

Supporting Medical Conversations between Deaf and Hearing Individuals with Tabletop Displays

Anne Marie Piper and James D. Hollan

Distributed Cognition and Human-Computer Interaction Laboratory
Department of Cognitive Science, University of California, San Diego
9500 Gilman Dr., La Jolla, CA 92093
{apiper, hollan}@ucsd.edu

ABSTRACT

This paper describes the design and evaluation of Shared Speech Interface (SSI), an application for an interactive multitouch tabletop display designed to facilitate medical conversations between a deaf patient and a hearing, non-signing physician. We employ a participatory design process involving members of the deaf community as well as medical and communication experts. We report results from an evaluation that compares conversation when facilitated by: (1) a digital table, (2) a human sign language interpreter, and (3) both a digital table and an interpreter. Our research reveals that tabletop displays have valuable properties for facilitating discussion between deaf and hearing individuals as well as enhancing privacy and independence. The contributions of this work include initial guidelines for cooperative group work technology for users with varying hearing abilities, discussion of benefits of participatory design with the deaf community, and lessons about using dictated speech on shared displays.

Categories and Subject Descriptors

H.5.3 [Information Interfaces and Presentation (e.g., HCI)]: Group and Organization Interfaces - *Computer-supported cooperative work*.

General Terms

Design, Experimentation, Human Factors.

Keywords

Computer-supported cooperative health care, multitouch, tabletop groupware, assistive technology, deafness, multimodal interfaces, speech recognition.

INTRODUCTION

This paper presents the design and evaluation of Shared Speech Interface (SSI), an application for an interactive tabletop display that enables and supports communication between a deaf patient and a hearing, non-signing medical doctor. Currently, medical facilities provide a sign language

interpreter to facilitate communication between a deaf patient and hearing doctor. Deaf patients must plan ahead to ensure that an interpreter is available. For most patients, privacy is a central concern. Some deaf patients, depending on their comfort level with interpreters, prefer and actually use an alternate communication channel (e.g., email or instant messenger to discuss sensitive medical issues with their physician). Increasing communication privacy is one motivation behind the work reported here.

While other viable communication tools for the deaf community exist, tabletop displays with speech recognition have potential to facilitate medical conversations between deaf and hearing individuals. Consultations with physicians often involve discussion of visuals such as medical records, charts, and scan images. Interactive tabletop displays are effective for presenting visual information to multiple people at once without necessarily designating one person as the owner of the visual. Taking notes while meeting with a physician is problematic for deaf individuals because it requires simultaneously attending to the doctor's facial expressions, the interpreter's visual representation of speech, and notes on paper. A tabletop display allows all active participants to maintain face-to-face contact while viewing a representation of conversation in a central location. Our implementation incorporates keyboard input by the patient and speech input by the doctor, allowing the physician to speak and gesture as they discuss medical details and visuals with the patient. SSI leverages the affordances of multimodal tabletop displays to enhance communication between a doctor and patient, potentially transforming a challenging situation into a constructive and collaborative experience.

Our work on SSI provides practical experience designing a shared communication device for users with varying hearing and speaking abilities. We examine the challenges of representing speech visually to multiple users around a tabletop display. We also provide design guidelines for using dictated speech with shared display systems and discuss the implications of our work on multimodal tabletop displays for the broader CSCW community.

BACKGROUND

Loss of hearing is a common problem that can result from noise, aging, disease, and heredity. Approximately 28 million Americans have significant hearing loss, and of that group, almost six million are profoundly deaf [17]. A

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CSCW'08, November 8-12, 2008, San Diego, California, USA.

Copyright 2008 ACM 978-1-60558-007-4/08/11...\$5.00.

primary form of communication within the United States deaf community is American Sign Language. ASL is not a visual form of English; it is a different language with its own unique grammatical and syntactical structure. Sources estimate that ASL is the fourth most commonly used language in the U.S. [17]. While ASL is widely used in the U.S., no one form of sign language is universal. Different countries and regions use different sign languages. For example, British Sign Language is different from American Sign Language, although both countries have English as their official and primary spoken language.

