HV
2502.5
.K66
2006

HWAET (Hand Waving Apparatus for Effective Turn-Taking)

by

Bertha Konstantinidis, BASc (Ryerson University, 2000)

A project

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Engineering

in the Program of

Electrical and Computer Engineering

Toronto, Ontario, Canada, 2006

© Bertha Konstantinidis 2006

PROPERTY OF
RYERSON UNIVERSITY LIBRARY

UMI Number: EC53506

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform EC53506
Copyright 2009 by ProQuest LLC
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I hereby declare that I am the sole author of this project.
I authorize Ryerson University to lend this project to other institutions or individuals for the purpose of scholarly research.

Sept. 28,06

Bertha Konstantinidis Date

I further authorize Ryerson University to reproduce this project by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Sept. 28,06

Bertha Konstantinidis Date

HWAET (Hand Waving Apparatus for Effective Turn-Taking)

Master of Engineering, 2006

Bertha Konstantinidis

Electrical Engineering, Ryerson University

ABSTRACT

This project describes the design and testing of an attention-getting device, called HWAET (Hand Waving Apparatus for Effective Turn-Taking). HWAET was designed to be used by sign language participants to indicate their desire to have a turn to speak during a sign language interpreted video conferenced meeting with hearing participants.

User studies were conducted with ten sign language users, fifteen hearing subjects and four different sign language interpreters to examine the use of HWAET during a video conference. These studies showed that deaf participants using HWAET found it easier to indicate and take their turn independently than those relying solely on a sign language interpreter (not using HWAET). Also, interpreters commented that they did not have to manage turn-taking when HWAET was used compared to the control group (where HWAET was not used).

ACKNOWLEDGEMENTS

I would like to express my gratitude to all those who gave me the possibility to complete this project. It would be impossible without the people who supported me and believed in me.

I would like to express my deep and sincere gratitude to my supervisor, Professor Deborah Fels, PhD., P.Eng. Her wide knowledge and her logical way of thinking have been of great value for me. Her understanding, encouraging and personal guidance have provided a good basis for the present project.

I would like to thank Jason Yeung, and Melanie Yeung for their assistance in developing the industrial design of HWAET. A special thanks to Emily Price who assisted in the studies.

A big thank-you to the deaf and hearing participants who were part of the studies, without you my research could not be completed.

I have furthermore to thank the "labbies" who believed in me and encouraged me in all the time of research for and writing of this project. In particular, I am thankful to Danny Lee, Raisa Rashid, John-Patrick Udo and Carmen Branje.

I would like to share this moment with my parents, John, brother, and in-laws. Especially, I am grateful to my son Menelaos for the inspiration he provided and whose patient love enabled me to complete this work.

Finally, I wish to acknowledge the funding support provided by Canadian Centre on Disability Studies and NSERC.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

To most hearing people the word deaf implies a disabling condition that involves the auditory inability of a person (Reagan, 1995). However, many deaf people do not see themselves as having a disability. Instead, they are a distinct cultural sub-group with unique cultural, linguistic and social systems just like any other cultural group (Reagan, 1995). They are recognized in many countries as a minority similar to other subcultures such as West Indians or East Asians.

As with other sub-cultures, deaf culture exists within the larger hearing society. In order to ensure that the intersection between deaf and hearing cultures is fruitful, informed and equal, deaf and hearing people must be supported in managing their relationship.

One important identifying characteristic of the deaf subculture is the language used in that culture - sign language, an example being American Sign Language (ASL). Sign languages are linguistically complete, natural language systems that express vocabulary and grammar visually and spatially by a series of hand gestures, facial gestures, such as eyebrow motion and lip-mouth movements, and body movements that change in time and space (Stokeo, 2001). They have no relation to spoken language (Stokeo, 2001).

Communication using sign languages is visual rather than sound based normally found with spoken communication. Communication between deaf and hearing individuals can then be extremely difficult because sign language users use gesture to communicate while hearing individuals rely primarily on speech and text. There are three main methods in which this is facilitated; 1) lip-reading; 2) text-based methods and 3)

interpreter services (Gallaudet, 2006). When lip-reading, a deaf person can only maintain a connection with one hearing individual at a time as they are unable to simultaneously "hear" two different voices or "look" at two different people and easily determine which person is speaking (Sacks, Schegloff & Jefferson, 1974). Hearing individuals may take this ability for granted and, because of this, be unaware of how this ability provides them with a major advantage over deaf individuals within a predominantly verbal world: the ability to instantaneously recognize and shift their attention between speakers. Because not all participants can be seen at the same time, deaf individuals are unable to pick up on verbal cues, which hearing participants constantly use to indicate turn-taking (Coates & Sutton-Spence, 2001).

Examples of text-based methods include older technology such as the Teletype (TTY) or telephone relay services (Ultratec, 2006) to support telephony, Internet-based services such as email or chat; or use of Personal Digital Assistants (PDAs)/Blackberries™ mobile phones as text displays. Text-based communication methods require some level of written language typing skills and there is no real way to communicate using sign language.

Another solution is to employ a sign language interpreter. Sign language interpretation is required when people who are deaf must interact with people who are hearing such as in business meetings, for court, and for accessing social and medical services. It is essential for providing equal access to these activities and services for people who are deaf, and in many western countries it is required through legislative initiatives (for example, see the Americans with Disabilities Act, US Department of Justice, 2003).

Sign language interpreters are usually preferred by deaf people, because they provide a more effective mean to communicate with hearing people. The person who is deaf and the person who is hearing are able to completely express themselves in their own language.

A deaf person can have access to interpreters through an interpreting service available in his/her community. However, there are several limitations with this model of service and communication: 1. Interpreters must be available (and paid) at all times; 2. Many services do not have the resources (either human or financial) to offer full sign language interpreting by qualified interpreters; 3. Interpreters are usually located in urbanized regions because the demand for that service can be constant and contained within a reasonable area making it resource-efficient; 4. Mis-matches between the knowledge and skill of the interpreter and the interpreting circumstances often occur. For example, interpreting for a medical situation can require specialized knowledge by the interpreter. As a result of not having access to sign language interpreter services for schooling or doctor/social services appointments, deaf adults and children can lack the ability to read, write or verbally communicate (Yarger, 2001). These limitations and challenges further add to the difficulties encountered by people attempting to communicate in mixed participant meetings or encounters.

Sign language services in rural areas are particularly problematic because the demand for interpreter services may be low and/or dispersed throughout a large geographical area, and often there are no interpreters in rural areas at all. If there is an interpreter, she might not have the qualifications needed to interpret in a specific educational or medical environment further aggravating an already difficult situation.

Recently, interpreting services have introduced video-conferencing (VC) (e.g., My Video Interpreter, 2004; and SignTalk, 2004) as a way of mitigating access to interpreter services for people in rural or non-service areas. In addition, it has been used to provide remote access to interpreters with specialized knowledge in medicine, education or other social services that would otherwise be unavailable in a particular community or area. However, VC presents new complications to already difficult communication situations between hearing and deaf people because the medium and the technology introduce additional limitations. For example, eye gaze, peripheral cues, control of the floor, having side conversations, or pointing to things that must be supported through a video conferencing interface are weaknesses of video conferencing in general.

In face-to-face meetings, the eye contact and gestures of each speaker often play critical roles in a facilitating smooth turn-taking whether between hearing people or hearing and deaf people (Vertegaal, Van der Veer, & Vons, 2000). However, most VC systems do not support turn-taking very well, because they do not convey eye contact correctly, and inhibit a smooth group turn taking process (Vertegaal, Weevers, & Sohn, 2002). The reason why this happens is that cameras are usually positioned on top of the monitor that displays the participant's face. Most systems broadcast images of a user from single camera unit to all other participant (e.g., Gaze-2: VC system) so that it is difficult for any one participant to establish eye contact.

Displayed as a two-dimensional image, the remote participant is unable to use recognized and accepted turn taking conventions (such as hand-raising) to establish his desire for a turn, since the camera's shot size is limited to his face. Were the remote

participant to raise her hand, it would most likely go unnoticed by other members of the group.

In this project, I present different turn-taking models for face-to-face communication and VC mediated communication for hearing and sign language users and that are used as a basis for a system developed to support turn-taking between deaf and hearing participants in a VC. These models describe the variety of cues and behaviours that are used to understand what is going on in a meeting such as who is being addressed, who is trying to take the floor, who is anticipating floor changes and who is participating in a conversation. HWAET (Hand Waving Apparatus for Effective Turn-taking) a physical waving hand is proposed and described as a solution to the problem of turn-taking management in video conference with mixed participants. The results of a usability study that examined the use of HWAET in an interpreter-mediated video conference between deaf and hearing participants are then discussed.

CHAPTER 2: LITERATURE REVIEW

The opportunity to participate in social interaction is one of the most fundamental requirements for practicable social organization (AMI 2006). Conversations or other context-specific forms of interaction (such as debate, interview, ritual, meeting etc.) form an integral part of a human society and culture. However, one important element of successful conversations regardless of cultural context is a common method for gaining attention that includes knowing when it is an acceptable opportunity to speak, and relinquish the floor (AMI 2006). In order to understand how turn-taking is managed in a conversation, different cues and behaviours must be considered.

In face-to-face meetings, attention is established through gaze, gesture and posture (Whittaker & O'Connaill, 1997). A speaker uses gaze to indicate that the person(s) being gazed at is an addressee of her utterance. From the speaker's point of view, looking at the listener may serve the function of monitoring the attention level and the processing status of the incoming speech, and help to regulate the flow of conversation. From the listener's point of view, looking at the speaker serves several functions: to provide feedback for the speaker's monitoring activity, to inspect the speaker's behavior (facial expression, posture, hands and arms gestures) for information about the speaker's attitude and emotion, and to monitor for nonverbal cues for turn-taking.

In this literature review, I provide an overview of turn-taking patterns in different types of conversations, including turn-taking in face-to-face and deaf conversations, with mixed participants and over video conferencing. In addition, a brief review of common video communication technologies is provided along with the limitations such as

inadequate turn-taking mechanisms that interfere with communication processes particularly between deaf and hearing participants.

2.1 Turn-taking in Face-to-face conversation

Turn-taking in conversation (TC) has had considerable attention in the research domain. TC is concerned with how participants jointly determine who will speak, who will listen and how transitions are made between these roles (Sacks, Schegloff, & Jefferson, 1974; Walker & Whittaker, 1990; Whittaker & Stenton, 1988). In general, turn-taking in conversations is said to be successful when one person speaks at a time and gaps between speakers are fairly short in time, Levinson (1983) suggests that these gaps are usually in milliseconds. Although some meetings might have some planned structure where speaking order is predetermined, (using a formal agenda, a chair person who moderates the conversation and mitigates turn-taking as in the Robert's Rules of Order, 2000) successful turn-taking occurs when interlocutors are aware of each other's movements and activities (Heath & Luff, 1991; Kendon & Ferber, 1973; Kraut, Fish, Root, & Chalfonte, 1993; Whittaker, Frohlich, & Daly-Jones, 1994).

