
Summit to Create a Cyber-Community to Advance Deaf and Hard-of-Hearing
Individuals in STEM (DHH Cyber-Community)
June 25-27, 2008

<http://www.ntid.rit.edu/cat/summit>

June 11, 2008

Educational Captioners and Interpreters:

Section of Working Paper on Captioning

(Reflects input from group members regarding captioning)

Overview of Remote Captioning Services

In remote captioning, an intermediary operator in an office in a remote location produces text with (a) a stenographic machine, (b) a QWERTY keyboard, or (c) automatic speech recognition while listening to the information via an audio source as it is being spoken by a teacher. As used here, *captioning* is the umbrella term for speech-to-text services and *remote captioning* is the umbrella term for remote speech-to-text services. Often when a speech-to-text service is used, interpreting and/or notetaking services have *not* been provided to deaf/hard-of-hearing (d/hh) students unless there are several of these students who have diverse needs in a class (Marschark, et al.; Stinson & Antia, 1999). Generally the goal of *steno-based services*, which employs a 24-key steno-machine and is also called Communication Access Real-Time Translation (CART), is to produce a verbatim display of the spoken message (Preminger & Levitt, 1998; Steinfeld, 1998; Smith, 2003; Stinson, Stuckless, Henderson & Miller, 1988). This approach is based upon theory concerning phonetics and syllables.

Keyboard-based approaches often use computer-enhanced high-speed typing to provide access to the spoken information. The two most common of these approaches, C-Print and Typewell, use word abbreviations with a QWERTY-based keyboard to enable the service provider to, as closely as possible, capture the meaning of the spoken word (Harkins & Bakke, 2003; O'Neill & Laidler, 2004; Stinson, Elliot, McKee & Francis, 2000; Stinson & Stuckless, 1998). In the last few years, *automatic speech recognition-based* (ASR) services have begun to be used as a support for communication access and learning for students who are d/hh or who have other disabilities

(Aylesworth, 2006; Viable Technologies, 2005). Automatic speech recognition may be used to provide a word-for-word translation, in a manner similar to a stenographic-based system, or a to provide a meaning-for-meaning translation, in a manner similar to a keyboard-based system.

Needs

Captioning service providers typically provide a real-time text display on a laptop computer for the student or other client to view in class, meeting, etc. However, support service needs still go unanswered in a variety of education, work, and extracurricular settings. Following are needs that have been identified for remote captioning services:

- Software that can translate either steno or voice into text.
- Software that allows for displaying formulas with symbols, not just words
- If the speaker at the remote location, such as the teacher, wears a microphone, then s/he needs to repeat what others are saying so that the provider hears the discussion.
- Settings that involve multiple speakers can be challenging for remote providers with respect to hearing all the comments or relying on the person with the microphone to repeat the comments. Other challenges with this type of situation are in the identification of speakers. The provider is dependent on the speakers identifying themselves or relying on their judgment as to whether the speaker is a male or a female, which may be difficult to ascertain if the sound quality is poor. Another consideration is the student who is receiving the text, sees only “male speaker or female speaker”, and may not be in a position at the table to visually determine who exactly is speaking.
- The provider typically does not see the figures, charts, etc. that the teacher (or other speaker in the remote setting) uses in the classroom, and this may affect the provider’s comprehension of information, as well as the resultant display. There need to be improved ways for providers to prepare via textbooks, syllabus, PowerPoints, faculty and student names, etc.

