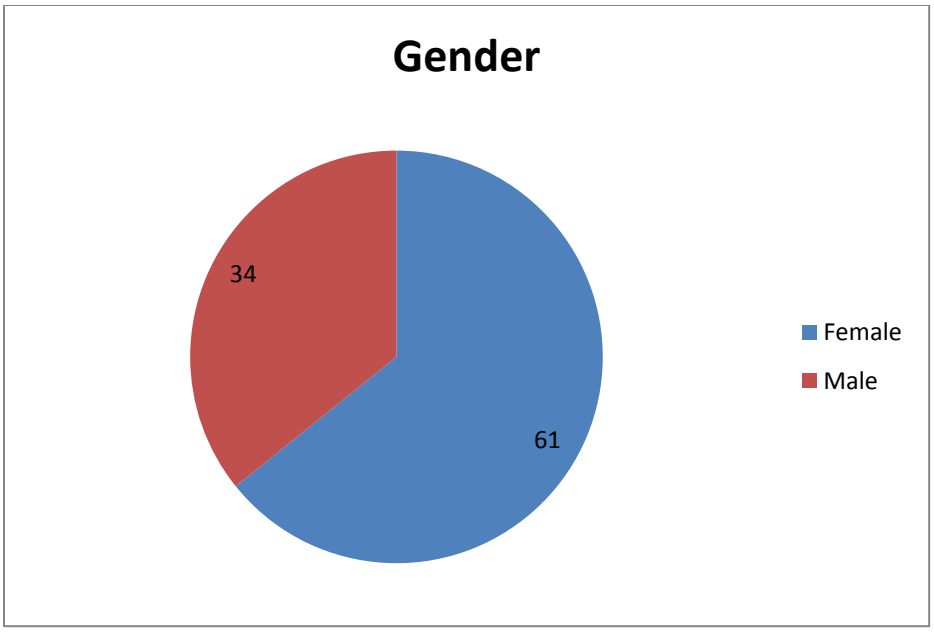
Appendix R – Demographic Background Charts (Study 2)