Using techniques or conversation regulators such as gaze direction and pauses, speakers provide listeners with frequent opportunities to offer feedback about what was just said (Kraut, Lewis, & Swezey, 1982; O'Connaill, Whittaker, & Wilbur, 1993; Yngve, 1970) – to show acceptance (Clark & Schaefer, 1989) or to clarify their level of understanding (Walker & Whittaker, 1990; Whittaker & Stenton, 1988). This feedback is visibly conveyed by the conversations partner(s) using behaviour or actions performed with the eyes (gaze, facial expression, hands and arms (gestures) and posture (Whittaker & O'Connaill, 1997). Feedback, such as a participant uttering "yes", can also be provided

using verbal acknowledgements. These various actions are described within a turn-taking context of hearing conversations.

Gaze is an indicator of attention (Whittaker & O'Connaill, 1997). Speakers tend to look away from listeners while talking, otherwise it would signal "turn-yielding" to the listeners who are competing for the floor (Whittaker & O'Connaill, 1997) (floor refers to a conversational space that is occupied and then relinquished in some way by a person who is speaking called "the speaker"). When speakers tend to be at the end of their utterance, they begin to gaze at their listeners signalling that they are ready to finish speaking. The speaker waits for acknowledgement from the listener and then the listener takes over the conversation (Isaacs & Tang, 1993; Whittaker & O'Connaill, 1997). In face-to-face interaction, Isaacs and Tang (1993) found that the next person to speak was determined by the eye gaze or by suggestion of the preceding speaker.

Gesture refers to the set of dynamic movements and shapes formed by a person's hands and arms during communication (Whittaker & O'Connaill, 1997). Gesture serves to coordinate turn-taking transitions. For example a participant will extend his hand to the person who wishes to take a turn in the conversation.

The termination of any speaker's gestures is intended (and interpreted) as a "turn-yielding" cue. Gestures such as a person extending her hand towards another participant who wants to take a turn may also be used to indicate that the speaker still wishes to hold the floor. Similarly, listeners can signal that they have something to say (Goodwin, 1981). Hand gestures, such as hand raising or waving are conventional gestures used to gain attention. Isaacs and Tang (1993) noted that participants also used gestures to "reserve" a conversational turn. For example, two participants start speaking at the same

time; one of the speakers can point to a document, looses his turn, but keeps his finger on the document to reserve his turn.

In conversations, *posture* is the information supplied by the inclination and orientation of a conversational participant's body, in particular her trunk and upper body (Whittaker & O'Connaill, 1997). Though posture is more of an indicator to the degree of interest or engagement of a conversational participant, posture can also be use by listeners to signal a desire to interrupt (Whittaker & O'Connaill, 1997). For example, the participant can lean forward towards the conference table to indicate a desire for a turn to speak.

Most of the research carried out and models proposed have concentrated on spoken language with hearing people, and not conversation between deaf people or between deaf and hearing people. Nevertheless, the visible behaviours and conversational organization discussed in the previous paragraphs is similar among hearing or deaf people, or even when other mediums, such as video conferencing, are used to enhance remote collaborations (Video Interpreting Adhoc Committee, 2005). However, important differences exist and can cause difficulties in mixed conversations. These differences will be discussed in the next section.

2.2 Turn-taking in Deaf Conversation face-to-face

The medium of communication for the deaf population is visual rather than sound based because of the use of signed languages. Participants in a sign language conversation can attend to only those sources that they can see. This might not be a problem when there are only two participants, however, where there are more than two participants, not all contributors can be seen at all times by everyone in the group. This

makes the conversation more complicated, and demands more cognitive resources and attention from participants in signed conversations than in hearing conversations. Turn-taking then becomes more complicated to manage.

Researchers seem to differ in how turn-taking is managed in deaf conversations. Cokely and Baker (1980) describe the use of conversation regulators such as waving a hand or tapping a fellow speaker to get attention in deaf interaction as the main turn-taking strategy. The addressees can acknowledge the signer, accept the next signer's right to the floor and be ready to receive the signing communication.

Others such as Baker (1977) and Mather (1996) assert that deaf speakers will not initiate signing unless eye contact is established. Similarly, Siple (1978) and Swisher, Christie, and Miller (1989) suggest that deaf interaction uses the social norm of requiring that the addressee maintain eye gaze on the signer's face during signing. Their findings conclude that deaf people do not start signing, that is taking a turn in the conversational floor, unless they have established eye contact with a listener.

Coates and Sutton-Spence (2001) performed a study of two groups of four deaf participants to examine the different modes of conversational organization applicable to sign language speakers. In each group, two participants sat next to each other and faced the other two participants opposite of them. This arrangement meant that all the signers could turn to see each other, take and relinquish control of the conversation, and "hold the floor". The seating arrangement created a signing environment that allowed researchers to observe what happened when natural visual contact with all participants was hindered. The researchers observed that participants did not maintain direct eye contact at all times and that speaker(s) holding the floor could usually see more than one

person in a single visual field. This suggested that deaf participants can attend to more than one source of conversation at a time especially in a social setting; however, it might not be feasible or appropriate in a business meeting or doctor's appointment. An intervention of a physical hand or a person managing the meeting can assist in the turn-taking process.

It seems that non-speaking participants need to maintain eye contact with speakers, while speakers will keep a regular "eye" on all participants, but may look away from them at times. This can also be found in a hearing conversation (Coates & Sutton-Spence, 2001). However, in a deaf conversation it might mean that, if a person wants to make a contribution, she may run the risk of not being seen by anyone other than the current speaker. From their study, (Coates & Sutton-Spence, 2001) found that deaf people are willing to take a chance of not being seen while signing, as long as they are available to contribute and take part in the conversation. The reason for this is that signs can be held for some time and be repeated over and over, unlike words; an advantage that deaf interactants have over hearing interactants while negotiating for a turn (Coates & Sutton-Spence, 2001).

In a mixed conversation with sign language interpreters and hearing participants, turn-taking becomes much more complicated because the sign language user's visual attention is with the interpreter and not with others in the meeting. As such, they cannot attend to subtle turn-taking gestures, such as shifts in eye gaze employed by hearing individuals, in order to know that the speaking floor is available. Turn-taking must therefore be slowed down and made more obvious or formal so that sign language users have a chance to keep up and participate. However, it is generally assumed that the

presence of a sign language interpreter creates equality of both parties as the interpreter understands the cultural and communication differences and can mediate between them (Van Herrweweghe, 2002). For example, it is the interpreter who resolves turn-taking problems resulting from differences (Roy, 1989).

To mediate turn-taking, the interpreter has four options: 1) to stop both speakers, halt the turn of one speaker, and allow the other speaker to continue; 2) to momentarily ignore the overlapping talk, and upon finishing the interpretation of one speaker, offer a turn to the other primary speaker; 3) to momentarily ignore one speaker's overlapping talk, hold (in memory) the segment of talk from that speaker, continue interpreting the other speaker, and then produce the "held" talk immediately following the end of the other speaker's turn; and 4) to completely ignore the overlapping talk (Roy, 1989). The interpreter incurs much responsibility for managing or controlling the conversation as well as translating it - creating a potential for imbalance in the conversational power dynamics and a slower conversation flow.

2.3 Videoconferencing Technologies

Videoconferencing (VC) was introduced in 1964 to allow two or more people at different locations to see and hear each other at the same time. It was marketed as being able to eliminate gridlock and enhance travel experiences for people who lived in cities (Molnar, 1969). Video communication through VC is designed to support communication between one or more users when separated by long distance (Fels & Weiss, 2000). Videoconferencing provides remote participants with much of the face-to-face familiarity that comes with physical presence, including elements of facial expression, body language, and eye contact.

Current VC configurations consist of two systems, one at each network connection. Each system contains a camera, some combination of audio headset, telephone handset, microphones and digitising devices (hardware and/or software), speakers, a small preview screen (picture-in-picture capability), monitors or large screen televisions and a connection through ISDN or IP. These subsystems can be PC-based such as iVisit and NetMeeting setups or can be dedicated hardware such as a PolyCom™ ViaVideo® II (see Figure 1). There are also cameras and telephones (videophones) with built-in videoconferencing software that can be attached to a phone line and a television, along with a microphone; the television speakers provide the sound output (Sorenson, 2005) but these were not considered in this work.

Figure 1 Video Conferencing System

In order to have effective video communication, the hardware must be properly configured and housed in an appropriate environment. There are numerous technical guidelines published to assist organizations and individuals in proper environmental assessments and setup for various configurations of video conferencing hardware (see Brightline, 2002; McAteer, 2000; and Polycom 2005 to name a few) but few of these

guidelines address the special considerations required to accommodate people with special needs, particularly sign language users and remote interpreters.

There is a standard set of common hardware technologies and configurations for video conferencing regardless of who the user is and how the system is used. The two most popular types of network transmission technologies used for video conferencing today are Integrated Services Digital Network (ISDN) (H.320, 1984) and Internet Protocol (IP) (H.323V5, 1996). ISDN uses digital point-to-point connection between phonelines, which provide reliable transmissions ranging from 128 kilobits per second (kbps) to 384 kbps. These transmission speeds allow audio-video signals to be consistently transmitted at near broadcast quality (broadcast quality video transmission is 29.95 frames per second (fps)). ISDN is offered by regional telephone carriers and involves the digitization of the telephone network to transmit voice, data, text, graphics, music, video and other source material over existing telephone wires.

IP videoconferencing involves using Internet Protocols and technologies to process and transmit live video and audio signals. Video conferencing using IP protocols is governed by the International Telecommunication's Union (ITU) H.323 video standard (Polycom, 2005). Refer to Appendix A for additional technical information on transmission protocols.

2.3.1 Bandwidth

Even though video transmission and video conferencing methods have improved significantly over the last few years, individual's using sign language still experience difficulties. As sign languages are visual languages, good image quality is required to display clear, smooth and unambiguous signs. Presently, many video conferencing

systems cannot deliver the necessary frame rate to transmit sign language intelligibly. Conventional video conferencing technology generally addresses the limitations of channel capacity by drastically reducing the frame rate, while preserving image quality (Foulds, 2004). This produces a jerky image that disturbs the display of the trajectories of the hand and arms that are essential in sign language. In particular, deaf users report the importance of facial expressions and clarity of handshape as significant aspects effecting comprehension (Elliot, Glauert, Kennway, & Marshall, 2000).

King, Dellon & Murray (2003) report that many video conferencing applications in deaf education require at least 384 kbps (or a framerate of 30 frame per second (fps)) for comprehensible sign language, which will support a video screen update of 30 frames per second, equivalent to VCR playback quality television. However, for ISDN and IP video conferencing, 384 kbps is still sometimes jittery or unclear, especially for fingerspelling (King, Dellon, & Murray, 2003). They also report that Gallaudet prefers 768 kbps for IP video conferencing.

Even with somewhat jittery images VC still holds much promise as an accessibility solution to improve access to urban-centred services such as social services, doctors and postsecondary education for the deaf population in rural or underserviced areas. It also allows deaf users to communicate non-verbally with each other. A third advantage of video conferencing is that it can support hearing people who wish to work with people who are deaf, but do not have easy access to interpreters in their physical setting.