- Instructors may use media to support their teaching and if this media is not captioned, the provider would need excellent sound and visual access to accurately represent the information in text for the student. Advanced viewing of this material is always helpful.
- If a web camera is used at the remote location, the provider can usually see references such as left, right, top, bottom with regard to comments made by the speaker (this, that, there) and refer the student to a specific location on the board or on the projection screen. The picture is not generally large enough or clear enough to offer additional visual cues for the provider that can support comprehension of the material.
- If the material is technical, the speaker may use jargon. If the provider is not aware of these technical words it may affect the production of accurate text.
- Students, teachers and other users need to be trained to set up and use technology and equipment with which they may be unfamiliar.
- Online training of students, faculty and staff on the distinction between word-for-word versus meaning-for-meaning services is desirable to ensure that the chosen service is appropriate for the particular classroom situation and student needs.
- Remote services need to have sufficient flexibility to effectively support individuals in many educational, work, and community settings. For example, community colleges in isolated settings that occasionally have d/hh students who need captioning services will benefit from services that are easy to install and use, work on a variety of display devices, and are inexpensive. Other settings, such as co-op experiences, often require meetings and professional development for which communication access is obviously crucial, but for which such access is often not available.
- Easy access by students or other users to the captioned material after the class or other event has been completed is desirable. It is desirable to have an electronic record of the transcript that the student can access after the visit. This transcript will help students recall what happened at the visit and further understand course material.
- Provision of captioning support services in nontraditional learning settings, such as laboratories or field visits where traditional delivery systems may be impractical is desirable. The absence of these supports

can preclude d/hh students from taking courses and participating in other activities with field visit components. Improvement of remote captioning would help overcoming this barrier. For example, students involved in medical or nursing field experiences will require mobility for remote captioning equipment and potentially nontraditional placement of the equipment. Pat Billies mentioned at the 2008 PEPNet conference the placement of a PDA in a cadaver next to the body part the instructor was discussing so the student could follow the lecture along with the visual information the instructors was providing within the cadaver.

- Disconnection and re-establishing of connections may be an issue. Disconnection may be more likely to occur if the connection is a wireless Internet connection, or through a cell phone.
- The student and provider must be prepared to resolve technical issues when they arise during the remote captioning process or have immediate access to a technician who can resolve the issues quickly.
- The lag time between the time a person, such as a teacher, speaks at a remote site and the display of text on the client's device may increase if cell phone lines are used.
- If speakers change and a polycom microphone is not feasible, the microphone may need to be changed from one speaker to another, such as on a field visit where the instructor gives the microphone to a guide at a place being visited. This could contribute delay and other possible difficulties.
- If a conference or USB microphone is used at the remote location, the provider can usually hear comments from several speakers. However, other sound issues can interfere with the sound quality such as chatting, coughing, rustling papers, or tapping near the placement of the microphone.
- If there is significant background noise in the setting, such as at a factory, this can affect the ability of the provider to hear the spoken message.
- Prep material is a necessary tool for supporting quality services. Experience in the topic area where services are needed or prep material is essential to support quality services with remote captioning.
- It may be desirable to have two-way communication between the speaker and the service provider, to obtain clarification, etc. This two-way communication may be with students via instant messaging or chat room
- D/hh students need a way to ask questions or make comments to the teacher or others at the remote site.

- A comprehensive clearinghouse, coordinating, or scheduling systems that allows d/hh users to easily arrange for their preferred service is desirable. Currently, a user contacts an individual company or agency ahead of time and that organization then arranges services. Thus, on demand, on-the-spot, service is not available.
- Captioning for online or distance learning is desirable.
- Software that allows ASCII text to flow through the Internet to a secure server is desirable.
- Improved Internet connections in the classroom that allow students to access a website to receive text is desirable.
- An encoder and a phone line at the remote site to receive text and merge it with an image. Sometimes a switcher is needed to switch between cameras and computer or other images is desirable.
- Improved was for service providers to hear what is being said via phone line or VoIP is desirable.
- A way for paying service providers is desirable.
- System for recruiting and training service providers is desirable.

Current Services

Remote CART, automatic speech recognition (ASR), and keyboard-based captioning services are currently available (Aylesworth, 2004; Preminger & Levitt, 1998; Rapidtext, 2006; Ultech, 2003). Features of current services can be subdivided into four categories: (a) features common to ASR, CART, and keyboard systems; (b) features unique to ASR systems; (c) features unique to CART systems; and (c) features unique to keyboard systems.

a. Features common to ASR/CART/keyboard systems.

- Current remote systems typically require the speaker, such as a teacher, to wear a microphone that communicates with a phone that relays the speech signal to a provider at a remote location. Some systems describe the voice connection as a conference call.