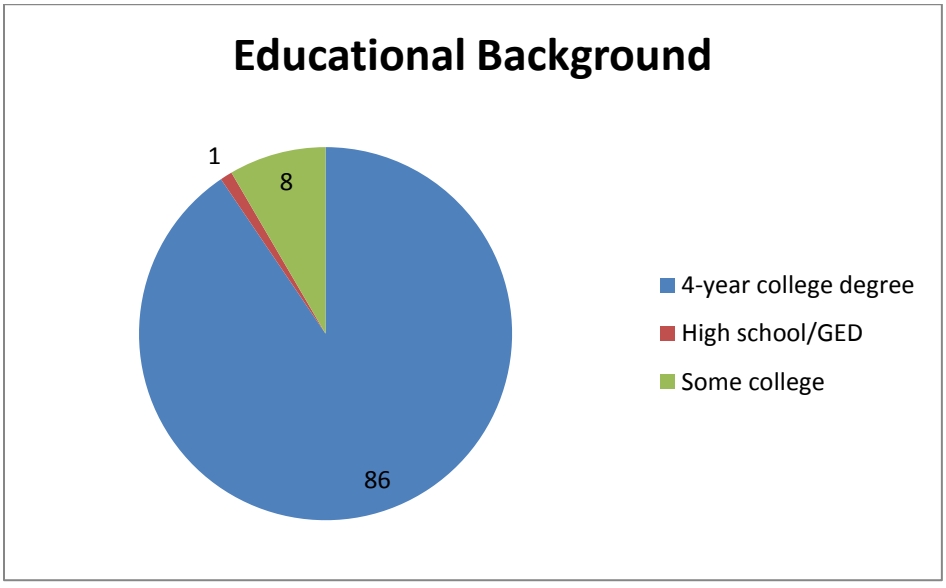


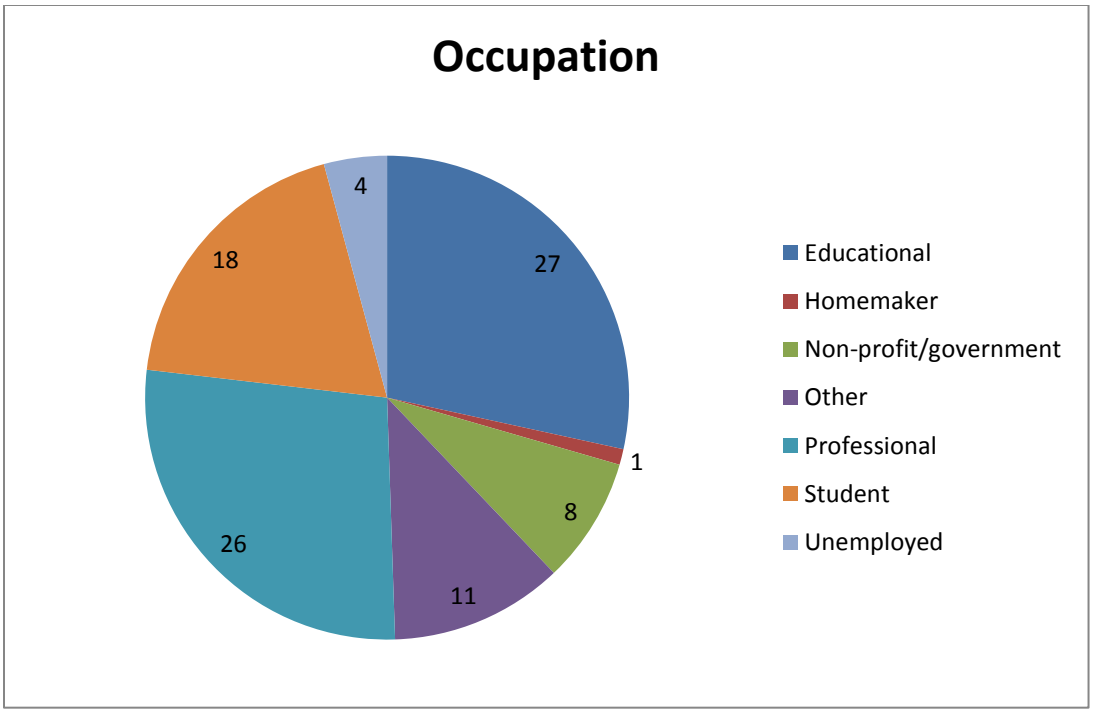


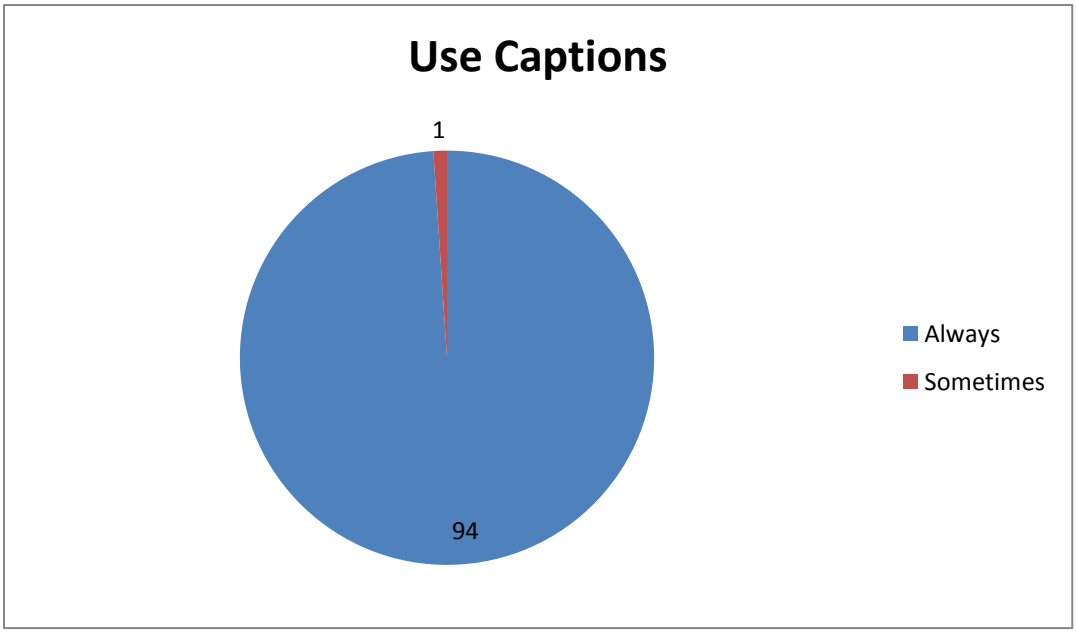

Hearing Device Used
36
38
5
16
Hearing aid
Cochlear implant
Both
Neither

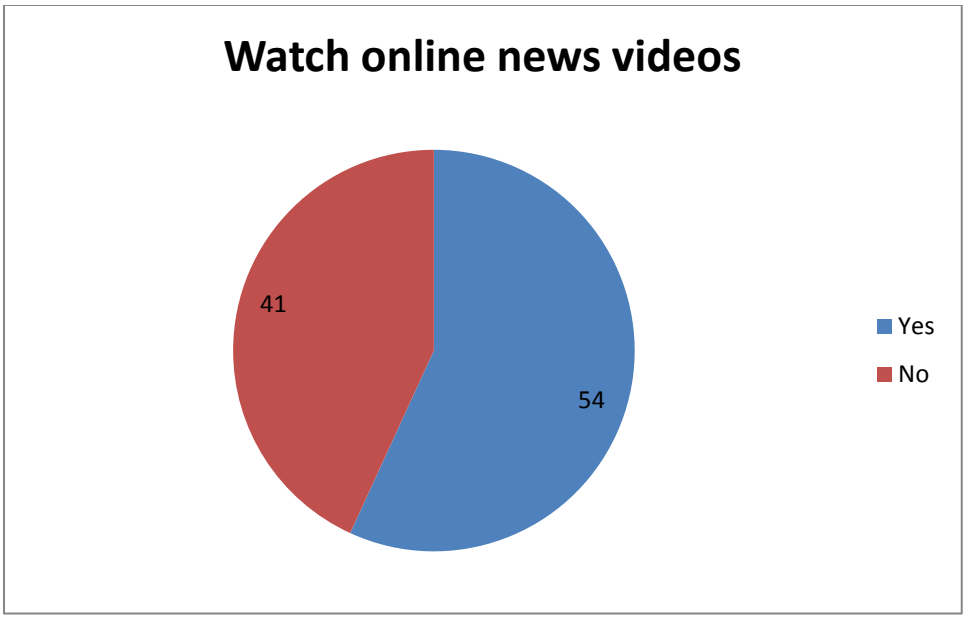

Age
6
12
54
23
18-29
30-39
40-49
50+

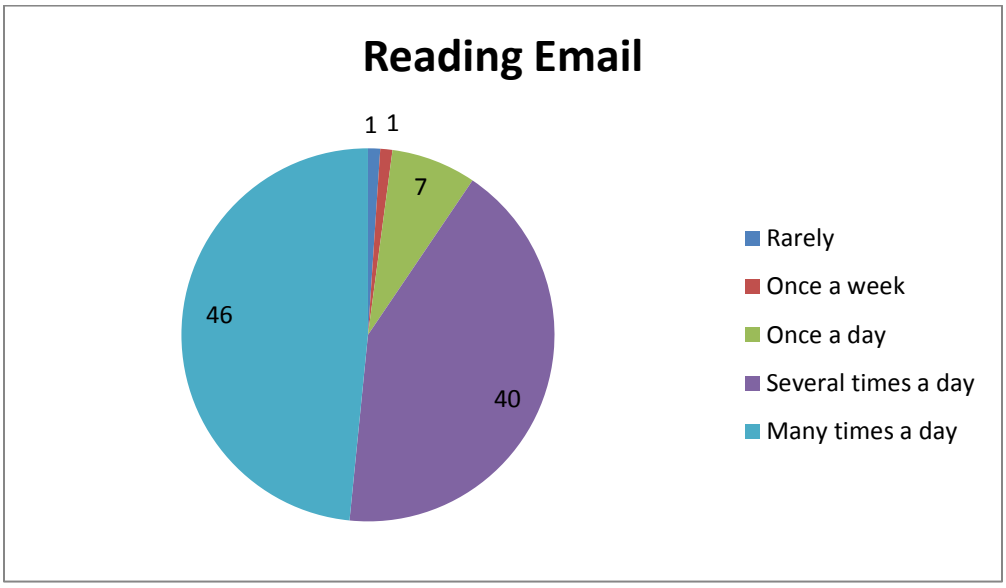

Gender
34
61
Female
Male

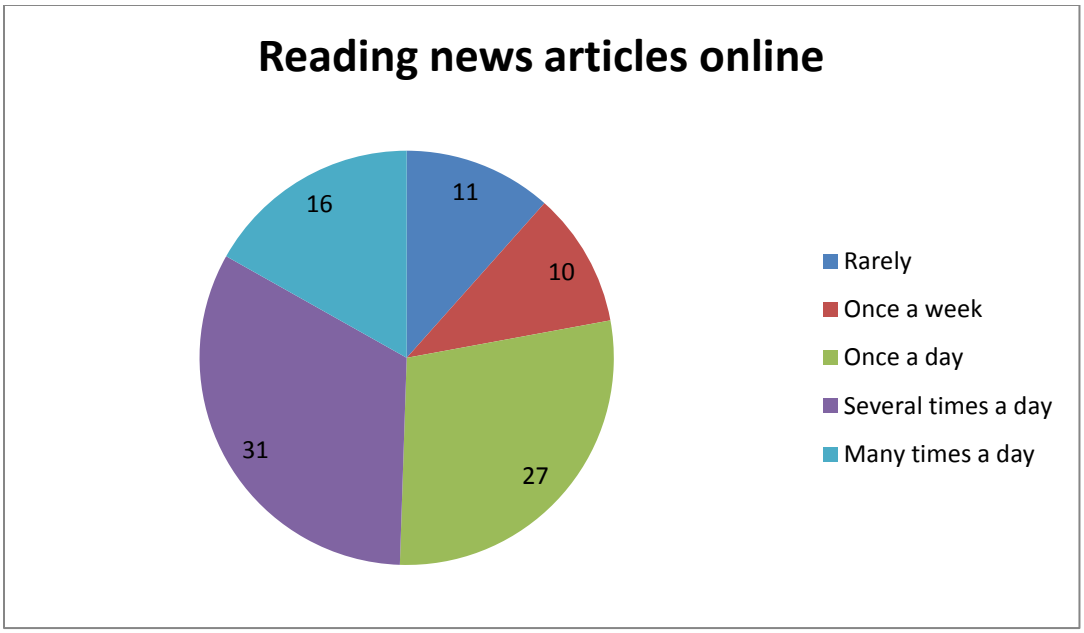

Educational Background
1
8
86
4-year college degree
High school/GED
Some college

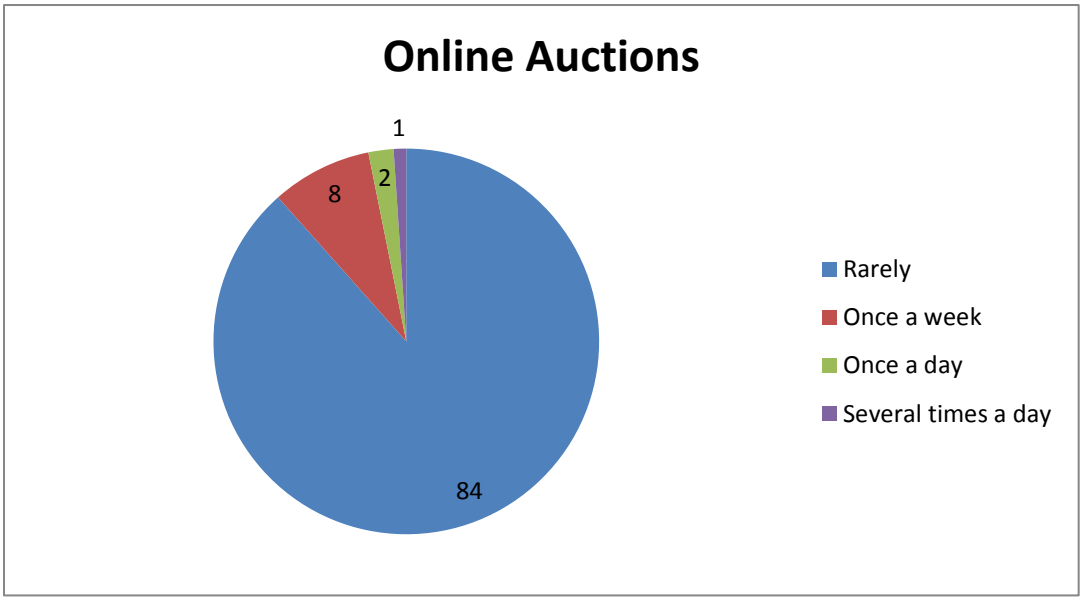

Occupation
4
18
27
1
8
26
11
Educational
Homemaker
Non-profit/government
Other
Professional
Student
Unemployed