2. 4 Turn-taking in VC

There remain a number of important issues that can be addressed to improve the communication flow within the current bandwidth and technical limitations. One specific area where VC seems to fall short in supporting human-to-human communication is in mediating the regulation of conversational turn-taking particularly the use of eye contact to indicate a turn is desired or completed (Coates & Sutton-Spence, 2001; Isaacs & Tang, 1993; Vertegaal, 1999). Most VC systems do not convey eye contact correctly due to poorly placed cameras and limited camera angles, and as a consequence hinder smooth turn-taking. As seen in Figure 2 the person in the remote image (on the screen) appears to be looking down. This is due to the vertical discrepancy between where the camera is located and where the person's gaze is directed. The person's gaze is directed at the screen while the camera is located above the screen. If the person was to look at the camera, it would appear to his communication partner that he was looking into his eyes. However, if he were to do so, the remote partner could no longer see the screen and the other person, preventing mutual gaze from serving its purpose as a synchronisation signal (Grayson & Monk, 2003). Also, it is sometimes difficult to have a clear view of the participant's face because of network capabilities, the camera being wrongly positioned or a low quality camera.

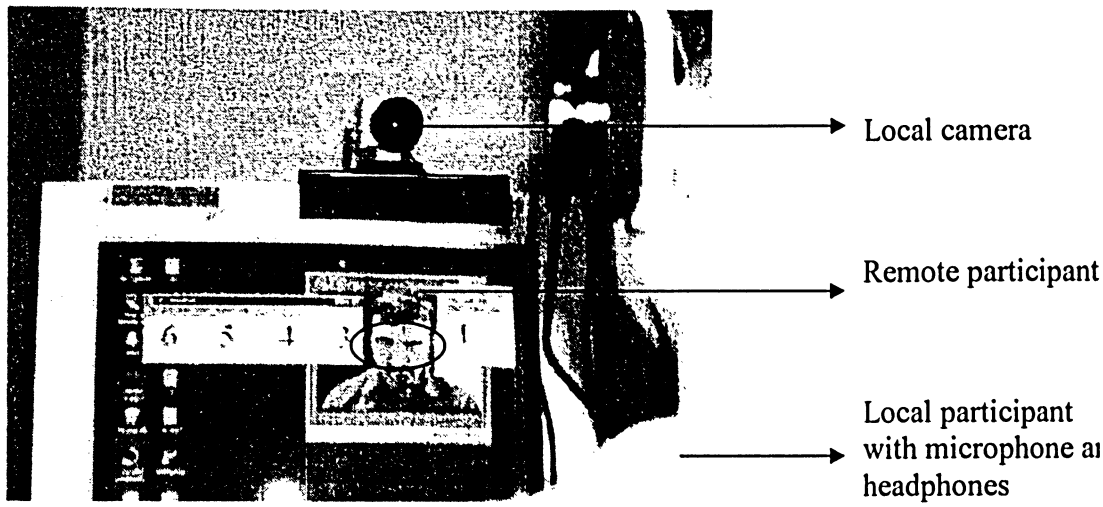

Figure 2 Typical Configuration of Conventional Desktop Video Conferencing, Grayson & Monk 2003

Researchers have shown that interacting remotely through video makes it difficult or creates difficulties for participants to: manage turn-taking, control the floor through body position and eye gaze, and notice motion through peripheral vision (Isaacs & Tang, 1993). Isaacs and Tang (1993) conducted a usability group of five participants using a desktop VC system. They found that during video conferences users addressed each other by name and started explicitly requesting individuals to take turns. O'Connaill, Whittaker and Wilbur (1993) also found that during a VC, more formal techniques such as addressing each other by name or using gestures to achieve speaker switching were typically used compared in face-to-face interaction. This was also attributed to the absence of eye contact.

Vertegaal et al. (2000) studied the effects of eye contact on turn-taking in three-person video conversations, and found that when eye contact was absent, participants took about 25% fewer turns. Without eye contact, 88% of the participants indicated they had trouble determining to whom their meeting partners were talking and were confused. These findings indicate that eye gaze seems to be an important determinant in defining a successful video conference.

There has been some research on mechanisms to mediate eye contact for VC. Grayson and Monk (2003) investigated using a half-silvered mirror to place the camera in the same virtual position as the monitor to eliminate the discrepancy between the position of the camera and the position of the eyes of the conversation partners. Vertegaal (1999) used an eye movement monitor to detect when one participant was looking at the image of another participant's face. With his GAZE system, he provided the illusion of mutual gaze by using prerecorded still images of the other participant looking in different directions. The system solves the problem of gaze awareness in multi-party conversations, where gaze awareness may be particularly important in regulating turn-taking.

The use of half-silvered mirrors, eye trackers or multiple cameras at each site makes for a complex and expensive system that remains the exclusive domain of research laboratories. The most common and cost-effective configuration for sharing video images is still the single camera placed on top of a PC monitor. Solutions for mediating communication for deaf people must fit with this more common and affordable configuration.

2.5 Turn-taking in VC with mixed participants

One of the strengths of video conferencing technology is that it offers sign language users the potential to remotely communicate using sign language with each other and with hearing people in conversations, meetings, and social gatherings because of the video aspect of the technology. Still, to communicate with spoken language users, a sign language interpreter must be employed, but this person can be physically located elsewhere.

In a face-to-face communication, the interpreter maintains eye contact with the person who is deaf so that they can maintain a separate interpersonal communication between them (Van Herreweghe, 2002). The interpreter does not look at other participants because the interpreter uses visual cues and attention to focus attention of the person who is deaf and because she must interpret when required. Remote interpreting can interfere with this practice, particularly when the person who is deaf is remote from the interpreter (e.g., the interpreter is physically present with the hearing participants or remote to all participants). Interpreters who are remote to all parties must always look at the viewing screen to watch the person who is deaf in order to translate to the hearing participants. This may aid hearing participants in attending to the deaf person, but they are likely watching the interpreter instead. Also, due to the limitations of the medium and the difficulties determining eye gaze and turn-taking, the interpreter and deaf participant may not establish the subtle, interpersonal communication signals that allow the interpreter to carry out her responsibilities of focusing attention on the person who is deaf, particularly for turn-taking.

When VC is introduced into the conversation mix, turn-taking mechanisms become cumbersome for the interpreter because there is reduced display real estate for signing due to the reduced field of view of the camera and small computer screens. Interpreters are constantly obliged to reduce their signing space to accommodate the camera's field of view. They can not sign outside of the area captured by the camera. As a result interlocutors constantly miss turn-taking cues resulting in communication errors such as overlapping each other or interrupting, having long moments of silence, and taking control of the floor for lengthy periods of time.

In a mix conversation with remote interpreting, turn-taking is also affected by problems that already exist in VC, such as awkwardness of using gesture, and eye gaze. These non-verbal cues heightened the difficulties of turn-taking, because of the reliance on the visual domain and the restrictions imposed by the VC hardware, such as camera angle and position.

Different approaches to support the turn taking issue with remote interpreting must be considered. Often the role of the interpreter is to facilitate turn taking, but it can be a difficult task because he may miss the subtle turn taking cues such as laughter, clearing the throat or specific gestures as described in section 2.2. Interpreters and hearing participants may miss these indicators because they are awkward to execute and notice over video conferencing due to the limited field of view of the camera and the low quality of the video image.

It may be possible to implement a technological solution that can support turn-taking. Fels and Weiss (2000) investigated the use of different technologies such as lights (flashing, spinning, etc), and a waving hand as a way of improving turn-taking for video conferencing. They found that a waving hand that can be activated when a participant wants to take a turn was most effective in gaining the attention of meeting participants. In this project, a waving hand regulator, called HWAET, was designed to support deaf people and interpreters during video conferencing by providing a physical system that a deaf person can use to indicate a desire to have the floor during meetings with each other or with hearing people. In the next chapter a detailed description of HWAET is provided followed by the report and discussion of a user study.

CHAPTER 3: HWAET DESCRIPTION

HWAET is an Old English word that means to gain the attention of an audience listening to the oral delivery of some content (Powell, 2005) and we could say that it is the Old English equivalent of today's "Yo!" used to gain attention of an audience. It seemed fitting that a device designed to be a regulator (see Figure 3) in a conversation to assist turn-taking be named HWAET

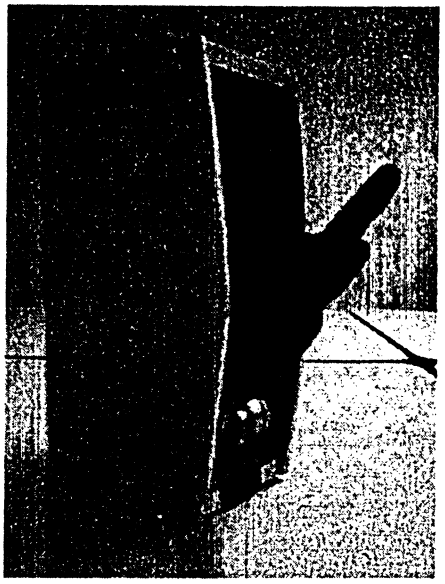

Figure 3 HWAET (Hand Waving Apparatus for Effective Turn-Taking)

HWAET was designed to represent a physical waving hand that simulates a swinging motion similar to what happens when a person mover her hand from an upright position down while bending the elbow through about 90 degrees. The hand would be located in the view of other participants in a video conferenced conversation and would be activated by the deaf participant in the remote location. Activation of HWAET indicates that the deaf participant wants to take a speaking turn.

The hand is enclosed in a protective box to hold the circuitry, and secure the hand's motor and solenoid. The hand is covered with a child's glove and inside the glove

are beads. The beads make a noise when the hand is activated by the deaf participant at the remote location to gain the attention of hearing participants who may not be in the line of sight of the video monitor or the hand.

A Sign Language Video Conference (SLVC) application was developed in Microsoft Visual Basic, using a NetMeeting OCX and NetMeeting Software Development Kit (SDK) to implement HWAET for a video conferencing application. SLVC was designed as a simple and easy to use VC so that the use of HWAET could be investigated during video conferencing without the need for complex, high fidelity hardware, and to allow control of various interface elements such as video window size and icon placement.

The data flow diagram (DFD) seen in Figure 4 is used to illustrate the functions and data flow of the HWAET system. The square boxes (terminators) describe "who" is carrying out a process (indicated by circles) or the external entity to which the system communicates. The DFD begins with a terminator that represents the deaf participant and that initiates a call to a remote location (meeting room terminator) using the SLVC software. The SLVC software process responds to the deaf participant's action by placing a call to the meeting room. The arrow into or out of a terminator describes the data flow in the DFD. In Figure 4, the process "Make Call" is the action the system is taking. The call is accepted in the meeting room and a double arrow indicates that there is two-way data flow, in this case, requesting a call and accepting it. Once the call is accepted, the call is then processed so that incoming/outgoing video (and audio) is the next process. This process has a dialog flow to the meeting room and deaf person, because both are sending and receiving incoming/outgoing video. Once there is a connection established

and incoming/outgoing video is processed, the deaf person can request a turn to participant in the conversation. The "Request a Turn" process sends a signal to activated HWAET. The "Meeting Room" participants notice HWAET and acknowledge the deaf person.