Within the deaf community, as with most communities, there is great variability among individual needs and abilities. Individuals who were born deaf may or may not have been raised in the aural tradition where they were taught to speak, read, and write in English. Others who suffer from late-onset hearing loss typically have fully developed vocal abilities but are not fluent in ASL and do not know how to read lips. This range of individual abilities and needs has led to a gamut of techniques for communicating with and adapting to the hearing world.

Adaptive Techniques

For the deaf population proficient in a spoken language such as English, writing has long been a central form of communication with the hearing world. Deaf individuals without an interpreter nearby often use handwritten notes and deictic gesturing as a means of communication. Telephone use was impossible for the deaf community until the invention of the Teletype and Text Telephone (TTY), a typing-based system that transmits individual lines of keyboard entry over phone lines. Adoption of the TTY and eventually the personal computer made typing an essential mode of communication within the deaf community. In recent years, the invention of webcams and increasing Internet bandwidth gave rise to communication through videochat with other ASL speakers and ASL interpreters.

ASL interpreters play a central role in enabling face-to-face communication between many deaf and hearing individuals. For the deaf population fluent in ASL, communicating through an interpreter is an optimal choice for many situations. Interpreters, however, are expensive and not always available. Furthermore, though interpreters are bound by a confidentiality agreement, the presence of a third person in a highly private conversation may reduce a deaf person's comfort and inhibit their willingness to speak candidly.

Related Work

Researchers have developed a variety of technologies to address communication barriers between the deaf community and hearing world. Researchers investigating tabletop technologies traditionally explore cooperative group work for only hearing populations and have yet to examine the value of tabletop displays for deaf populations.

Technologies for the Deaf

As early as 1975, researchers began investigating how cooperative computing environments, such as early forms of instant messenger, could facilitate communication between deaf and hearing individuals [27]. More recently, HCI researchers have examined how mobile devices, tablet computers, and video conferencing technologies can augment communication for deaf individuals. Schull investigated communication via a browser-based client on multiple co-located laptops that allows a deaf and hearing user to access a common browser window and share real-time chat information [21]. iCommunicator [7], a commercial product, enables communication in a similar way. Only a handful of initiatives, however, have attempted to facilitate shared face-to-face communication experiences. The majority of work has focused on single-user interfaces for distributed applications. For example, MobileASL enables two signing individuals to communicate with each other over cell phones with real time video [1]. Scribe4Me, a mobile sound translation tool, enables deaf individuals to request a transcription of the past 30 seconds of audio in their environment [10]. There is also work using peripheral displays to visualize various channels of auditory information for deaf individuals [11]. The Facetop Tablet project examined visual attention problems that deaf individuals experience when trying to view a presentation, watch an ASL interpreter, and take notes [13]. This project presents a viable approach for helping deaf individuals follow a conversation with hearing individuals but it does not explicitly help deaf individuals communicate or become active participants in the conversation. Past research has also examined enhancing communication for deaf individuals through gesture recognition with computer vision and wearable computers (e.g., [24]).

While these solutions address various communication challenges for deaf individuals, none provide a shared communication experience that bridges both speaking and listening barriers between deaf and hearing individuals. Furthermore, many of these communication technologies only provide an interface and feedback to one user. Our system, on the other hand, leverages the cooperative nature of interactive tabletop displays to enable a shared, co-constructed communication experience between users with varying hearing and speaking abilities.

Tabletop Displays

The field of CSCW has a rich history of research on tabletop displays and their utility for supporting group work. This body of research examines cooperative group work around tabletop displays with hearing populations and is the foundation for our research. We leverage techniques from research on social protocols around digital tables [16], the notion of shared and private spaces on tabletop displays [22], and how to present textual, pictorial, and auditory information to users [15][19].

Research on multimodal tabletop interaction, specifically integrating speech with touch input, also influenced our work. Work by Tse et. al. involving speech commands in group design and gaming experiences is a primary example [25][26]. This work focuses on facilitating interactions between hearing individuals but does not examine the use of multimodal tabletop displays for non-hearing populations.