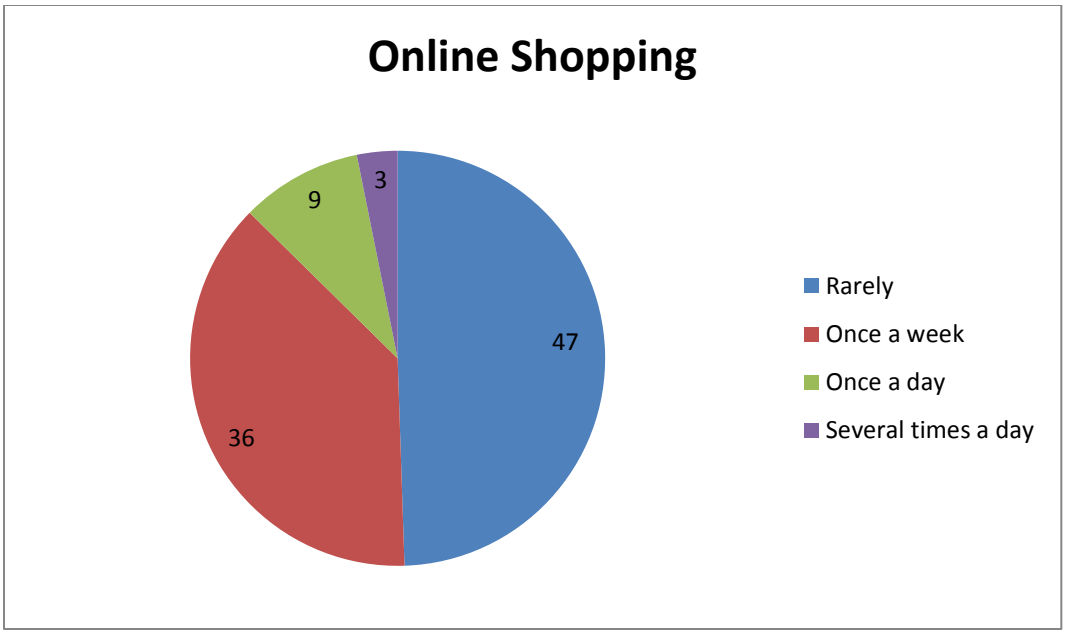

Use Captions
1
94
Always
Sometimes

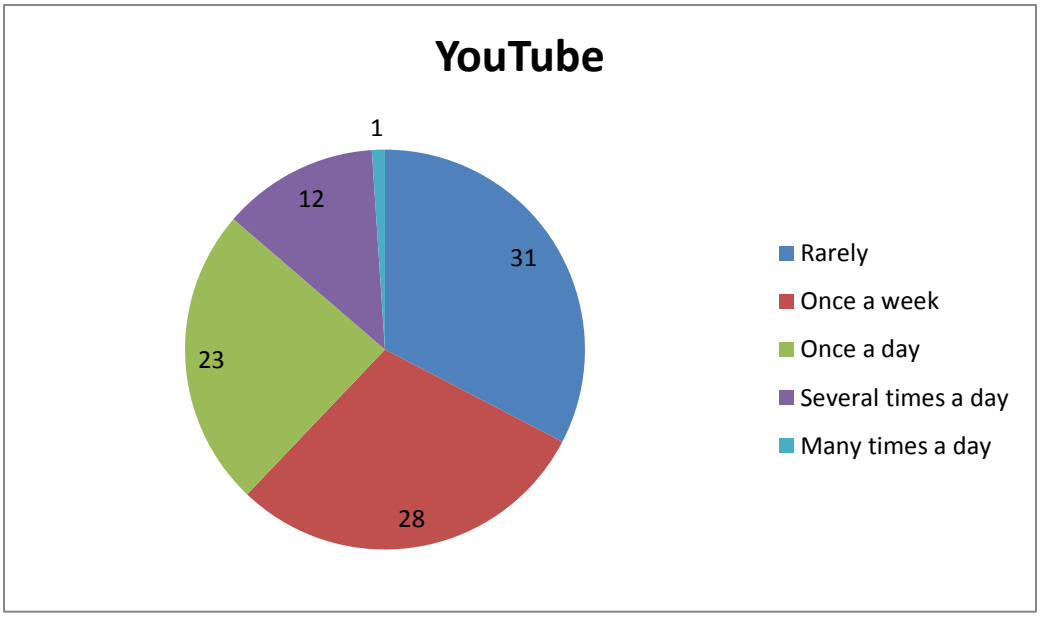

Watch online news videos
41
54
Yes
No

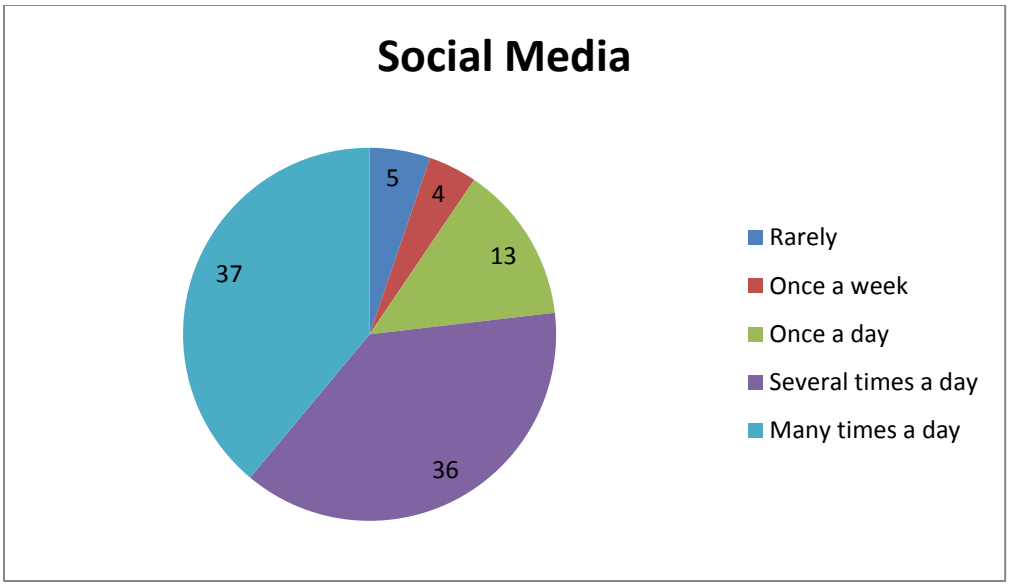

Reading Email
1 1
7
46
40
Rarely
Once a week
Once a day
Several times a day
Many times a day


Reading news articles online
11
10
27
31
16
Rarely
Once a week
Once a day
Several times a day
Many times a day


Online Auctions
1
2
8
84
Rarely
Once a week
Once a day
Several times a day


Online Shopping
47
36
9
3
Rarely
Once a week
Once a day
Several times a day


YouTube
31
28
23
12
1
Rarely
Once a week
Once a day
Several times a day
Many times a day


Social Media
5
4
13
37
36
Rarely
Once a week
Once a day
Several times a day
Many times a day

Appendix S – Single Factor ANOVA Results for Presentation Order

Accuracy: Story 1

Groups	*Count*	*Sum*	*Average*	*Variance*
1234	17	12.5	0.735294	0.089614
1324	16	10.75	0.671875	0.064323
2134	16	13	0.8125	0.029167
2314	15	10	0.666667	0.095238
3124	15	11.75	0.783333	0.06131
3214	16	10.5	0.65625	0.090625

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.345686	5	0.069137	0.963364	0.444692	2.316858
Within Groups	6.387209	89	0.071766			
Total	6.732895	94				

Accuracy: Story 2

Groups	*Count*	*Sum*	*Average*	*Variance*
1234	17	13.5	0.794118	0.040033
1324	16	11.33333	0.708333	0.046296
2134	16	11.5	0.71875	0.058218
2314	15	10.16667	0.677778	0.053439
3124	15	9.5	0.633333	0.084127
3214	16	11	0.6875	0.055093

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.231267	5	0.046253	0.829858	0.531829	2.316858
Within Groups	4.960546	89	0.055736			
Total	5.191813	94				

Accuracy: Story 3

Groups	*Count*	*Sum*	*Average*	*Variance*
1234	17	10	0.588235	0.09329
1324	16	8.25	0.515625	0.095573
2134	16	8	0.5	0.091667
2314	15	8	0.533333	0.06131
3124	15	9.5	0.633333	0.079167
3214	16	9.25	0.578125	0.097656