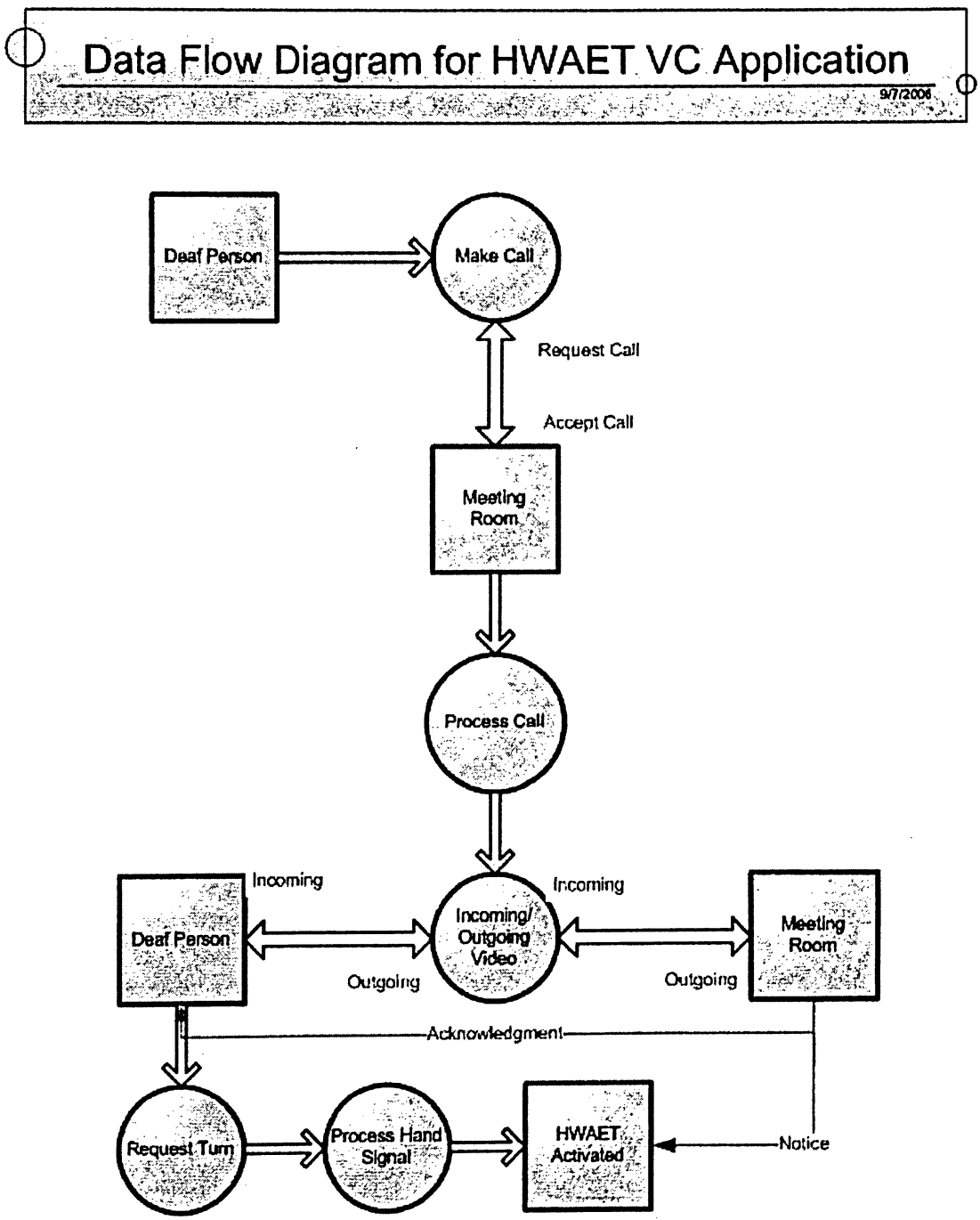

Figure 4 Data Flow Diagram for HWAET VC Application

The user interfaces of the SLVC application for deaf and hearing participants is seen Figure 5 and Figure 6 respectively. In the deaf person's interface, the larger window in the top left corner shows the video of the meeting participants in one location (named remote video) and the video of the deaf participant is shown at the bottom right hand side (named local video). The deaf participant uses the hand icon button to activate the physical hand to wave in the meeting location (remote location for deaf participant). The deaf participant can also terminate the conference call by pressing the "Hang Up" button. In the meeting location, the interface does not include the hand icon button. The incoming videos will display automatically, once the call was made.

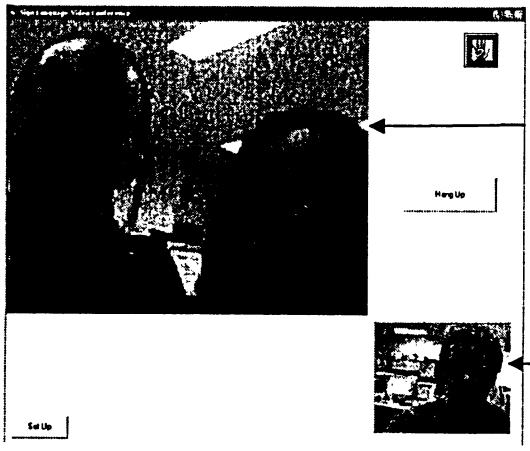

Figure 5 SLVC for Deaf Person

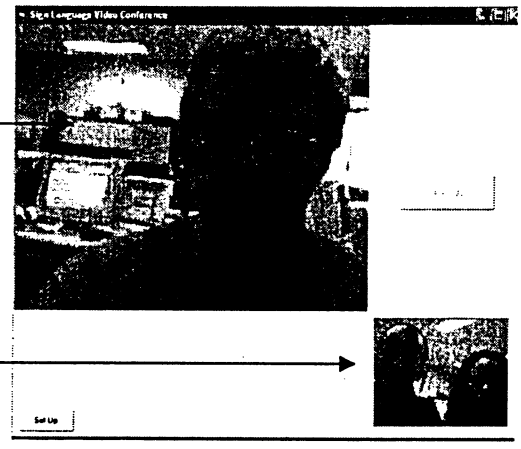

Figure 6 SLVC for Hearing Participants

The video conferencing cameras used were Logitech QuickCams. These inexpensive web cameras were used to emulate a typical video conferencing setup for remote interpreting situations.

3.1 Microsoft Visual Basic

Microsoft Visual Basic (VB) was used as a development environment because it allows rapid prototyping and interface development as well as supporting NetMeeting

video conferencing libraries and Component Object Model (COM) interface. VB is a semi-object oriented "event-driven" environment, which means that procedures are called automatically when the end user carries out actions such as choosing menu items, clicking the mouse, and dragging and dropping objects on the screen. It is also a graphical development environment making it relatively simple to quickly develop, test, and re-develop windows applications in response to user feedback.

3.1.1 COM development environment

The standard video conferencing libraries for the Microsoft Windows XP Professional version 2002 operating system are found in the NetMeeting development environment. NetMeeting is real-time collaboration tool that allows individuals to communicate in pairs or groups over the internet using audio, video and data communication. The NetMeeting libraries allow a good quality conferencing solution that includes text chat, whiteboard, file transfer, as well as point-to-point audio and video to be easily configured.

The COM development environment enables developers to add NetMeeting video conferencing functionalities such as managing calls to independent applications, to replace the NetMeeting User Interface in those applications, and to write applications that work within a NetMeeting conference. The COM development environment specifies an object model and programming requirements that enable COM VC objects (also called COM components, or sometimes simply objects) to interact with other COM or COM VC objects. These objects can be within a single process, in other processes, and can even be running on remote machines.

PROPERTY OF
RYERSON UNIVERSITY LIBRARY

A COM object is one in which access to an object's data is achieved exclusively through one or more sets of related functions, called *interfaces.* The functions of an interface are called *methods.* Further, COM requires that the only way to gain access to the methods of an interface is through a pointer to the interface.

COM *interfaces* are different from interfaces typically used in Visual C++ programming. A C++ interface refers to all of the functions that a class supports and that clients of an object can call to interact with it. A COM interface refers to a predefined group of related functions that a COM class implements, but a specific interface does not necessarily represent all the functions that the class supports. With these COM interfaces and VB, audio and video capabilities are used to create communication collaboration, such as SLVC.

3.1.2 Rapid Prototyping

For this project VB and the COM development environment were used for prototyping of HWAET because of the ease and agility of VB to constructing graphical interfaces. Also, making changes to the prototype based on feedback from users and design experts were easily accomplished in VB without the need to rebuild low-level software components. The development focus of this project was to develop and test prototypes, not to produce a final version of HWAET.

CHAPTER 4: USER STUDY

4.1 Research Questions

The primary objective of this project is to carry out a user study with HWAET to determine its effectiveness, usability and acceptability. The effectiveness of HWAET is determined by the level of success of attention-getting by sign language speakers during VC meetings. A second objective is to examine the turn-taking behaviour of deaf meeting participants when using the HWAET assistive device and compare it to turn-taking behaviour without HWAET during video conferenced meetings. The results from this study can contribute to improving the attention-getting mechanism, HWAET, and its software VC interface as well as to gaining a better understanding of turn-taking behaviour during VC between deaf and hearing participants.

The research questions to be addressed in this project are the following:

1) Can HWAET be successful in allowing people who are deaf to indicate the desire for taking a turn to speak during video conferenced meetings with hearing people?

2) Is HWAET acceptable and useable in supporting the need to gain access to the floor by a deaf participant in a video conference.

3) How does the role of the interpreter change as a result of using HWAET?

4.2 Methodology

4.2.1 Method

Ten sign language users (eight deaf and two hard of hearing), 15 hearing subjects, and four different sign language interpreters participated in a study designed to examine the use of HWAET during a video conferenced meeting. A pre-study questionnaire available in Appendix C was administered prior to beginning the study to determine

people's level of computer knowledge, experiences with attention-getting and turn taking with respect to remote interpreting. The questionnaire consisted of 17 questions: seven questions involving information about demographics (e.g. gender, age, education and deaf or hearing), and familiarity with computers, and four questions regarding communication behaviour. The other six questions were only directed to sign language users. These questions included information on strategies used to get attention from hearing communication partners, on experience and difficulties with video remote interpreting, and on difficulty with taking a turn to speak during a video remote interpreting. The hearing people answered them as "not applicable".

There were 14 males and 11 female subjects in total (six male and four female subjects in the sign language group, and nine male and six female subjects in the hearing group, all interpreters were female). The average age range of the sign language user group was the 25-34 years age group, and 19-34 years for the hearing subjects. The average level of education for the sign language users was a technical college diploma, and for the hearing group it was an undergraduate university education. Twenty-three subjects reported using a computer every day (the two other subjects reported using a computer a few days per week). Only three people reported having experience with video conferencing (two sign language users and one hearing person).

Sign language users were asked about their experience in meetings with hearing people with and without interpreters. These participants attended an average of one meeting per week which included hearing individuals and interpreters (two people reported having daily meetings with hearing people and interpreters). In addition, deaf people reported attending meetings with hearing people and no interpreter an average of

once per month (only one person reported having to attend meetings with hearing people and no interpreter on a daily basis). When asked what the most difficult tasks were during meetings with hearing people and interpreters, one person reported difficulties keeping up with the conversation, two people found it difficult to request a turn to speak, five people had difficulties trying to see the interpreter and other meeting participants simultaneously, one person reported difficulties with having people "get to the point", three people reported other reasons such as finding qualified interpreters.

Only three people reported having any experience with remote video conferencing and noted that trying to keep up with the conversation, seeing the interpreter and others simultaneously, and being misunderstood by their interpreter as their main difficulties.