- Placement of a polycom microphone in the class or meeting room is another option. In addition, it is recommended that the provider use telephone headsets or a speaker phone in listening to the message from the class or meeting room (National Court Reporters Association, 2002).
- In some remote systems, the provider produces the text and sends it to the consumer's computer (usually a laptop), which has special software for viewing the text, via the Internet (Aylesworth, 2004). The consumer's computer usually connects to the provider via the Internet using a wired or wireless local area network connection. This system may involve a USB microphone connected to a laptop computer with the student that relays the speech signal to the provider through a free software program (Skype), which is loaded on the student laptop computer and on the provider desktop or laptop computer.
- In other remote systems, the provider continuously uploads the text on an Internet site that the students' laptop then accesses. This requires minimal special software on the students' computer. These websites may allow simultaneous access by multiple users.
- Some providers describe the website as similar to a "meeting room" on the Internet. This room, or site, receives the feed from the provider and allows viewing of the text by the consumer (Caption First, 2008). Software packages such as NetMeeting, which is a free download, and WebEx allow this type of meeting (National Court Reporters Association, 2002).
- Some remote systems use a classroom interface device, such as the ccSatellite box in Caption Mic, to facilitate transfer of the speaker signal to the provider.
- Preparation materials are typically gathered by providers. The service provider may be given access to the professor's listserv and PowerPoint/class outline repository.
- Some systems include downloading of an applet or other small software package onto the web browser of the client's computer. This procedure allow better viewing of the caption's on the Internet site that is accessed by the client's computer (Caption First, 2008).
- Some systems allow highlighting sections of text for note taking purposes.
- Edited text may be placed in a repository for access by students and faculty.

- The system may include a web camera so that the provider can see what is going on in the classroom (however, note limitations stated above). In addition, web cameras are one way that the provider and client may communicate with each other (Aylesworth, 2004).
- If the provider and client both share a software, the provider and client may communicate with each other on an independent channel in the software.
- There are two types of displays: a full screen of text or an image with text. These displays can appear on a myriad of screens, computers, or handheld devices. All word-for-word or meaning-for-meaning systems and technology can be displayed as a full screen of text without an image as well as text with an image.
- University Disability Service Coordinators determine the needs of the student and contracts with the appropriate service provider.
- The institution decides on the display method (either full screen of text, text with an image, or another option) to be delivered via Internet or phone line and encoder.
- Students are empowered to make personal choices about their screen display, such as font size and color. For full-screen display, a student can download the PowerPoint presentation from the professor and layer the flowing text underneath it or alongside.
- Universities are billed for services.

b. Features unique to ASR systems

- The provider listens to the audio and “shadows,” (other terms are “echoes,” and “re-speaks”) the words into the ASR software, which converts the dictation into text.
- Dictation macros can enable the provider to insert names, or phrases into the text without requiring the provider to say the full phrase. For example, the provider may say “CLC” for “coordinated list of

chemicals,” because the provider has entered CLC into the system’s dictionary as a macro for the expanded phrase (Caption Mic, 2008).

c. Features unique to CART systems

- Captions may be added to Webcasts. The captioning signal is added to the broadcast signal before being sent to the Web. The technology involved in this arrangement is similar to that used for regular broadcast captioning (Caption First, 2008).
- Some CART providers use text streaming. This technology regulates the flow of text so that it appears on the screen in a consistent, even manner. When the text box technology is combined with a Webcast, it enables the consumer to combine the text with other features, or pods that are part of the Webcast, such as audio, video, and chat functions (National Court Reporters Association, 2002).

Potential Benefits of a Multimedia Cyberinfrastructure

A multi-media cyber infrastructure refers to high bandwidth connections between institutions that allow for collaboration using software, hardware, and other technologies. This working paper is intended to provide some initial ideas for development of a multimedia cyber infrastructure that provides remote communication support for d/hh students in STEM mainstream classrooms. The following possible benefits of a cyberinfrastructure are intended for a wide variety of class situations, from a standard class, to a laboratory, to a field trip.

- Inter-institution connections for a class are an economical way to teach many students.
- Allow a person to create on-demand service if needed.
- Allow coverage for a variety of times.
- Allow access to the service from a variety of places.
- Support for access to group communication situations.
- Allow choice from among a variety of services.
- Allow easy access to the saved text version of the captioning after the event.
- Allow the combining of captions with other forms of input.
- All ways for the d/hh client to ask questions; make comments; etc.