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.198828	5	0.039766	0.457681	0.806673	2.316858
Within Groups	7.732751	89	0.086885			
Total	7.931579	94				

Accuracy: Story 4

Groups	*Count*	*Sum*	*Average*	*Variance*
1234	17	11.16667	0.656863	0.060662
1324	16	10.16667	0.635417	0.045255
2134	16	10.33333	0.645833	0.036574
2314	15	8	0.533333	0.048413
3124	15	10.5	0.7	0.036508
3214	16	9.333333	0.583333	0.062963

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.26397	5	0.052794	1.084802	0.374348	2.316858
Within Groups	4.331352	89	0.048667			
Total	4.595322	94				

Time Required: Story 1

Groups	*Count*	*Sum*	*Average*	*Variance*
1234	17	388.039	22.82582	545.7844
1324	16	489.6075	30.60047	758.2317
2134	16	264.598	16.53738	284.3156
2314	15	253.668	16.9112	248.6166
3124	15	371.9358	24.79572	1583.782
3214	16	342.088	21.3805	495.1414

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	2180.163	5	436.0325	0.675473	0.643114	2.316858
Within Groups	57451.47	89	645.5221			
Total	59631.63	94				

Time Required: Story 2

Groups	*Count*	*Sum*	*Average*	*Variance*
1234	17	442.6155	26.03621	1069.433
1324	16	356.5807	22.28629	1277.823
2134	16	308.7637	19.29773	111.1579
2314	15	237.6005	15.84003	36.10082
3124	15	300.276	20.0184	168.6724
3214	16	446.4858	27.90536	460.6002

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1581.089	5	316.2179	0.589743	0.707792	2.316858
Within Groups	47721.47	89	536.1962			
Total	49302.55	94				

Time Required: Story 3

Groups	*Count*	*Sum*	*Average*	*Variance*
1234	17	976.5063	57.44154	29103.63
1324	16	221.832	13.8645	264.4774
2134	16	187.4746	11.71716	40.0821
2314	15	555.5572	37.03715	10208.6
3124	15	296.2458	19.74972	417.5071
3214	16	329.943	20.62144	1223.321

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	25201.97	5	5040.393	0.703853	0.622015	2.316858
Within Groups	637341.8	89	7161.144			
Total	662543.8	94				

Time Required: Story 4

Groups	*Count*	*Sum*	*Average*	*Variance*
1234	17	335.2577	19.72104	255.7367
1324	16	492.3915	30.77447	2602.631
2134	16	397.76	24.86	602.6678
2314	15	267.7075	17.84717	94.11437
3124	15	347.7325	23.18217	516.0712
3214	16	352.7344	22.0459	545.7079

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1617.771	5	323.5542	0.417947	0.835128	2.316858
Within Groups	68899.49	89	774.1515			
Total	70517.26	94				

Number of Rewinds: Story 1

Groups	*Count*	*Sum*	*Average*	*Variance*
1234	17	5	0.294118	0.970588
1324	16	2	0.125	0.25
2134	16	0	0	0
2314	15	0	0	0
3124	15	2	0.133333	0.12381
3214	16	2	0.125	0.25

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.963571	5	0.192714	0.692636	0.630328	2.316858
Within Groups	24.76275	89	0.278233			
Total	25.72632	94				

Number of Rewinds: Story 2

Groups	*Count*	*Sum*	*Average*	*Variance*
1234	17	0	0	0
1324	16	1	0.0625	0.0625
2134	16	14	0.875	12.25
2314	15	7	0.466667	3.266667
3124	15	0	0	0
3214	16	0	0	0

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	10.48443	5	2.096886	0.809922	0.545643	2.316858
Within Groups	230.4208	89	2.588998			
Total	240.9053	94				

Number of Rewinds: Story 3

Groups	*Count*	*Sum*	*Average*	*Variance*
1234	17	1	0.058824	0.058824
1324	16	0	0	0
2134	16	0	0	0
2314	15	0	0	0
3124	15	6	0.4	0.828571
3214	16	6	0.375	0.783333

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	2.92987616	5	0.585975	2.146944	0.067086	2.316858
Within Groups	24.2911765	89	0.272935			
Total	27.2210526	94				

Number of Rewinds: Story 4

Groups	*Count*	*Sum*	*Average*	*Variance*
1234	17	7	0.411765	1.507353
1324	16	1	0.0625	0.0625
2134	16	1	0.0625	0.0625
2314	15	0	0	0
3124	15	9	0.6	2.114286
3214	16	2	0.125	0.25

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	4.44682663	5	0.889365	1.333839	0.257313	2.316858
Within Groups	59.3426471	89	0.666771			
Total	63.7894737	94				

Appendix T – Captioning Style Analysis

significant at p < 0.01

Accuracy

Mean	*n*	*Std. Dev*	
0.7105	95	0.26376	Story1
0.6877	95	0.23349	Story2
0.5579	95	0.28818	Story3
0.6158	95	0.21481	Story4
0.6430	380	0.25780	Total

ANOVA table

Source	*SS*	*df*	*MS*	*F*	*p-value*
Treatment	1.38158	3	0.460526	7.27	.0001
Error	23.80804	376	0.063319		
Total	25.18962	379			

Post hoc analysis

p-values for pairwise t-tests

		Story3 0.5579	Story4 0.6158	Story2 0.6877	Story1 0.7105
Story3	0.5579				
Story4	0.6158	.1137			
Story2	0.6877	.0004	.0496		
Story1	0.7105	3.62E-05	.0098	.5326	

significant at p < 0.01
significant at p < 0.05
not significant

Tukey simultaneous comparison t-values (d.f. = 376)

		Story3 0.5579	Story4 0.6158	Story2 0.6877	Story1 0.7105
Story3	0.5579				
Story4	0.6158	1.59			
Story2	0.6877	3.56	1.97		
Story1	0.7105	4.18	2.59	0.62	

critical values for experimentwise error rate:

0.05	2.60
0.01	3.18

Time Required

Mean	*n*	*Std. Dev*	
22.2099	95	25.18688	Avg S1
21.4410	95	22.81805	Avg S2
24.9364	95	82.90679	Avg S3
22.2041	95	27.40773	Avg S4
22.6978	380	46.68339	Total

Source	*SS*	*df*	*MS*	*F*	*p-value*
Treatment	671.90743	3	223.969144	0.10	.9588
Error	825,297.60880	376	2,194.940449		
Total	825,969.51623	379			

Number of Rewinds

Mean	*n*	*Std. Dev*	
0.1	95	0.52	Group 1
0.2	95	1.60	Group 2
0.1	95	0.54	Group 3
0.2	95	0.82	Group 4
0.2	380	0.97	Total

Source	*SS*	*df*	*MS*	*F*	*p-value*
Treatment	0.89	3	0.298	0.31	.8156
Error	357.64	376	0.951		
Total	358.54	379			

Appendix U – Kruskal-Wallis/Mann-Whitney Analysis on Captioning Preference

significant at p < 0.01
significant at p < 0.05
not significant
p' (with Bonferroni adjustment)

1. The captioning was easy to read.

Median	*n*	*Avg. Rank*	
2.50	94	214.40	Story1P1
3.00	94	258.54	Story2P1
0.00	94	78.48	Story3P1
2.00	94	202.57	Story4P1
2.00	376		Total

152.124	H (corrected for ties)
3	d.f.
9.17E-33	p-value

multiple comparison values for avg. ranks
41.83 (.05) 49.84 (.01)

Captioning Style	*1*	*2*	*3*
2	z=-3.2341 p=0.00124 p'=0.00744		
3	z=8.8875 p=0 p'=0	z=10.5884 p=0 P'=0	
4	z=0.8712 p=0.19215 p'=1.15290	z=3.8225 p=0.00007 p'=0.00042	z=-8.2415 p=0 p'=0

2. The captioning made it easy to understand the story.

	Median	*n*	*Avg. Rank*	
	3.00	95	215.11	Story1P2
	3.00	95	251.96	Story2P2
	0.00	95	83.76	Story3P2
	2.00	95	211.16	Story4P2
	2.00	380		Total

135.612	H (corrected for ties)
3	d.f.
3.34E-29	p-value

multiple comparison values for avg. ranks

42.05 (.05) 50.11 (.01)