When asked about the advantages of video remote interpreting, five sign language users suggested that it would be easier than typing English words. They suggested that the disadvantages would be comfort with the cost and configuration of the technology, being misunderstood and lack of good interpreters.

4.3 Experimental Setup and Data Collection

Each participant was asked to read and sign a consent form (see Appendix B), or had the consent form sign language interpreted, and acknowledged their consent on video tape. The consent form was approved by the Ryerson Ethics Research Board. All participants were informed of the intention to collect verbal and non-verbal information through note-taking and videotaping, and that the data would be analyzed and presented in summary form.

Subjects were placed in groups of five with three hearing subjects, one sign language user and one sign language interpreter. The interpreter and the three hearing

participants where located in one room, while the deaf individual was located in a remote usability laboratory. One of the hearing participants acted as a chair for the meeting. The chair's role was to facilitate the discussion. This configuration simulated one possible standard video remote interpreting configuration. Although only this one configuration was used, it represents a likely scenario for people in underserviced or rural areas (deaf person living in underserviced area, services including interpreter services are available in a remote urban location. However, regardless of the configuration, eye contact between the interpreter and the deaf must be maintained as it is in the study configuration.

Some of the hearing subjects participated in multiple sessions due to availability and training requirements for hearing subjects. Having hearing subjects participate more than once did not seem to have any affect on being aware of HWAET. Hearing participants who participated twice were still surprised when the hand was activated by the sign language. Hearing participants were really involved in the discussion. Where this was the case, data for these subjects was only collected from their first session. There were eight experimental groups that employed the HWAET technology and two control groups that did not. The meetings were twenty minutes long and used a variety of topics such as gun control, public transit, and global warming.

During the meetings, data were collected by videotaping and screen capture. All of the participants were videotaped as a group in the meeting room, and in the remote setting (usability lab), the computer screen was captured. The computer screen contained one video image of the sign language user in the usability lab and a second image containing the other meeting participants in the meeting room (picture-in-picture) as seen

in Figure 5. In addition, the meeting chair and the observer at the remote usability laboratory took notes on any technical issues or behaviours from hearing and deaf participants.

Following the meeting, participants were asked to complete a post study questionnaire designed to capture information about their experiences and their opinions about the remote interpretation experience, taking turns during video conferencing and, for those using it, their opinion of HWAET.

4.4 Data Analysis

Thematic outcome measures for the video commentary were derived by two independent reviewers and then focused into five measures by agreement. Having these five measures, will assist on analysing the effectiveness and usability of HWAET. These measures shown in Table 1 were used in matrices of the type proposed by (Miles & Huberman, 1994).

Two of these measures, "Use of HWAET" and "Turn-taking", were selected as important indicators of HWAET's ability to support turn-taking between the deaf participant in remote location and the hearing individuals in the meeting room. Participant comments attributed to these measures were also identified as positive or negative. The positive and negative subcategories served to indicate a deaf person's ability to operate the physical hand and her success/failure in gaining a turn for speaking in a mixed meeting with hearing people and interpreter. A positive outcome in the "Use of HWAET" measure indicated that the deaf participant activated the physical hand to gain attention from hearing participants or to request a turn in the conversation. A negative outcome indicates the participant did not use HWAET and used other means

instead (e.g., the interpreter or putting their hand up). In the "Turn-taking" measure, a positive outcome was marked when the deaf participant was acknowledged by the hearing participants to obtain the "floor". A negative outcome was indicated when there was a confusion as to whose turn it was and/or the deaf participant did not have a chance to take a turn.

The "Inclusion" category was selected to record how the subject was included in the discussion or whether hearing participants were addressing the interpreter instead of the deaf person. Finally, the measures "Technical Issues" and "Signing Issues" were selected to document the frequency of technical issues such as video quality, computer problems, hand malfunctioning, and difficulties with signing quality such as speed or legibility which could have disrupted a subject's ability to participate in the meeting discussion. There were no positive outcomes for "Technical and Signing Issues".

Table 1: Description of Event Categories

Theme	Description	Positive	Negative
Use of HWAET	Use of HWAET to gain attention to take a turn	Deaf person uses HWAET to interject in the meeting.	Deaf person uses own hand or other gesture/strategy to gain attention
Inclusion	The deaf person is included in the conversation.	Hearing person addresses deaf person by looking at the camera/screen not the interpreter.	Hearing person addresses interpreter instead of the deaf person.
Turn-taking	Taking a turn to speak	Deaf person is able to take a turn successfully. Smooth indication and transition to next speaker.	Deaf person is unable to take a turn successfully. Confusion arises as to whose turn it is to speak.
Technical Issues			Difficulties with video quality, or software problems with the computers. Issues with the HWAET malfunctioning.
Signing Issues			Signs were not legible, or person signing too quickly

In order to determine the inter-rater reliability of the video analysis procedure and the operational definitions, two evaluators were instructed on the video analysis procedure, and trained in the coding categories. Each evaluator then independently rated the same two randomly selected video data sets and their ratings analysed. The single measures The Intra Class Correlation (ICC) for all themes and for the positive/negative categories was 0.8 or better. This indicates that the inter-rater reliability was very good. All subsequent analyses were carried out by a single evaluator.

Due to the relatively small number of measure counts and subjects, non-parametric analyses were used to compare the control and experimental data. A descriptive and qualitative presentation of the results follows.

CHAPTER 5: RESULTS

As the control group did not use HWAET, only the turn-taking and inclusion measures overlapped between the two groups. A Mann-Whitney analysis was carried out between the means of the total positive and negative number of instances for the Turn-taking and Inclusion measures. No significant differences were found.

The length of time that people spent speaking during a turn was then examined. Duration was calculated to include that act of gaining attention and any activity that occurred as a result of that attempt. A Mann-Whitney was carried out on the duration of the turns between deaf participants in the control and experimental groups. There was a significant difference between the control group and the experimental group in the time spent in turn-taking ($U= 1613$, $p< 0.05$).

Table 2 shows the mean duration (in seconds) for the "Turn-taking" category in the experimental and control groups. The control group took longer turns (mean = 26 sec, SD = 1.1 sec) compared with the experimental group (mean = 21 sec, SD = 11.6 sec).

Table 2: Mean Duration and Standard Deviation (seconds) for Turn-taking in the experimental and control group.

Group	Mean Duration (sec)	Standard Deviation (sec)
1. Experimental	21	11.6
2. Control	26	1.14

Table 3 shows the frequency of positive and negative events of "Use of HWAET", "Inclusion" and "Turn-taking" event measures and the ratio of positive to negative events for all participants in the experimental group. Table 4 shows the frequency of positive and negative events and the ratio for only the "Turn-taking" measure. There were similar numbers of positive to negative events for the "Use of HWAET" and the "Inclusion" measures for the experimental group (ratio near one).

The ratio of positive to negative events for the "Turn-taking" measure indicated that the experimental group had six times more positive than negative instances of turn-taking while the control group had only 1.2 times more positive instances than negative ones.

Table 3: Frequency of occurrence of positive and negative events for the experimental study.

Categories	Number of events	Number of positive events	Number of negative events	Ratio (positive / negative events)
1. Use of HWAET	91	44	47	0.9
2. Inclusion	9	4	5	0.8
3. Turn-taking	105	90	15	6.0

Table 4: Frequency of occurrence of positive and negative events for the control group.

Categories	Total number of events	Number of positive events	Number of negative events	Ratio (positive / negative events)
Turn-taking	24	13	11	1.2

"Technical" and "Signing Issues" were measured for the experimental and control groups. There were a total of 13 negative events during the experimental study (eight technical and five signing issues), and none during the control. Most of the "Technical Issues" were caused by poor video quality and latency due to low bandwidth. Poor video quality and latency produced choppy or blurry looking signs making it difficult to understand the interpreters' and the deaf person's signing. Several of the technical problems resulted in the need to reconnect from the remote location, which increased the amount of time spent on correcting technical difficulties. The average duration of technical issues was 42.3 sec, with a standard deviation of 2.7 sec where duration refers to the time from when participants first noticed that there were some technical difficulties until the problem was corrected. The "Signing Issues" lasted an average duration of 8.4

sec with a standard deviation of 2.0 sec, where duration equals the length of time required to clarify a mis-understood signed communication over the VC.

5.1 Post-study questionnaire

A nine question post study questionnaire (see Appendix D) was completed by the ten sign language users (eight users from the experimental group and two from the control group). All participants were asked about the difficulties communicating their ideas, knowing when to take a turn, and gaining attention in the group. The experimental group was asked about the ease of use of HWAET (hand-mechanism).

Specifically, four questions used a five point Likert scale ranging from "Very Difficult" to "Very Easy" to rate the ease of use of HWAET, and level of difficulties with different communication functions including using a remote interpreter. One question used a three point Likert scale (ranging from "Not included" to "Included") to rate level of inclusion in the video conferencing meeting. The final forced-choice question used five point Likert scale question ranging from "Much More Confusing" to "Much Less Confusing" to rate the confusion level when taking a turn compared with a regular video conferencing conversation or face-to-face meeting depending on the participant's experience. The three remaining questions were open ended questions that asked for suggestions/improvements and any problems the subjects had with HWAET.

Seventy-five percent (six of eight) of the participants in the experimental group rated getting attention during the meeting as easy and 13% rated it as very difficult. In comparison, all (100%) the subjects from the control group rated their ability to get attention as either neutral or difficult.

Thirty-eight percent (three of eight) of the participants in the experimental group rated their ability to know when to take a turn as easy, and 50% (four of eight) percent as difficult, while all (100%) participants in the control group found it difficult or neutral.

Four of eight participants (50%) reported that HWAET was easy to use and did not have any problems learning how to use it, and two of eight subjects found it difficult. Two of eight participants (25%) found HWAET to be less confusing when requesting turn-taking in video conferencing, three people (38%) did not find a difference and three (38%) sign users found it somewhat confusing.

CHAPTER 6: DISCUSSION

The time taken on the turn-taking task is an important indicator of the enabling effect of HWAET. There was a significant difference in the time spent on turn-taking between the HWAET and control groups where the deaf participants in the control group spent significantly more time on turn-taking tasks than the HWAET users. However, the deaf participants in the control group were not speaking longer. In fact, they were required to interrupt to take a turn and the sign language interpreter often had to speak to facilitate that interruption whereas in the experimental condition, the sign language interpreter did not need to speak to indicate the deaf person wanted a turn. This need to interrupt on behalf of the deaf person added extra time to the turn taking and also increased the confusion by the hearing group, as they attempted to discover who was trying to speak.

Although there were no other statistically significant quantitative results, the descriptive and qualitative results seem to indicate that having an attention getting technology enables deaf users to participate more successfully in video conferenced meetings. The deaf participants in the experimental group were better able to gain attention because the hand was successful at indicating that the deaf person wanted to take a turn to speak and the hearing individuals understood that this was the case. There seemed to be fewer confusions, missed or ignored turn-taking indications by the hearing group when the hand was used. There were also six times more positive turn-taking instances than negative ones for the experimental group using HWAET, and most of the participants reported that using it to gain attention was easy. This might imply that deaf

participants can effectively use HWAET to indicate their desire to have the floor to contributions in meetings.