- Maintain consistent, reliable service.
- Provide service on a variety of display devices from smart phones to desktop computers.
- Distance learning or remote students can use a chat or IM function or an audio connection to comment or ask questions. A designated person can read aloud all typed comments and questions.
- A cyberinfrastructure allows for easy access to the saved text as well as a time-coded on-demand replay of the text and video.
- Stored text allows for exams to be developed quickly and easily.
- Institutions can "share" laboratory exercises and expenses with the use of Webcam and time-coded text. Though not hands on for all students, the knowledge can still be shared.
- Institutions can share resources, including faculty, guest speakers, and tutors.
- Use of webcams used with high bandwidth inter-institution Internet connections allows for completed information exchanges of words and formulas written on the board in class as well as facial expressions of the professor and students.

Associated challenges to creation of the cyberinfrastructure are as follows:

- Creation of technology that would coordinate the various options to permit choice. and that would allow the various options to work together.
- Coordination of agencies that provide services to allow near continuous availability of services.
- Creation of a scheduling system that would make on-demand or near on-demand services feasible.
- Creation of storage access to allow use of the saved text produced with the captions as appropriate.
- Storage of audio/video/text files
- Slowdown or shutdown of Internet
- Developing retrieval procedures and policies for specific files or classes.
- Collaboration between private (for profit) agencies and public (non profit) educational institutions
- Creation of policies that are appropriate for all clients and all environments.
- Cooperation between agencies/institutions with instructors who may not fully understand the remote process or appreciate the importance of sharing educational material with the provider ahead of time. It

could be challenging for the instructor to receive a request for prep material from a provider who is not associated with the institution at which the class is being taught.

- If one institution experienced campus wide technical issues, would other institutions experience similar issues solely based on their technical connection for the remote captioning?

Educational Captioners and Interpreters:

Section of Working Paper on Interpreters

Overview of Educational Interpreting

The presence of interpreters in mainstream classrooms began in earnest in 1975. Public Law 94-142, subsequently the Individuals with Disabilities Education Act (IDEA) promulgated the practice of educating deaf children in the "least restrictive environment", i.e., mainstream classrooms. These regulations led to a significant increase in the number of deaf children attending public schools (Moores, 1987) and a concomitant surge in the need for interpreters in those schools.

Hurwitz (1991) was one of many researchers who noted that formal training for the task of educational interpreting was lacking, a condition which persists to this day. Jones (2004) provides a useful set of terms and definitions that will clarify our topic.

K-12 Educational Sign Language Interpreter:

“Educational Interpreter” means a person who uses sign language in the public school setting for purposes of facilitating communication between users and nonusers of sign language and who is fluent in the languages used by both deaf and nondeaf persons (CO 2002, 22-20-116 (2), in CDE 2002).

[An educational sign language interpreter] ...is a professional, who facilitates communication and understanding among deaf and hearing persons in a mainstream environment. The interpreter is a member of the educational team and is present to serve staff as well as students, hearing as well as deaf people, by minimizing linguistic, cultural, and physical barriers. The title, “Educational Interpreter,” is recommended by the National Task Force on Educational Interpreting, and is intended to imply that a person holding this title is a professional with specialized preparation in deafness, whose primary role is interpreting, but who is also qualified to provide certain other educational services (New York 1998).

Interpreting:

“...the process of changing messages produced in one language immediately into another language. The languages in question may be spoken or signed, but the defining characteristic is the live and immediate transmission” (Frishberg 1990, 18).

“Interpreting...refers either to the general process of changing the form of a message to another form, or to the specific process of changing an English message to American Sign Language (ASL), or vice versa” (Winston 1989, 147).

Research shows (Jones, Clark & Soltz 1997), however, the term “interpreting” in the K-12 arena refers to transliterating (between two codes of English: one spoken, one signed).