Captioning Style	*1*	*2*	*3*
2	z=-2.5712 p=0.00508 p'=0.03048		
3	z=8.5051 p=0 p'=0	z=10.0723 p=0 P'=0	
4	z=0.2335 p=0.40905 p'=2.4543	z=2.7585 p=0.00289 p'=0.01734	z=-8.1726 p=0 p'=0

3. I have confidence in the accuracy of the captioning.

Median	*n*	*Avg. Rank*	
2.00	95	213.41	Story1P3
2.00	95	234.53	Story2P3
0.00	95	103.49	Story3P3
2.00	95	210.57	Story4P3
1.00	380		Total

87.596	H (corrected for ties)
3	d.f.
7.19E-19	p-value

multiple comparison values for avg. ranks

42.05 (.05) 50.11 (.01)

Captioning Style	*1*	*2*	*3*
2	z=-1.5198 p=0.06426 p'=0.38556		
3	z=7.3085 p=0 p'=0	z=8.1265 p=0 P'=0	
4	z=-0.0449 p=0.48405 p'=2.9043	z=1.3852 p=0.08226 p'=0.49356	z=-6.3706 p=0 p'=0

4. From the captions, I feel that I fully understood the story.

Median	*n*	*Avg. Rank*	
2.00	95	215.35	Story1P4
2.00	95	228.48	Story2P4
0.00	95	91.66	Story3P4
2.00	95	226.51	Story4P4
1.00	380		Total

109.596	H (corrected for ties)
3	d.f.
1.34E-23	p-value

multiple comparison values for avg. ranks
42.05 (.05) 50.11 (.01)

Captioning Style	*1*	*2*	*3*
2	z=-0.8192 p=0.20611 p'=1.23666		
3	z=7.889 p=0 p'=0	z=8.7439 p=0 P'=0	
4	z=-0.839 p=0.20045 p'=1.2027	z=-0.0435 p=0.48405 p'=2.9043	z=-8.1383 p=0 p'=0

5. I like this style of captioning.

	Median	*n*	*Avg. Rank*	
	1.00	95	194.66	Story1P5
	3.00	95	268.85	Story2P5
	0.00	95	104.61	Story3P5
	1.00	95	193.88	Story4P5
	1.00	380		Total

113.948	H (corrected for ties)
3	d.f.
1.55E-24	p-value

multiple comparison values for avg. ranks

42.05 (.05) 50.11 (.01)

Captioning Style	*1*	*2*	*3*
2	z=-5.1252 p=0 p'=0		
3	z=-6.1107 p=0 p'=0	z=9.7649 p=0 P'=0	
4	z=0.0567 p=0.47608 p'=2.85648	z=4.7453 p=0 p'=0	z=-5.6503 p=0 p'=0

6. Compared to captions on TV, I prefer this caption style.

Median	*n*	*Avg. Rank*	
1.00	95	187.04	Story1P6
2.00	95	250.73	Story2P6
0.00	95	122.47	Story3P6
1.00	95	201.77	Story4P6
1.00	380		Total

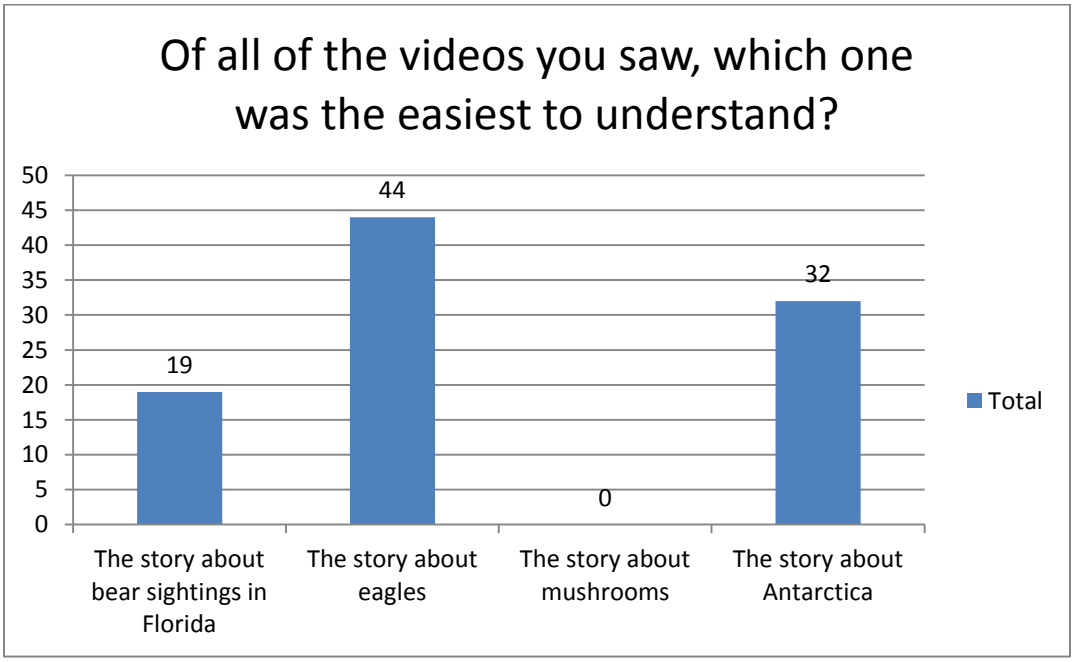
72.717	H (corrected for ties)
3	d.f.
1.12E-15	p-value

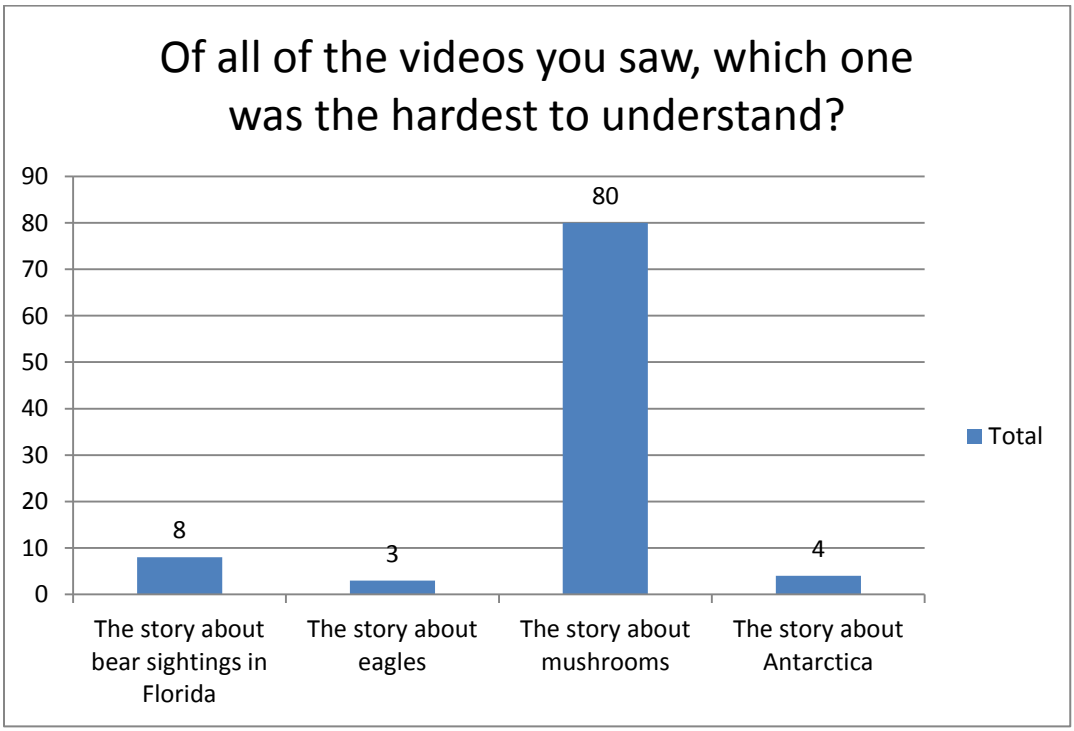
multiple comparison values for avg. ranks