In the post-study questionnaire, 75% of people in the experimental group rated their ability to gain attention as easy and only one person found it difficult while all people in the control group found gaining attention in the meeting was either difficult or neutral. The one person who found it difficult to use the hand reported that he was accustomed to working through an interpreter and preferred that method. However, this person did think that HWAET was a good idea but that he likely needed additional time to become accustom to using it. Positive comments by deaf participants in the experimental group such as [I got] "immediate attention from the group" were commonly expressed by the remaining participants during the study. It seems that HWAET facilitated turn taking by allowing the deaf person to gain the immediate attention of the other meeting participants.

In the control group, the interpreters spent about one-third of their time and effort trying to orchestrate smooth turn-taking using techniques such as pointing to the hearing or deaf person who wanted to speak next, and keeping her finger pointed until the turn was recognized. Often the interpreter would wait until a hearing person had finished and then indicate that the deaf participant wanted to speak. As a result, the interpreter became the focus of attention for the turn, and experienced difficulty managing the turn-taking situations when more than one person wanted to speak. At times, this seemed to disrupt the flow of the meeting in that people were waiting or were confused about who was going to speak next.

In the experimental group, when the deaf person used HWAET, the hearing participants finished their turn and then focused their attention on the video image of the deaf person rather than on the interpreter. The interpreter was not required to manage the conversation turn-taking at all. The flow of the conversation was smoother because people were not waiting for the interpreter to indicate the need for a turn. In addition, the interpreter had less work to do to manage the conversation and did not need to interrupt on behalf of the deaf person. This empowered the deaf person to manage her own turn-taking desires and did not force her to rely on the interpreter to indicate the desire to have the floor. Similar results regarding children participating in school via video conferencing have been reported by (Fels & Weiss, 2001)

The deaf participants did, however, rely on the interpreters to indicate that HWAET was functioning which resulted in a high number of negative "Use of HWAET" events. While there was feedback on the video conferencing display, it did not seem to be noticed. In two of the studies, the hand was placed so that the deaf participant could see it in their remote screen which then seemed to alleviate the need to confirm that the hand was indeed functioning with the interpreter.

In the post-questionnaire, hearing participants commented that the hand startled them and did get their attention. However, in one meeting where there was a particularly intense conversation, the hearing participant acknowledged the hand, but the deaf participant was not able to have their turn until the hearing person finished and until another hearing participant who immediately "jump-in" to take their turn was finished. As a result, the deaf participant waited for her turn. In this situation, the interpreter had to remind the hearing participants that the deaf person needed a turn.

Even though the interpreters were not formally interviewed in this study, because the study was designed to gather impressions of HWAET from deaf participants and it was an exploratory study, they did make comments about HWAET during the meetings and afterwards. The interpreters believed that HWAET was a great idea. They mentioned, however, that deaf people would need time to become accustom to the novelty of a physical hand because it is such common practice for deaf people to rely on the interpreter to mitigate turn-taking. The interpreters also reported that they had to facilitate the turn-taking when HWAET was not used and that this practice was usually expected from them from the deaf community. Interpreters suggested that when they assist in turn-taking management conversations can become confusing especially during video conferencing, because it increases to the high cognitive load and corresponding fatigue levels to their interpreting role. They mentioned that the Canadian Hearing Society (CHS) attempts to resolve the turn-taking issues over VC by employing a conference manager at one site. This person is not the meeting chairperson and is only responsible for managing all of the technology as well as maintaining a formal speaker's list (by monitoring people's desire to speak). This approach is reported by CHS as better than using the interpreter to manage the turn-taking during the meeting but it is a costly solution. Employing a technology such as HWAET might be at least as efficient as this method. Further study is required to compare the method used at CHS, which is using a conference manager with HWAET.

The results in the usability, effectiveness and ability to learn HWAET is promising with deaf and hearing people. Deaf participants found HWAET easy to use and were able to take a turn in the meeting without relying on the interpreter. Even

though many of the deaf participants were not exposed to video conferencing, they believed that HWAET was a usable technology in combination with VC particularly as VC becomes more popular for people who are deaf.

In this study, when deaf participants activated HWAET, they were quickly acknowledged by the hearing participants in the remote location and had ready access to the floor as a result. Once the hearing participant finished their comment, the floor was given to the deaf person. This shows that HWAET was acceptable by the hearing and deaf participants. The interpreter did not need to have to point her finger to indicate that the deaf person needed a turn to speak. This alleviated the interpreters from the extra task and workload of facilitating turn-taking and instead allowed them to focus on translating tasks. The conversation did not have to slow down because of turn-taking issues, and sign language users seemed able to keep up with and participate in the meeting.

CHAPTER 7: LIMITATIONS

While the results of this study seem to support the use of a physical attention-getting technology to facilitate turn-taking by deaf participants in video conferenced meeting, there were several important limitations. This preliminary study used a fairly small number of subjects and an unbalanced number subjects in each group (control and study). To examine whether there are statistically significant differences between groups using an attention-getting device and those who do not use one, additional subjects are required in the experimental and control situations.

Using a larger study set also presents the opportunities for further analysis. In this preliminary study, all of the participants were unfamiliar with each other, and none was familiar with the interpreters. Often deaf people use preferred interpreters and have an established system between them. Studying the use of HWAET in the situation where the interpreter and deaf person know each other may show that HWAET changes this established relationship.

A second limitation is that the study used a controlled setting and contrived and researcher determined but relevant topics. Participants did not have a choice regarding the topic. The chairperson for the meeting was also predetermined and was always a hearing person. Deaf participants were not necessarily comfortable with the hearing participants and the topic might have not been of their interest. In this study, two deaf participants had recently arrived to Canada and were not familiar with the selected topics (e.g., global warming). They were less able to participate fully in the discussion resulting in fewer turn-taking opportunities. A field study that involved actual VC meetings with relevant and participant determined topics would be a natural next step in examining the

use of HWAET. A more lively discussion from hearing and deaf participants would likely occur providing an increase in the quantity and perhaps quality of the data.

A third limitation was the use of the laboratory setting. This was a controlled and closely monitored environment in an informal meeting facility and a usability laboratory. I expect that actual meetings take place in a variety of settings that include formal and semi-formal spaces but rarely would a usability facility be available for the remote participant. In an actual setting, the remote individuals (deaf participant or interpreter) might be located in their home or at a location where video remote interpreting is offered. Also, in the laboratory setting it was uncertain whether participant's behaviour was typical meeting behaviour or not, except in one of the experimental group where the heated discussion occurred. A field study where the deaf person can be at their home or in their normal setting and can connect to CHS with access to an interpreter and other hearing participants in a meeting would be desirable. In this case, deaf and hearing participants' behaviour could be considered more representative of actual meeting behaviour.

Another limitation is that only one group configuration was used in this study; the remote person was the deaf participant, and the interpreter and hearing participants were physically located together in a separate location. There are other typical scenarios that should be considered for a more comprehensive analysis of HWAET. Two other possible scenarios are: 1) the interpreter is remote from both parties (hearing person and deaf person physically located together); or 2) the interpreter is physically located with the person who is deaf and the hearing participant(s) are remote. For example, the Canadian Hearing Society (CHS) in Toronto provides remote interpreting services for their staff

who are deaf and who hold meetings with other CHS branches in Northern Ontario. In this service, it is typical for the interpreter and deaf person to be connected remotely through a video conferencing system and all other participants be face-to-face. HWAET can likely aid in allowing the deaf person to be acknowledged and engaged in the meeting with the hearing participants while the interpreter focus on translating. In this scenario, a future study could be used to examine whether the interpreter's role changed similar to that seen in the laboratory study. For example, if the interpreter is not physically in the same room, will meeting behaviours with HWAET be different than when the interpreter is physically present. Each different scenario requires unique considerations regarding the behaviour and perception of the interpreter/deaf person pair. However, regardless of scenario the interpreter and deaf person must have constant eye contact and must be able to see each other's signs at all times, turn-taking using VC is difficult and may require intervention.

A final limitation is that we used a prototype of HWAET with little industrial design. While participants did not generate comments regarding the aesthetics or robustness of HWAET, specific user testing regarding the appearance of HWAET could be carried out. In addition, industrial design, fabrication and electro-mechanical testing such as fatigue and benchmark testing could be carried out on.

CHAPTER 8: CONCLUSION + RECOMMENDATIONS

In this project, a first prototype of an attention-getting device, called HWAET, to assist people who are deaf with turn-taking when using remote interpreting via video conferencing was developed and tested. HWAET represents a physical waving hand that it is positioned in the view of all participants in a video conferencing meeting. The hand can be activated by the remotely located deaf person to obtain the attention of other participants in a meeting. Although the system was designed by a research team my responsibility included developing the user interface and functionalities of the video conferencing application, and conducting and analysing the user studies.

Once HWAET was designed, user studies were required to in order to evaluate the effectiveness and usability of HWAET with deaf and hearing participants and interpreters in VC mediated meetings.

The research questions were thus related to whether HWAET can be practical and acceptable to gain access to the floor by a deaf participant in a VC meeting with hearing participants and whether the interpreter's role changes when using HWAET in meetings.

A review of conversational patterns suggested that turn-taking is of critical importance in any type of communication. Turn-taking is usually accomplished through different cues, such as gestures, posture or eye contact. However, in a mixed conversation with hearing and deaf members facilitated by interpreters, some of these cues can be missed or not used. When VC is used for meetings between geographically separated participants, turn-taking becomes a greater challenge.

User testing with HWAET, VC, deaf and hearing participants and sign language interpreters was carried out to explore the research questions. In these studies, a deaf

person was located in a usability laboratory and the hearing participants, interpreter and meeting chair were located in a different and separate room (meeting room). A comparison between turn-taking effectiveness was carried out between HWAET facilitated meetings and ones without. The deaf person connected to the meeting room through video conferencing system for a twenty-minute discussion on a pre-determined and current topic. Observations and data collected through video tapes and post-study questionnaires indicated that using HWAET was effective in allowing people who are deaf to independently indicate a desire for taking a turn to speak in a video conferencing meeting with hearing people and rely less on an interpreter to facilitate that task. Once HWAET was activated, hearing people acknowledged the deaf person and the turn was given to her once the hearing person finished his turn. The interpreter then was also able to concentrate on her interpreting duties rather than being required to perform meeting management functions as well.

In the study group, deaf participants were asked to use HWAET to indicate a desire for taking a turn and although some people were reluctant to use it at first, as the meeting progressed most of the deaf participants became accustom to HWAET and used it frequently and successfully.

The interpreters reported they did not need to mitigate turn-taking and did not have to keep track as to whose turn was next. In the control group (and in common practice), the interpreter had to point her finger to the next speaker in order to coordinate the turn-taking in the meeting.

HWAET is a first step in providing technology solutions to support independent turn-taking in video conferencing for people who are deaf. It is incredibly important that

people who are deaf have access to interpreters and this can be achieved through the use of video conferencing and HWAET. Turn taking has, however, been identified as a major impediment to equal communication among all group participants. In my project, I investigated the use of an attention getting mechanism in order to overcome this problem. While the scope of this study was limited, it did provide research and analysis that opens up avenues for further research.

8.1 Future Research

Based on the initial user studies, HWAET shows promise as an attention-getting system for VC. However, further evaluation is required to consider more diverse groups of deaf users of different ages and gender, different group configurations, and the system's performance in actual meetings. Different configurations could include having the interpreter be remote from hearing and deaf people who are face-to-face together, or having the interpreter physically located with the person who is deaf and hearing participant(s) are remotely. Actual meeting situations could include using HWAET in settings as hospitals, social worker's offices, homes or CHS.