Transliterating:

“...is a specific form of sign language interpreting. It is the process of changing one form of an English message, either spoken English or signed English, into the other form. The assumption in transliteration is that both the spoken and the signed forms correspond to English, the spoken form following the rules of standard English and the signed form being a simple recoding of the spoken form into the manual code of expression” (Winston 1989, 147).

Transliteration incorporates features of American Sign Language (ASL) to enhance clarity. Ability to transliterate implies a knowledge of ASL features so that they can be incorporated into a transliteration.

While there is a small but growing body of research into the qualifications, efficacy, roles, and responsibilities of educational interpreters, there is general agreement that an alarming number of interpreters working in K-12 settings are ill-prepared for their work. In a survey from 2001, Jones reports that 50% of educational interpreters held no certification; 58% were not evaluated for interpreting skills prior to being hired; 31% had *never* been evaluated for interpreting skills; and 38% had never received in-service training. These numbers were largely corroborated by a subsequent study (Peterson and Monikowski, 2006) of educational interpreters in Ohio.

Needs

Clearly education is chief among the needs of these interpreters. There are relatively few baccalaureate programs that specialize in the training of educational interpreters, an odd fact given that fully 20 years ago it was found that more than 50 percent of graduates from interpreter education programs (IEPs) went to work in K-12 and postsecondary settings (Stuckless et al.) . A recent survey found there to be 130 IEPs in this country. Of that number, only 30 were baccalaureate programs and of those programs, only a very few specialized in educational interpreting.

The perception of inattention to quality has also led to educational interpreters suffering something of a stigma in the interpreting community, a phenomenon little studied but much discussed, to wit the recent brouhaha when the Registry of Interpreters for the Deaf recently decided to confer full voting status on educational interpreters who met certain qualifications.

Those qualifications centered on the candidates' scores on the Educational Interpreter Performance Assessment (EIPA), a dual-purpose instrument that both measures specific criteria of interpreting competence and provides diagnostic results and recommendations (Seal, 1998). Since its inception the EIPA has become the most reliable metric for classroom interpreting, and is now used widely across the country.

The early results from this wide netting of data on educational interpreting confirm long held suspicions about the quality of interpreting in many venues. On the positive side, however, this broad implementation holds great hope for the eventual identification and acceptance of standards in this field.

Possible Directions for the Future

The advent of Video Relay Interpreting has, in a very short period of time, changed fundamental considerations in the field of sign language interpreting. Video Relay Interpreting can be defined as a technological innovation whereby the interpreter and the two interlocutors are all in separate locations. The signing consumer and the interpreter can see each other on videophones. The hearing consumer and the interpreter

are linked by phone line. In this way the interpreter can serve the communication needs of both consumers without having to be physically present with either of them.

The growth of video work is unprecedented. In 2002, Video Relay Service (VRS) companies provided 530,053 minutes of interpreting service. In 2006 that number had grown exponentially to 44,326,554, and continues to grow apace. Projections for the year ending in June of 2008 are for the provision of 65,139,834 minutes of service (NECA, 2007). This explosive growth has meant that an ever-increasing number of interpreters are committing to a steadily escalating workload in the VRS environment. Given that the demand for this service far outstrips the number of qualified interpreters able to perform it, stress on the population of interpreters has already been felt. There is reason to believe that this situation will get much more serious before it improves.

The effect of fatigue on the efficacy of sign language interpreters and interpreting is little studied and poorly understood. Moreover, the work that has been done in this regard is focused on traditional (non-VRS) interpreting. While VRS interpreting has much in common with traditional interpreting, it also has some major differences (Taylor, 1995). VRS interpreters work in circumstances that are unique to video interpreting, e.g., working from a video source, working in a cubicle, working in a rapidly changing series of contexts, working in an overlapping series of roles (interpreter, customer service agent, operator, technician).

VRS interpreting needs to be differentiated from Video Remote Interpreting (VRI). Where VRS interpreting requires that the two parties being interpreted for be in different locations, VRI does not have that restriction. The absence of this stipulation means that VRI work has potential in classrooms. However, much remains to be known about such an application. While it might serve well in lecture settings, it would be largely unserviceable in other situations, like small-group discussions or any of the sundry dialogue-based interactions that happen in classrooms.