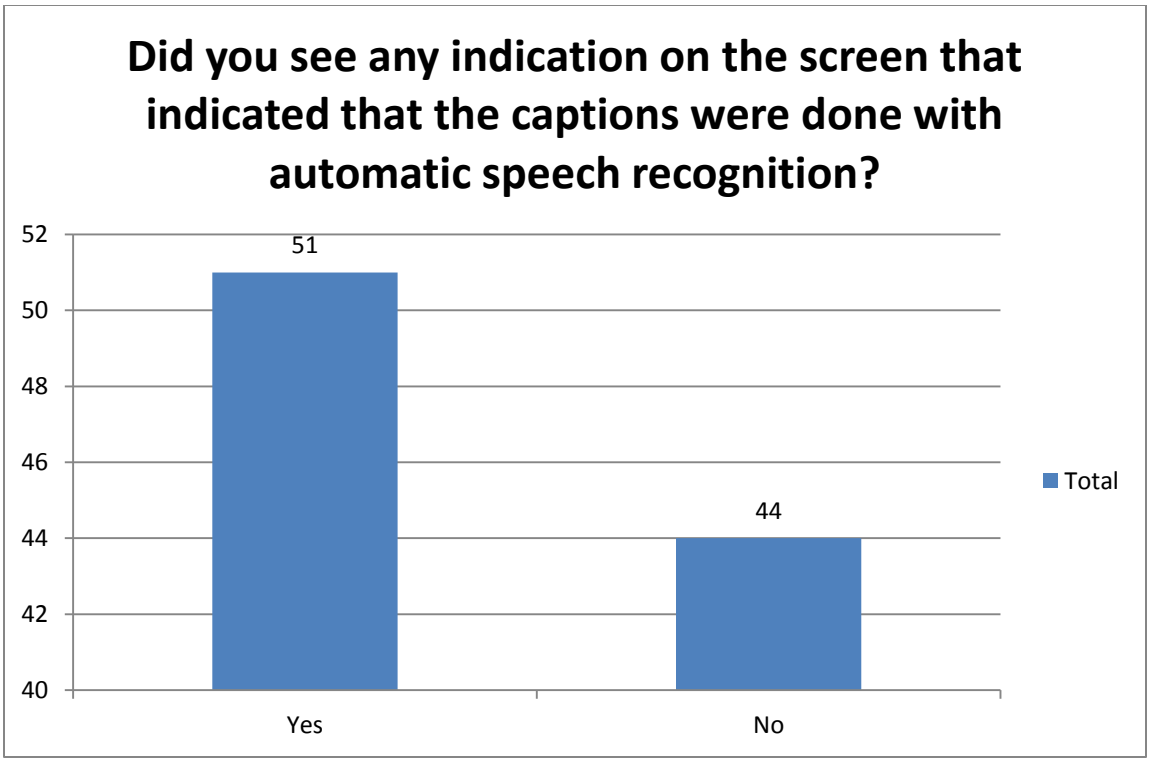
42.05 (.05) 50.11 (.01)

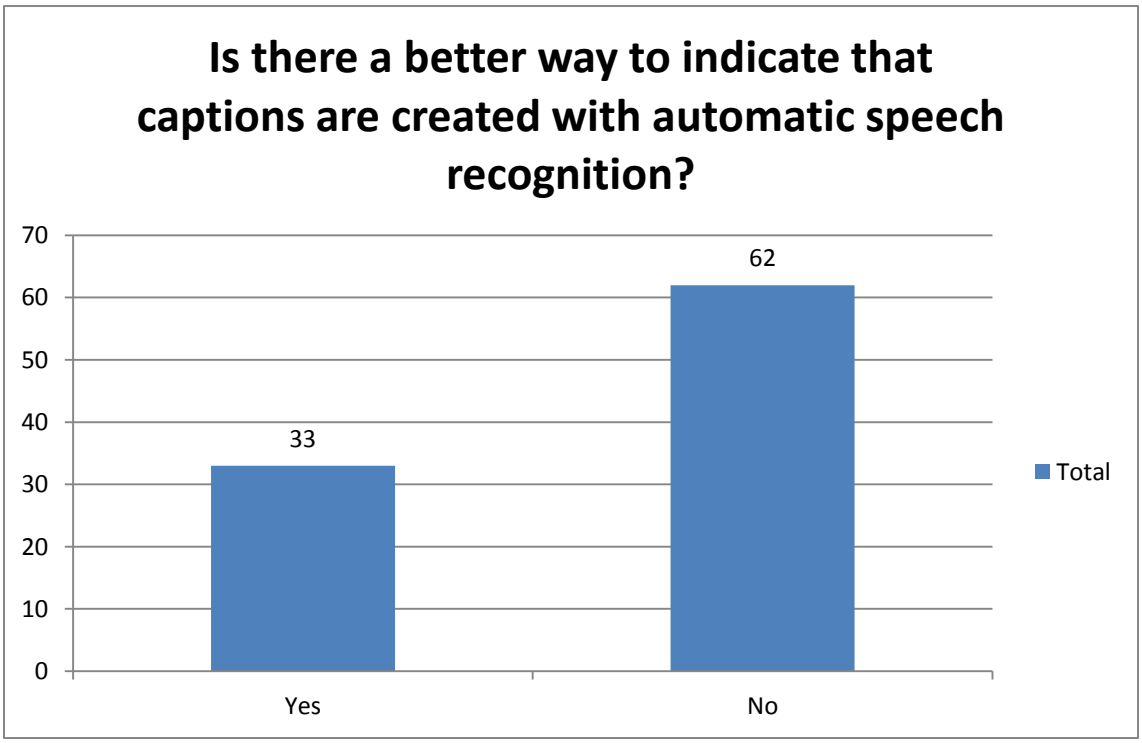
Captioning Style	*1*	*2*	*3*
2	z=-4.3046 p=0 p'=0		
3	z=4.4062 p=0 p'=0	z=7.7676 p=0 P'=0	
4	z=-0.9683 p=0.16602 p'=0.99612	z=3.0197 p=0.00126 p'=0.00756	z=-4.8746 p=0 p'=0

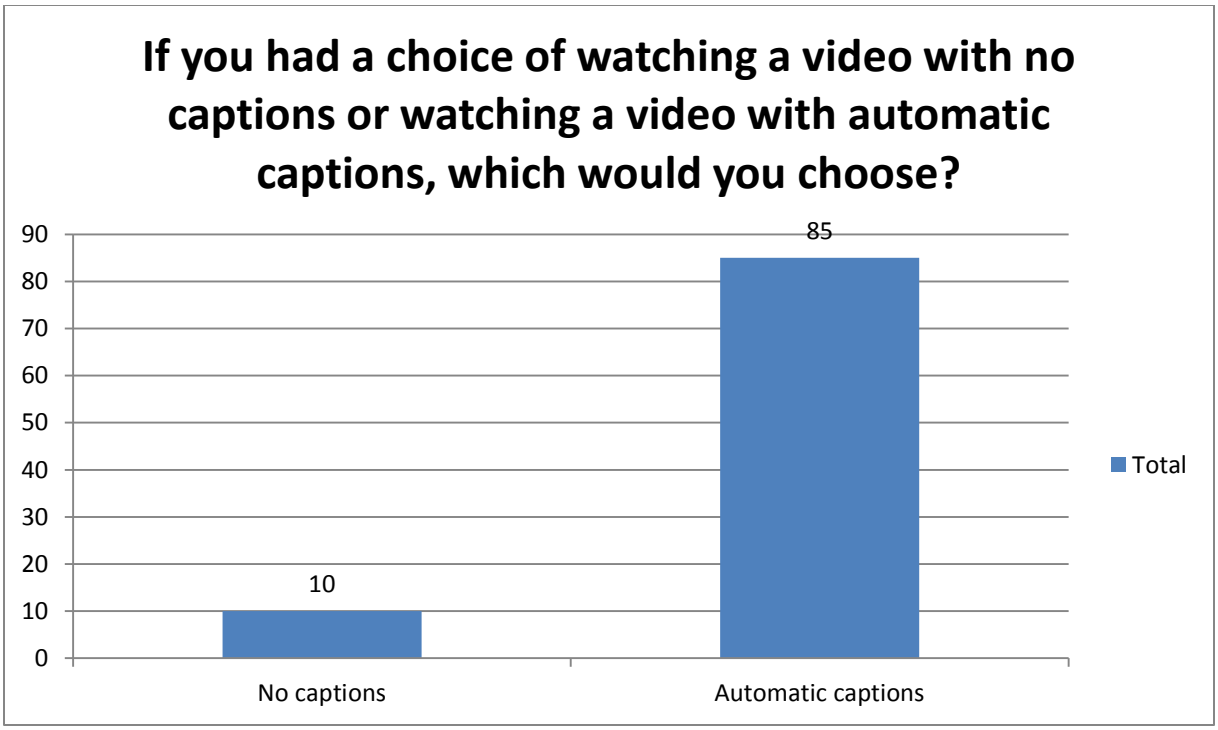
Appendix V – Post Test Responses (Study 2)