During the study, deaf participants only used HWAET for a short time, about 20 minutes. Using HWAET over for a longer meeting period and over a long term will eliminate the novelty and the long term effectiveness of HWAET use could be evaluated (e.g., when people become accustom to using HWAET is there a change in the effectiveness of the system?). More study of hearing participants can also demonstrate the effectiveness of HWAET. For example, the post-questionnaire for hearing participants could include additional attitudinal data about any changes in participation of deaf people in meetings over time. Finally, interpreters could also be studied to

understand the impact of HWAET on their role in meetings in detail (e.g., impact on fatigue rates, relationship with deaf person, etc.).

Furthermore, other research questions that might be relevant to study are: 1) If HWAET used in a social setting, will it be usable, because of the familiarity between friends and 2) Is HWAET usable and successful in allowing deaf people request a turn when to speak in different configurations with the interpreter.

Although there is much more work to accomplish in understanding and supporting the communication needs of deaf meeting participants with video conference, I have taken one step towards alleviating the turn-taking process in video conferencing meeting between hearing participants and people who are deaf in remote areas. By using HWAET, deaf people have an opportunity to be acknowledged and participate in VC meetings.

APPENDIX A: VIDEO CODECS

For video conferencing systems, audio and video information is segmented into chunks, encoded and compressed, put into a series of data packets and sent over the network to the remote end at basically constant intervals. The data packets arrive at their destination at varying times, if at all, and often out of order. To keep the "real time" impression of an interactive videoconference, the packets must arrive, on time and in time to be re-ordered for delivery through the videoconferencing terminal (ViDe 2005). This compression/decompression is called codecs and requires high processing power and reliable performance.

Personal computers (PC) are not powerful enough to do high-quality full-screen, full-motion video compression and decompression, especially in addition to all the other applications that a PC might be running at the same time a videoconference is desired. Vendors have integrated hardware-based codecs in their videoconferencing products. These codecs are specifically designed to offload the compression and decompression task from the PC, allowing the endpoint overall to achieve good performance (Vide 2005). The codec is often included in the PCI bus card in a PC, such as Zydacron OnWAN and the VCON Escort Series. These products can support a high frame rate (15 - 30 frames per second) and extended range of call quality/bandwidth settings (128K - 1.5Mb) so that videoconference quality seen at one's desktop can equal that of larger and more expensive room-based conferencing systems (Vide 2005). Most recently, vendors have introduced "plug and play" external devices that plug into a USB port. These USB devices, such as camera include the power to encode and provide the necessary transmission to compress video to pass from the camera to the PC.

H.323 is a communication standard produced by International Telecommunication's Union (ITU). It is an outgrowth of the traditional H.320 technology but optimized instead for the Internet. H.323 has since been revised to include voice-over IP and IP telephony, as well as gatekeeper-to-gatekeeper communications. IP video conferencing uses the Internet protocol and technologies to process and transmits live video and audio signals. The IP video conferencing signals must share the network with all the other Internet traffic resulting in inconsistent and unpredictable quality of audio and video signals. However, because network has become cheaper, IP-based video conferencing systems can transmit from 384 kbps to 768 kbps. There is a need for transmission rate to be higher on IP networks than on ISDN networks, because of packetizing.

APPENDIX B: INFORMATION AND CONSENT FORM

Principal Investigators: Deborah Fels, Ph.D., P.Eng. Ryerson University
(416)-979-5000 ext. 7619 or dfels@ryerson.ca
Bertha Konstantinidis, Ryerson University
(416)-979-5000 ext. 7620 or bkonstan@ryerson.ca

Project Title: Evaluation of attention getting and turn taking in video remote interpreting with deaf users

Information Form

Purpose of Study

The purpose of this study is to evaluate the usability of a robotic hand designed to allow sign language speakers to gain the attention of participants in video conferencing meetings. The system consists of two computers connected together using video conferencing technology. We will simulate the situation by having one meeting participant (interpreter or sign language speaker) in a separate computer room from the rest of their group in a meeting room. The participant in the separate room will see and interact with other participants in the meeting room. Participants will be asked to have a discussion on current topic.

Experiment tasks

The study will be carried out over one, one-hour session at a time convenient to you at The Centre for Learning Technology at Ryerson University. You will first be asked to complete a pre-study questionnaire to collect some background information, your current computer experience, techniques for gaining attention, communication needs for deaf users and remote interpreting. The second part consists of a group discussion on a current topic through video conferencing, where the participant in the separate room participates in the discussion and interrupts five to six times during the discussion to request the speaking floor.

After completing the study, you will be asked to complete a post-study questionnaire to give your opinion and impressions and ease of use of the hand mechanism in the video conferencing. You do not have to answer all of the questions if you do not want to. There will be no penalty if you do not answer all of the questions.

All of the comments you say or sign will be recorded on the video camera focused on the computer screen containing the remote participant. A second interpreter will provide simultaneous interpretation into English so that the researchers and developers who are hearing can benefit from what you say. The timing of what you have said in relation to the video will be measured and recorded.

You will be involved for approximately one (1) hour including the short training session and the completion of the questionnaires. We will be compensating you $30.00 for your time and travel expenses.

Your participation in this research study is entirely voluntary. You may withdraw from the study at any time without penalty. Whether or not you participate will have no effect on services provided to you by Ryerson or on your relationship with or academic status at Ryerson.

Principal Investigators:	Deborah Fels, Ph.D., P.Eng. Ryerson University (416)-979-5000 ext. 7619 or dfels@.ryerson.ca Bertha Konstantinidis, Ryerson University (416)-979-5000 ext. 7620 or bkonstan@ryerson.ca
Project Title:	Evaluation of attention getting and turn taking in video remote interpreting with deaf users

Information Form (cont.)

Confidentiality
All data will remain confidential. We will use number codes to link data with personal information so that people reviewing the data will only see the number identifiers and not the personal information. The data will be securely locked in a storage cupboard at the Centre for Learning Technologies at Ryerson University, and will only be viewed by the project development team. Data will only be presented in summary form and no one individual will be identified. Video-tape data will be destroyed after five years.

Expected Benefits
It is not foreseen that you will personally benefit from participation in this study. However the results from this research will contribute to the development of a system that will improve the ability of deaf users and remote interpreters to participate in meetings using video conferencing. The long-term benefits of such a system may be more consistent and satisfying remote interpreter services for a person who is deaf.

Risks and Discomforts
There are only minimal risks associated with participation in this study involved in the potential discomfort with the amount of effort required to carry out the study tasks of participating in a meeting using video conferencing, and with completing the questionnaires associated with the study. A technician and a researcher will be available at all times during the study to assist with the video conferencing equipment. These data will only be used to analyze the usability of the robotic hand and inform future versions of it. Only members of the project team will have access to all the data, which will be stored in the project office in a locked filing cabinet.

The attached consent form concerns only the study described above and not other normal activities such as attending class at Ryerson University.

Opportunities for Feedback
Copies of any conference proceedings or publications arising from this research will be available in the Ryerson libraries or in the publication section from http://www.ryerson.ca/clt.

We sincerely appreciate your co-operation. If you have any questions or concerns, please do not hesitate to call Deborah Fels at 979-5000 ext. 7619 or Bertha Konstantinidis at 416.979.5000 ext. 7620. In addition to the principal researcher and his supervisor, The Research Ethics Board may also be contacted should there be any complaints or concerns about the project, c/o Office of Research Services, Ryerson University, 350 Victoria St., Toronto, ON M5B 2K3, Tel: 416-979-5042.

Principal Investigators: Deborah Fels, Ph.D., P.Eng. Ryerson University
(416)-979-5000 ext. 7619 or dfels@.ryerson.ca
Bertha Konstantinidis, Ryerson University
(416)-979-5000 ext. 7620 or bkonstan@ryerson.ca

Project Title: Evaluation of attention getting and turn taking in video remote interpreting with deaf users

Information Form (cont.)

Voluntary Participation

I understand that my participation is voluntary and that I have the right to withdraw from this project, or from parts of it, at any time by contacting Deborah Fels or Bertha Konstantinids. Whether or not I participate in this study will have no affect on my future relations with Ryerson University

Principal Investigators: Deborah Fels, Ph.D., P.Eng. Ryerson University
(416)-979-5000 ext. 7619 or dfels@.ryerson.ca
Bertha Konstantinidis, Ryerson University
(416)-979-5000 ext. 7620 or bkonstan@ryerson.ca

Project Title: Evaluation of attention getting and turn taking in video remote interpreting with deaf users

Consent Form to Participate in Study

I acknowledge that the research procedures described above have been explained to me and that any questions that I have asked have been answered to my satisfaction. I have been informed of the alternatives to participation in this study, including my right not to participate and the right to withdraw without penalty. I hereby consent to participate in the study and to be video-taped during the study. I have received a copy of the information sheet.

Signature of Participant: _____

Name of Participant (please print): _____

Date: _____

The details of this study were explained to me by:

Name of Investigator: _____

Date: _____

APPENDIX C: PRE-STUDY QUESTIONNAIRE

Pre-Study Questionnaire

The purpose of this questionnaire is to collect information about your previous experience with computers and video conferencing.

These questions will be presented by an ASL-speaking interviewer.

Part I – Demographics

1. Do you identify yourself as (check one):
   - ❑ Hearing
   - ❑ Hard of hearing
   - ❑ Deaf
   - ❑ Cochlear Implant

2. Are you:
   - ❑ Male
   - ❑ Female

3. Please indicate your age range:
   - ❑ Under 18
   - ❑ 19 – 24
   - ❑ 25 – 34
   - ❑ 35 – 44
   - ❑ 45 – 54
   - ❑ 55 – 64
   - ❑ over 65

4. What is your highest level of education completed?
   - ❑ No formal education
   - ❑ Elementary school
   - ❑ High School
   - ❑ Technical/College
   - ❑ Undergraduate degree from university
   - ❑ Graduate School

Part II – Communication

5. How often do you use a computer?
   - ❑ Never
   - ❑ Occasionally (about 1 per month)
   - ❑ Sometimes (once per week)
   - ❑ Often (a few days per week)
   - ❑ Frequently (every day)

6. How often do you use the internet?
   - ❑ Never
   - ❑ Occasionally (about 1 per month)
   - ❑ Sometimes (once per week)
   - ❑ Often (a few days per week)
   - ❑ Frequently (every day)

7. For what types of activities do you use the Internet: (categories, check as many as apply):

Email	
searching for information	
Video remote interpreting	
Video conferencing	
chat-rooms	
Games	

Other, please specify

_____

8. How often do you attend meetings with hearing people and an interpreter who is physically present?
   - ❑ Never
   - ❑ Occasionally (about 1 per month)
   - ❑ Sometimes (once per week)
   - ❑ Often (a few days per week)
   - ❑ Frequently (every day)

9. What is the most difficult task when attending meetings with hearing people and interpreters (check one)

- ❑ Trying to keep up with the conversation
- ❑ Getting a turn to speak
- ❑ Seeing the interpreter and other people in the meeting at the same time
- ❑ Having people get to the point

- ❑ Other, please specify _____

10. How often do you attend meetings with hearing people without an interpreter at the meeting?

- ❑ Never
- ❑ Occasionally (about 1 per month)
- ❑ Sometimes (once per week)
- ❑ Often (a few days per week)
- ❑ Frequently (every day)

11. How familiar are you with video remote interpreting?

- ❑ Don't know what it is/never used it
- ❑ Know about it but never used it
- ❑ Somewhat familiar (used it a couple of times)
- ❑ Familiar
- ❑ Very familiar (use it daily)

12. What is the most difficult task when attending video conference meetings with hearing people when the interpreter is remote (check one)

- ❑ Trying to keep up with the conversation
- ❑ Getting a turn to speak
- ❑ Seeing the interpreter and other people in the meeting at the same time
- ❑ Having people get to the point
- ❑ Managing the video conferencing equipment

- ❑ Other, please specify _____

- ❑ Never been in this situation.