DESIGN PROCESS

We employed a participatory design process [20], involving members of the deaf community and domain experts in all aspects of the design and evaluation of SSI. When designing for special needs populations, it is critical to gain community support and involve domain experts and community members from the beginning [6][18]. As we prototyped SSI, we concurrently met with deaf individuals, linguists studying deaf culture and sign language, and medical professionals who communicate regularly with deaf individuals. In fact, the idea for SSI came from a medical doctor. We began by conducting interviews to investigate current technologies for communication and problems with existing technologies.

Communication Challenges

As previously mentioned, sign language interpreters are an important link enabling communication between deaf and hearing individuals. The physical presence of an interpreter, however, may inhibit a deaf patient from candidly discussing private medical issues. Furthermore, other problematic aspects of communicating through interpreters include accuracy of interpretation and challenges for the deaf individual when attending to multiple channels of visual information.

Karen¹ is a Professor of Communication who studies deaf culture and sign language. She was born deaf to two deaf parents but was raised in the aural tradition by attending hearing schools. Karen has adapted exceptionally well to the hearing world by reading lips and developing her vocal abilities. She does not need an interpreter for most situations, but in an interview mentioned her deaf husband's privacy concerns, "My husband for example, lip reads his doctor... he doesn't want an interpreter. If he needs an interpreter he wants me to come." Karen describes another situation and reiterates the need for a more private and independent communication medium:

"I know one situation where a therapist needed to see a deaf person on an emergency situation and [the patient] wasn't comfortable going through an interpreter. So they had two TTYs. They actually took the TTYs and were typing back and fourth. But the problem is it only had one line at a time. It was hard to remember what the issue was if you kept losing it."

¹ All names were changed to preserve anonymity.

When an interpreter is provided by a medical facility, the deaf individual usually does not have a choice about who will interpret. The deaf community is well-connected in our city, and one deaf woman said that she would not feel comfortable going through an interpreter that she knew well or going through a male interpreter. She emphasized the need for a better option when a deaf patient is not comfortable with the interpreter.

Beyond issues of privacy and autonomy, ensuring accuracy of communication with an interpreter is critical. It is often the case that interpreters are not experts on the content they are interpreting. For example, one deaf individual described having an interpreter in Biology class who knew little about Biology. This made understanding the interpretation extremely difficult and impacted her ability to learn in the classroom. When it comes to health care issues, receiving accurate and full interpretation is essential. Dr. Stevens is a physician and professor at our university hospital. She established a program to teach medical students about deaf culture and how to create a deaf friendly medical environment. Dr. Stevens comments on the challenges of relaying information through an interpreter:

"We're assuming that interpreters have a lot of medical information, and they may not. They may be miscommunicating. I always think about the doctor who told the patient, take three of these a day, and the interpreter didn't know to explain take one morning, lunch, and dinner."

Facilitating conversation with an interpreter also creates challenges for deaf individuals to share their attention between the conversation and note-taking. Dr. Stevens explains:

"With an interpreter... the patient has no ability to make a record of what's being said because the eyes are on the interpreter and they can't be on your paper, on the interpreter, on the doctor getting your emotions. What's your body language saying? You're showing me my knee but I'm looking at your face to see how bad this really is... So most of us go home with notes written down, but if you're deaf, it's hard to do both."

The goal behind our design for SSI is to explore an alternative to human interpreters as well as to augment conversation that occurs through an interpreter.

Technical Implementation

We prototyped SSI on a MERL DiamondTouch table [4] using the DiamondSpin toolkit [23]. The DiamondTouch table is a multiuser, multitouch top-projected tabletop display. Users sit on conductive pads that enable the system to uniquely identify each user and where each user is touching the surface. Our system enables conversational input through standard keyboard entry and a headset microphone. The audio captured from the microphone is fed into a speech recognition engine (currently this is Microsoft

Windows' default recognizer, but our application is easily adapted to any off-the-shelf recognizer). SSI uses the Java Speech API [8] and CloudGarden software [3] to interface with the speech recognition engine and send converted speech-to-text into the main application running on the DiamondTouch table.

Because understanding and analyzing conversation becomes increasingly complex as the number of speakers increases, we decided to implement the first version of SSI for two users only, one hearing and one hearing-impaired user. We wanted to investigate whether tabletop displays are an effective communication tool for this basic case of two users prior to expanding to larger group work scenarios. The discussion section at the end of this paper provides ideas for expanding and adapting the SSI technology for more than two users.