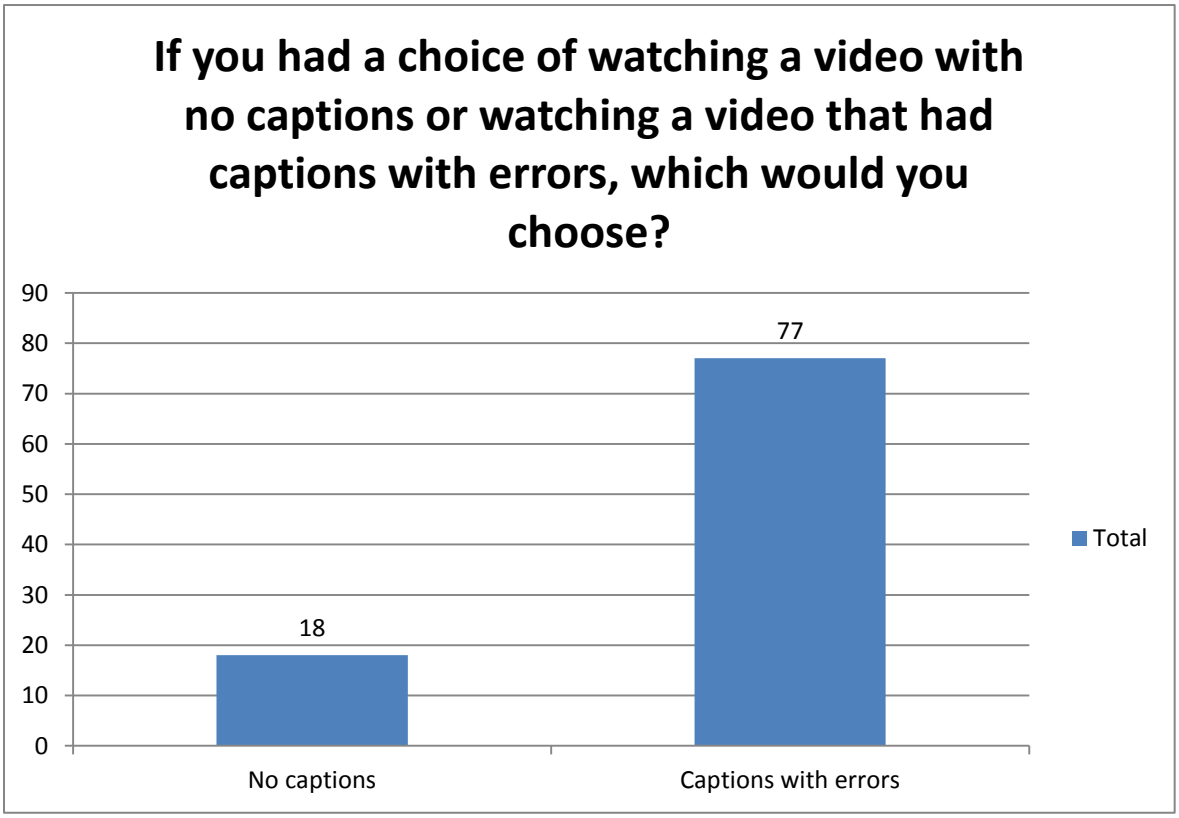


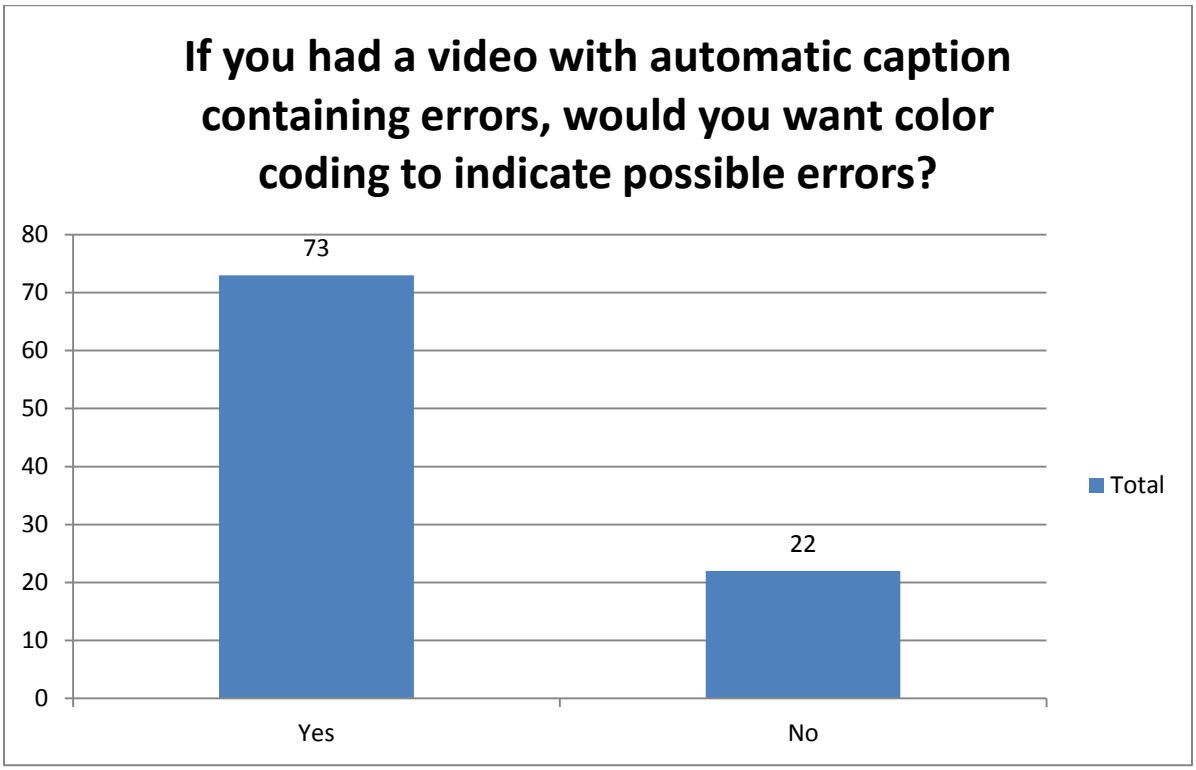


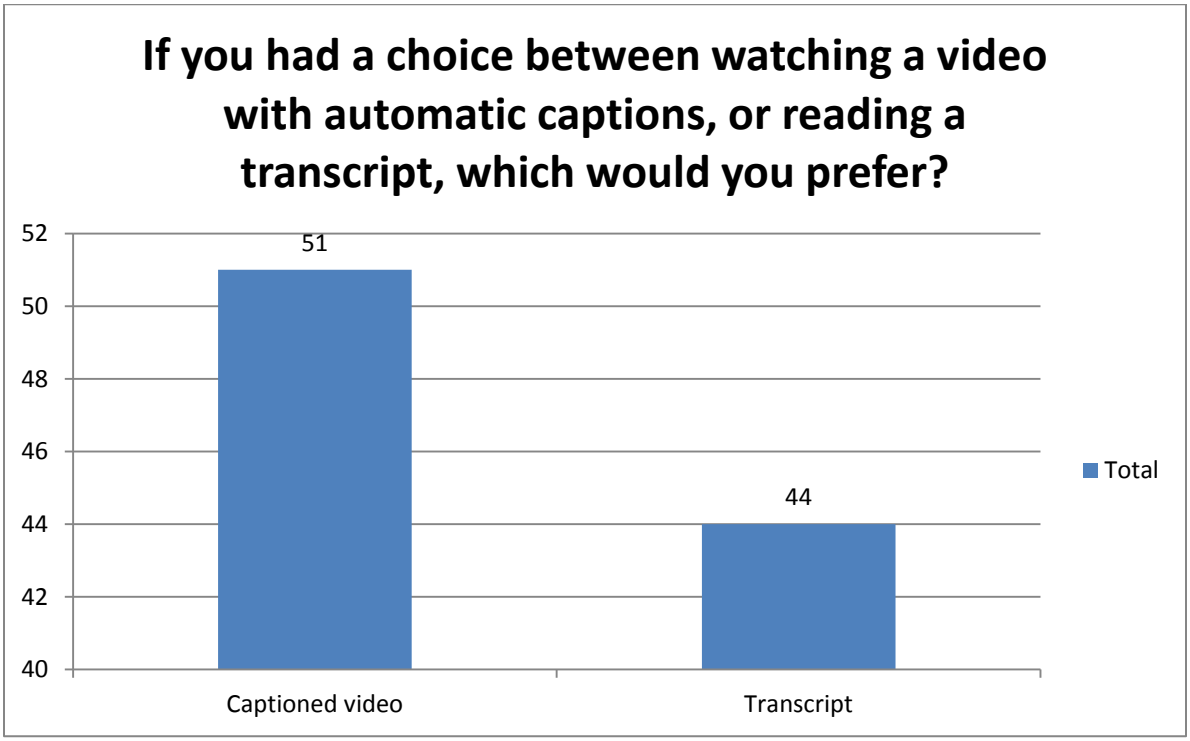

Did you see any indication on the screen that indicated that the captions were done with automatic speech recognition?
52
50
48
46
44
42
40
51
44
Yes
No
Total