13. What is the most difficult task when attending video conference meetings with

hearing people when you are remote and the interpreter is with the hearing people (check one).

- ❑ Trying to keep up with the conversation
- ❑ Getting a turn to speak
- ❑ Seeing the interpreter and other people in the meeting at the same time
- ❑ Having people get to the point
- ❑ Managing the video conferencing equipment

- ❑ Other, please specify _____

- ❑ Never been in this situation

14. On average, how difficult is it for you to get a turn to speak during a meeting with hearing people?
    - ❑ Very Difficult
    - ❑ Difficult
    - ❑ Just Right
    - ❑ Easy
    - ❑ Very Easy

15. On average, how difficult is it for you to get attention from other participants?
    - ❑ Very Difficult
    - ❑ Difficult
    - ❑ Just Right
    - ❑ Easy
    - ❑ Very Easy

16. What do you like about video remote interpreting?

17. What do you dislike about video remote interpreting?

APPENDIX D: POST-STUDY QUESTIONNAIRE

Post-Study Questionnaire

The purpose of this questionnaire is to gather your opinion about your experiences with remote video interpreting and the use of the hand mechanism to gain the attention of the group.

These questions will be presented by an ASL-speaking interviewer.

1. In your opinion, how included in the meeting were you?

   - ☐ Not included
   - ☐ Somewhat included
   - ☐ Included

2. Was the hand-mechanism easy to learn how to use?
   - ☐ Very Difficult
   - ☐ Difficult
   - ☐ Just Right
   - ☐ Easy
   - ☐ Very Easy

3. How difficult was it to use the hand mechanism?
   - ☐ Very Difficult
   - ☐ Difficult
   - ☐ Just Right
   - ☐ Easy
   - ☐ Very Easy

4. Rate your level of difficulty with the following communication functions. Use a rating of 1 to 5 where 1 is very hard and 5 is very easy.

	Very Hard (1)	Hard (2)	Not hard, not easy (3)	Easy (4)	Very Easy (5)
Getting attention during the meeting					
Knowing when to take a turn					
Contributing to the conversation					
Getting my ideas across					
Working with the interpreter					

5. Compared with your ability to indicate that you want to take a turn in the conversation with regular video conferencing, rate your ability to indicate you want a turn with the hand mechanism?

- ☐ Much More Confusing
- ☐ Somewhat more confusing
- ☐ Not Different
- ☐ Less Confusing
- ☐ Much Less Confusing

6. Compared with your ability to work with the interpreter, rate your ability to work with the interpreter in the video conference meeting?

- ☐ Very difficult
- ☐ Somewhat more difficult
- ☐ Not more difficult
- ☐ Somewhat easier
- ☐ Much easier

7. What problems did you experience with using the hand mechanism?

8. Do you have any suggestions for ways to improve video conferencing for sign language users?

9. Other Suggestions?

REFERENCES

American with Disabilities Act. (2003). *US Department of Justice*, 2003, from http://www.usdoj.gov/crt/ada/adahom1.htm.

Baker, C. (1977). Regulators and turn-taking in American Sign Language discourse. In L. Friedman (ed.), *On the Other Hand* (pp. 215-236). New York: Academic Press.

BlackBerry™. (2006). *Our Devices*. Retrieved May 20, 2006, 2006, from http://www.discoverblackberry.com/

Brightline. (2002). *Bsic Principles of video-conferencing lighting*: Brightline LP.

Clark, H., & Schaefer, E. (1989). Contributing to discourse. *Cognitive Science, 13*(2), 259-292.

Coates, J., & Sutton-Spence, R. (2001). Turn-taking patterns in deaf conversation. *Journal of Sociolinguistics, 5*(4), 507-529.

Cokely, D., & Baker, C. (1980). *American Sign Language*. Silver Springs Mary-land: T. J. Publishers.

Elliot, R., Glauert, J. R. W., Kennway, J. R., & Marshall, I. (2000). *The development of language processing support for the ViSiCAST project.* Paper presented at the ACM SIGACCESS Conference on Assistive Technologies, New York, New York.

Fels, D., & Weiss, T. (2000). Towards determining an attention getting device for improving interaction during video-mediated communication. *Computers in Human Behaviour, 16*(2), 99-122.

Fels, D., & Weiss, T. (2001). Video-mediated communications in the classroom to support sick children: A case study. *International Journal of Industrial Ergonomics, 28*, 251-263.

Foulds, R. A. (2004). Biomechanical and perceptual constraints on the bandwidth requirements of sign language. *IEEE Transactions on Neural Systems and Rehabilation,*12(1), 65-72.

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

Grayson, D. M., & Monk, F. A. (2003). Are you looking at me? Eye contact and desktop video conferencing. *ACM Transactions on Computer-Human Interaction, 10*(3), 221-243.

Heath, C., & Luff, P. (1991). *Disembodied conduct: Communication through video in a multi-media office environment.* Paper presented at the Conference on Human Factors in Computing Systems, New Orleans, Louisiana.

Herreweghe Van, M. (2002). Turn-taking mechanisms and active participation in meetings with deaf and hearing participants. In C. Lucas (Ed.), *Turn-taking Fingerspelling, and Contact in Signed Languages* (Vol. 8). Washington, D. C.: Gallaudet University Press.

Isaacs, E., & Tang, J. (1993). *What video can and can't do for collaboration.* Paper presented at the Proceedings of the First ACM International Conference on Multimedia, Anaheim, California.

Kendon, A., & Ferber, A. (1973). A description of some human greetings. In R. Michael & J. Crooks (Eds.), *Comparative Ecology and Behaviour of Primales* (pp. 591-668). London: Academic Press.

King, C. M., Dellon, J., & Murray, J. (2003). *Gallaudet University Digital Video Applications*. Retrieved June 27, 2003, from http://academictech.gallaudet.edu/events/2003/ntid

Kraut; R., Fish, R., Root, B., & Chalfonte, B. (1993). *Informal communication in organizations.* Paper presented at the Groupware and Computer Supported Co-operative Work, San Mateo, CA.

Kraut, R., Lewis, S., & Swezey, L. (1982). Listener responsiveness and the co-odination of conversation. *Journal of Personality and Social Psychology*(43), 718-731.

Mather, S. (1996). Initiation in visually constructed dialogue: Reading books to three to eight-year old students who are deaf and hard-of-hearing. In C. Lucas (Ed.), *Multicultural Aspects of Sociolingusitics in Deaf Communities* (pp. 109-131). Washington, D.C.: Gallaudet University Press.

McAteer, E. (2000). *Video Conferencing Guidelines*, Feb. 27, 2004, from http://www.gla.ac.uk/lncs/main/video.htm

Miles, & Huberman. (1994). *Qualitative Data Analysis*: Sage Publishing.

Molnar, J. P. (1969). Picturephone Service - A New Way of Communicating. *Bell Laboratories: Record, 47*(5), 134-135.

O'Connaill, B., Whittaker, S., & Wilbur, S. (1993). Conversations over video conferences: An evaluation of the spoken aspects of video mediated communication. *Human Computer Interaction, 8*(4), 389-428.

Polycom, I. (2005). *White Papers*. Retrieved Feb. 26, 2004, from http://www.polycom.com/common/pw_item_show_doc/0,667,00%2BenUSS_01DBC.pdf

Powell, K. (2005). *Hwaet.net*. Retrieved Sept. 24, 2005, from http://hwaet.typepad.com/

Reagan, T. (1995). A sociocultural understanding of deafness: American Sign Language and the culture of deaf people. *International Journal Intercultural Relations*(19), 239-251.

Roy, C. (1989). *A sociolinguistics analysis of the interpreter's role in the turn exchanges of an interpreted event.* Georgetown University, Washington, D. C.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organisation of turn-taking in conversation. *Language, 50*, 696-735.

Siple, P. (1978). Visual constraints for sign language communication. *Sign Language Studies, 1*(9), 95-110.

Sorenson, C. (2005). Sorenson VRI.

Stokeo, W. C. (2001). The study and use of sign language. *Sign Language Studies, 1*(4), 369-406.

Swisher, V., Christie, K., & Miller, S. (1989). The reception of signs in peripheral vision by deaf persons. *Sign Language Studies, 6*(3), 99-125.

Van Herreweghe, M. (2002). Turn-taking mechanisms and active participation in meetings with deaf and hearing participants. In C. Lucas (Ed.), *Turn-taking Fingerspelling, and Contact in Signed Languages* (Vol. 8). Washington, D. C.: Gallaudet University Press.

Vertegaal, R. (1999). *The Gaze groupware system: Mediating joint attention in multiparty communication and collaboration.* Paper presented at the Conference on Human Factors in Computing Systems, Pittsburgh, Pennsylvania.

Vertegaal, R., Van der Veer, G. C., & Vons, H. (2000). *Effects of gaze on multiparty mediated communication.* Paper presented at the Proceedings of Graphics Interface, Montreal, Quebec.

Vertegaal, R., Weevers, I., & Sohn, C. (2002). *GAZE-2: An attentive video conferencing system.* Paper presented at the Conference on Human Factors in Computing Systems, Minneapolis, Minnesota.

Video Remote Interpreting Adhoc Committee . (2005). *Emerging video interpreting and video relay; standards and regulations*. Retrieved Nov. 20, 2005, from http://www.rid.org/VIC_SPP11_05.doc

Walker, M., & Whittaker, S. (1990). *Mixed initiative in dialogue: An investigation into discourse segmentation.* Paper presented at the Proceedings of 28th Annual Meeting of the Conference on Computational Linguistics, Pittsburgh, Pennsylvania.

Whittaker, S., Frohlich, D., & Daly-Jones, O. (1994). *Informal communication: What is it like and how might we support it?* Paper presented at the Conference on Human Factors in Computing Systems, Boston, Massachusettes.

Whittaker, S., & O'Connaill, B. (1997). The role of vision in face-to-face and mediated communication. In S. Wilbur, K. Finn & A. Sellen (Eds.), *Video-Mediated Communications* (pp. 23-49): LEA Press.

Whittaker, S., & Stenton, P. (1988). *Cues and control in expert-client dialogues.* Paper presented at the Proceedings of the 26th Annual Meeting of the Association of Computational Linguistics, Buffalo, New York.

Yarger, C. C. (2001). Educational interpreting: Understanding the rural experience. *American Annals of the Deaf, 146*(1), 16-30.

Yngve, V. (1970). *On getting a word in edgewise.* Paper presented at the Proceedings of the Sixth Meeting of the Chicago Linguistics Society, Chicago.