Interface Design

We chose the form factor of a tabletop display because it allows the doctor and patient to maintain face-to-face communication. For the deaf patient, reading and making facial expressions is an important communication channel. For the doctor, eye contact with the patient reveals the patient's understanding and emotional state. A wall mounted display or even a traditional computer monitor would not enable face-to-face interaction in the same way; however, smaller horizontal devices such as a shared tablet computer may also be worth exploring.

We explored multiple tabletop interface designs before proceeding with the design tested in the evaluation section. From conversation research, we know that orientation and position of speakers is an important aspect of communication. Certain positions afford easier eye contact and information sharing, especially when speakers are positioned around a table. Sitting across from someone gives direct access to facial expressions but makes sharing textual information problematic, although some solutions for reorienting and rotating text have been explored [19]. While sitting side-by-side makes sharing textual information much easier, the critical activities of making eye contact and reading lips becomes challenging. We chose to seat users at an angled position (see Figures 4 and 6) to facilitate eye-contact, lip reading, and information sharing.

As each person contributes to the conversation, either by speaking into a microphone or typing on the keyboard, their speech appears on the tabletop display in front of them. We refer to these fragments of conversation as "speech bubbles." Speech bubbles are color-coded by user and moveable around the display. In ASL, speech has a visual and spatial element, and signers often point back to where a previous gesture occurred. The moveable nature of speech bubbles in our design is an analog for the spatial nature of ASL signs. Seating users at angled positions makes uniformly orienting speech bubbles towards the bottom of the display a natural choice. We deferred the decision of the specific locations

where speech should appear on the display until we had feedback from our initial prototype. In the first version (see Figure 1) we presented speech bubbles close to the user who entered the speech but not in an ordered fashion. We wanted to examine how people would organize speech bubbles on the display and determine whether preserving turn-taking is important.

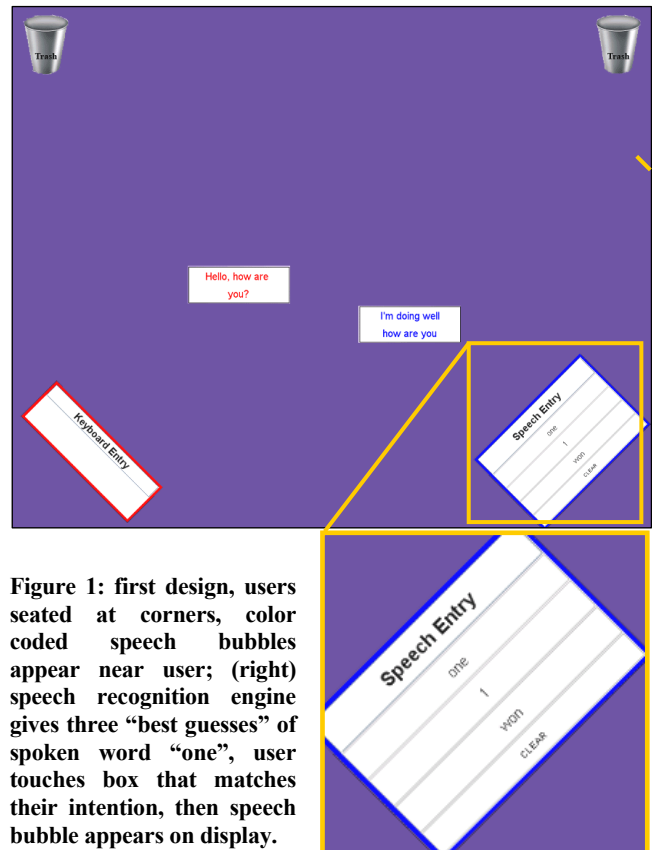


Figure 1: first design, users seated at corners, color coded speech bubbles appear near user; (right) speech recognition engine gives three "best guesses" of spoken word "one", user touches box that matches their intention, then speech bubble appears on display.