Is there a better way to indicate that captions are created with automatic speech recognition?
70
60
50
40
30
20
10
0
33
62
Yes
No
Total


If you had a choice of watching a video with no captions or watching a video with automatic captions, which would you choose?
90
80
70
60
50
40
30
20
10
0
85
10
No captions
Automatic captions
Total


If you had a choice of watching a video with no captions or watching a video that had captions with errors, which would you choose?
90
80
70
60
50
40
30
20
10
0
77
18
No captions
Captions with errors
Total


If you had a video with automatic caption containing errors, would you want color coding to indicate possible errors?
80
70
60
50
40
30
20
10
0
73
22
Yes
No
Total


If you had a choice between watching a video with automatic captions, or reading a transcript, which would you prefer?
52
50
48
46
44
42
40
51
44
Captioned video
Transcript
Total

BRENT SHIVER

2802 Forest Park Dr, Auburn, MA 01501 · 312.957.8950 · bshiver@cs.depaul.edu

EDUCATION

DePaul University Chicago, IL
Ph.D. in Computer Science December 2013
Emphasis in Database Systems, Human-Computer Interaction, Software Management

M.S. in Computer Science June 2004
Concentrations in Database Systems, Human-Computer Interaction

Gallaudet University Washington, DC
B.S. in Computer Science May 2000
summa cum laude, Business Administration minor

Northwestern University Evanston, IL
B.A. in Psychology June 1996
Full wrestling athletic scholarship recipient

INTERESTS

Deaf accessibility and accommodation technologies, human-computer interaction, software management, web technologies, software development, database systems.

EXPERIENCE

Allston Trading, LLC Chicago, IL
Software Developer, Technology Services 12/2012 – 3/2013

- Led an effort to upgrade the position and profit & loss (P&L) system that involved redesigning the architecture and protocol to support mandates including enhanced security and bandwidth reduction.
- Assisted with operational support that involved handling software and data issues, requests from compliance and audit, position and P&L concerns, and trade reconciliation.

Software Developer, Trade Infrastructure Development 12/2010 – 12/2012

- Created Java Swing GUI desktop application that displayed position and P&L data. It allowed traders and managers to create customized and extensible views using Jython that matched their objectives. API support was also incorporated so developers can create separate monitoring and reporting tools.
- Additional responsibilities included designing, developing, testing, deploying, and debugging software features that support high-frequency, low-latency trading and research infrastructure including quote feeds, metadata support, trade simulation, monitoring tools, trade reconciliation, logging and reporting.

Software Developer, Back Office Technology 9/2007 – 12/2010

- Designed, developed, and supported a comprehensive, mission critical system that disseminated real-time and historical P&L and position data to the entire firm. This system evolved into an indispensable tool for traders and managers since it allowed for trading, trade reconciliation, and risk management with total confidence. It was written in Java that employed multi-threads, sockets, and databases.
- Promoted a firm-wide standardization of P&L tracking including fee calculations, currency rate conversions, contract valuation, and daily settlements and incorporated them into the P&L system.
- Delivered an automatic procedure using Java and SQL Server that handled position adjustments resulting from options exercises, assignments, and expirations.
- Designed and constructed a Java-based system that automatically tracked FX currency positions and value dates; it became an invaluable utility for the back office and FX trading desk.

Software Developer, Operations 8/2003 – 9/2007

- Created and developed a Java server application that functioned as an interface between trading gateways and databases; it stored critical information about orders and fills into SQL Server databases.
- Built an internal cross server in Java that acted as an exchange between trading clients; it saved money by not incurring commissions and fees, while reducing risk by bypassing external exchanges.
- Assembled a database-driven, PHP-based Intranet system that provided news, incident reporting, human resource forms and information, announcements, daily trading reports, and change logs.
- Served as firm's impromptu DBA during first year until a full-time position was created and filled.
- Developed and deployed company's first public website.

Foresight Survey Systems International, Inc Lake Zurich, IL

Programmer, Lead Retrieval and Association Services 2/2002 – 1/2003

- Designed and developed online registration and barcode-based badge production system using PHP and MySQL database for use at convention and trade shows.
- Enhanced company's existing lead retrieval software written in C to make compatible with MySQL database and added new time-saving and user-friendly features.
- Reconstructed company's public website with a new look and navigation enhancements.

Lucent Technologies Naperville, IL

Member of Technical Staff 1, Core Systems and Reliability 6/2000 – 12/2001

- Wrote shell scripts for automated software and hardware testing, and conducted tests on each load to reveal errors.
- Prepared and monitored testing statistics to check progression towards completion date, and investigated possible interruptions to the process.
- Contributed to analysis, design, documentation, and implementation of subsystem features in C that managed fault recovery and maintenance functions for 5ESS Switch Module.
- Redesigned and developed department website as a summer intern.

RESEARCH PROJECTS

DePaul University Chicago, IL
Lab Member, American Sign Language Project 4/2003 – 8/2013

- Created facilitation techniques and instrument templates for use with deaf test participants in usability studies. They became an effective and reusable testing procedure that measures the legibility of ASL displayed on the synthesizer product.

REFEREED CONFERENCE PROCEEDINGS

- Schnepp, J., Wolfe, R., Shiver, B., McDonald, J., Toro, J. (2011, October 23). SignQUOTE: A Remote Testing Facility for Eliciting Signed Qualitative Feedback. Paper presented at Second International Workshop on Sign Language Translation and Avatar Technology (SLTAT), Dundee, Scotland, UK.
- Shiver, B. (2011, May 13). Utilizing automatic speech recognition to improve Deaf accessibility on the Web. Proceedings of the DePaul CDM School of Computing Research Symposium (SOCRS 2011) Chicago, IL. pp 4-8.
- Wolfe, R., Alba, N., Billups, S., Davidson, M.J., Dwyer, C., Jamrozik, D.G., Smallwood, L., Alkoby, K., Carhart, L., Hinkle, D., Hitt, A., Kirchman, B., Lancaster, G., McDonald, J., Semler, L., Schnepp, J., Shiver, B., Suh, A., Young, J. (2006, March 20-25). An Improved Tool for Practicing Fingerspelling Recognition. Paper presented at Technology and Persons with Disabilities Conference, California State University at Northridge, Los Angeles, CA.

POSTER PRESENTATIONS

- Shiver, B. (2013, May 31). Improving Deaf Accessibility to Web-based Multimedia [poster presentation]. DePaul CDM Research Symposium, Chicago, IL.
- Schnepp, J. & Shiver, B. (2011, October 24-26). Improving Deaf Accessibility in Remote Usability Testing. Poster presented at the 13th International ACM SIGACCESS Conference on Computers and Accessibility, Dundee, Scotland, UK. pp 255-256.
- Shiver, B. (2007, May 5). Towards Building a Tool to Remedy Literacy Challenges Facing Deaf Learners [poster presentation]. DePaul CTI Research Symposium, Chicago, IL.

GRANTS, FELLOWSHIPS, AND AWARDS

- DePaul University travel grant to present at ACM ASSETS conference in Dundee, Scotland, 2011
- DePaul College of Computing and Digital Media (CDM) Research Grant, 2008 & 2013
- Sun Certified Programmer for the Java Platform, Standard Edition 5.0, 2010
- DePaul CDM Ph.D. Tuition Waiver Award, 2007 – 2011
- District of Columbia NASA Space Grant Consortium Fellowship Award, 2001 – 2009
- U.S. Army Research Institute (ARI) Fellowship, The Consortium of Universities, 1999 – 2000
- Who's Who Among Students in American Universities & Colleges, 1999 & 2000
- Fannie Mae Scholarship for Outstanding Scholastic Achievement Award, 2000
- Co-Winner of NASA Space Grant Consortium Essay Competition, 1999
- National Academic Advisors for Athletes Association Achievement Award, 1996
- Three-time NCAA Division-I Wrestling Championships Qualifier, 1994 – 1996
- Four-year starter and letter winner, Northwestern University Wrestling, 1992 – 1996