With our design, both users could have easily entered conversation through keyboards. We hypothesized that it was important for the physician to speak naturally to the patient. The patient could then attend to the physician's body language and read their lips. The doctor's facial expressions and body language are masked when speech is entered through a keyboard. As noted by Dr. Stevens, reading the physician's body language is an essential part of communication in medical conversations.

While speech recognition engines are constantly improving, transcribing natural language into text is still problematic. We wanted the speaking user to be able to control the speech that the system displays, so SSI provides the speaking user with three "best guesses" of their speech from the recognition engine (see Figure 1). The user touches the phrase that matches their intended speech and the phrase appears on the interface as a speech bubble. While this requires an extra step for the speaking user, it allows greater fidelity in displayed speech and prevents unintended

conversation from appearing on the display. Users can delete previous parts of conversation by dragging speech bubbles to the trashcans in the top corners of the display. In general, we wanted to allow users to correct miscommunications and remove unwanted speech.

Prototype Review

As part of our participatory design process, eight people reviewed our design, four of whom are deaf and the other four are actively involved in the deaf community. An interpreter was present to facilitate communication between the deaf participants and hearing participants and researchers. Our questions at this stage of design addressed user position at the table, text size and color, speech bubble behavior, and the use of keyboard entry. This preliminary feedback revealed several key issues about the design of SSI. We discuss these and explain how they influenced subsequent designs and evaluation.

Overall User Interface Design

We found that positioning users at an angled position around the table works well. One deaf person said, “It feels too much like teacher and student when people sit across from each other.” Maintaining eye contact and reading shared speech bubbles was not a problem with the current configuration. Reviewers said that the text size and colors worked well, but because the speech bubble text is fairly large, the display cluttered quickly. Displaying medical images in the background was also received positively, especially for stepping through slides of an MRI or indicating fracture points in an x-ray image (see Figure 2).



Figure 2: Navigation tabs at top, speech bubbles are colored and offset by speaker and appear in order, conversation visuals are displayed behind all speech bubbles.

Organizing Conversation, Managing Clutter

In our first design, the interface became cluttered quickly, even with a trashcan to delete unwanted conversation. Reviewers said the trashcan was useful for correcting

misunderstandings and removing unwanted speech bubbles (especially large speech bubbles). However, there was still a need to organize conversation and several people suggested being able to create a new page.

We considered zooming and scrolling interfaces to increase screen real estate, but instead we created a simple tabbed design that leveraged users’ knowledge of web browsers (see Figure 2). One deaf man said, “The tabs are a great way to know what you’re going to work on and how you’re going to move forward. You can go back and refer to something without having to search for it...the tabs are a nice way to do that.” In response to his comment, a deaf woman said, “exactly, I think it’s good because it’s not overwhelming...it’s very deaf friendly. It’s very visual.” Overall, reviewers liked the tabbed design and found it easy to understand, so we proceeded with this design.

In our first design, we also presented speech bubbles at an arbitrarily location in front of the user who contributed the speech. Reviewers said that it was difficult to anticipate where speech would appear and that it is important to preserve turn-taking. We modified our design to preserve turn-taking by presenting speech bubbles in a linear pattern, offset and color coded to indicate speaker (see Figure 2).

Pre- and Post-Discussion Use

Dr. Stevens explained that doctors often know several key discussion points before they enter a meeting with a patient. She said that it would be extremely helpful to have the nurse enter talking points to structure and pace the conversation. These could be available on the doctor’s side of the interface. Dr. Stevens also mentioned that it would be good to add labels to the tabs to show the topics for discussion (e.g., “welcome” and “update since last visit”). Reviewers unanimously wanted the dialogue to persist after the appointment. They mentioned saving the conversation for the next visit, printing it to take home and share with family, and receiving a copy of it via email.

Diversity within the Deaf Community

One critical finding that came out of our prototype review is that there is great diversity within the deaf community and that our system would work better for certain subpopulations. Members of the deaf community thought that SSI would work well for deaf individuals who feel comfortable using English and for individuals who are Hard of Hearing. They said SSI would be problematic for deaf individuals with low English literacy or low confidence in English communication.

Design Modifications

Based on feedback received in the preliminary evaluation, we kept the text size and font the same, added a feature to display speech in an ordered fashion that preserves turn-taking, and proceeded with the tabbed interface design. We made a slight modification to the look of the speech bubbles by adding a tail on each bubble that indicates the user who added it. After testing speech recognition with multiple

people, including medical professionals, we added a button to bring up a virtual keyboard so the doctor could type certain words that are difficult for the voice recognition system to recognize.

EVALUATION

Our evaluation examines the role of tabletop displays in facilitating medical conversations between a deaf patient and a hearing, non-signing doctor. We present a user study that compares communication through the digital table, an interpreter, and both the digital table and an interpreter.

Methodology

We conducted a laboratory experiment with eight deaf participants (mean age=33, stdev=11.4, range=[22,52]; 3 males) and one medical doctor (age=28, female). All eight deaf participants were born deaf or became deaf before the age of one. Three participants identified English as their native language and five identified ASL. All participants were fluent in ASL and proficient at reading and writing in English. Each deaf participant conversed with a medical doctor and a professionally trained ASL interpreter about a sample medical issue. We chose to have each deaf participant work with the same doctor. This resembles the real-world scenario where one doctor has similar conversations with multiple patients throughout the day. None of the participants had used a tabletop display prior to participating in this evaluation. Our evaluation was conducted in our university laboratory around a DiamondTouch table. Each session was video taped by two cameras from different angles to capture participants' interactions with each other and the digital table. The computer unobtrusively recorded all user interactions with the tabletop surface. Three researchers were present for the testing sessions and took notes.

Procedure

The doctor trained the speech recognition engine prior to the first testing session. At the beginning of each testing session, the deaf participant and doctor performed a brief period of training together to get to know each other and adjust to the task. Then the patient and doctor discussed a medical issue using SSI (Digital Table), a human interpreter (Interpreter), and both SSI and the interpreter (Mixed). At the beginning of each conversation, the deaf participant received a discussion prompt in English text that described a medical topic (e.g., nutrition). Each discussion prompt had a corresponding medical visual. For consistency, the experimenter manually preloaded medial visuals into the system under the third tab. The other two tabs were blank conversation space. A paper version of the visual was provided for the Interpreter condition. Medical professionals worked with us to ensure that the discussion prompts reflected authentic conversations that might occur in a normal patient interaction but whose content did not require participants to discuss information that might be too personal. Both the order of conditions and discussion prompts were randomized between subjects. An experimenter ended each conversation after 8 minutes (based

on the length of actual medical consultations). After completing the three conditions, the deaf participant completed a survey about their experience.

Analysis

After the experiment, two researchers reviewed videos of the conversations (24 total, approximately 200 minutes of talking) and examined interaction based on transcription techniques described by Jefferson [8] and McNeill [12]. The video data revealed extremely rich and complex multimodal interactions between the patient, doctor, interpreter, and digital table. These interactions merit a separate analysis and are topics for future work. In this paper, we characterize high-level themes relevant to the design of SSI and summarize survey results.

Findings

Overall, participants indicated that digital tables are a promising medium for facilitating medical conversations. We observed a rich use of gesture to augment communication in the Digital Table condition. Survey data indicated that our application was good for private conversations and enabled independence. However, the interaction overall is limited by the technology and certain aspects of communication were lost in practice. Specifically, imperfections in the speech recognition engine made conversation in the Digital Table condition considerably slower than in the Interpreter condition.

Conversation and Gesture Analysis

There were several key differences in communication between the Digital Table and Interpreter conditions. The Digital Table condition allowed for asynchrony in communication, whereas the interpreter acted as a broker of conversation and thus encouraged synchronous interactions. Dialogue in the Interpreter condition was verbose and elaborated, while speech in the Digital Table condition was more concise and typographic in nature. We observed equitable participation levels in the two conditions, the doctor and patient each contributed to about half of the conversation.

Slower conversation in the Digital Table condition was likely due to problems with the speech recognition process. The system took one second to determine and display the speech in textual form. Then the interface required the doctor to tap on the phrase she wanted to add. The best speech recognition result occurred when the doctor broke her natural speech into short phrases, but this also slowed communication. On some occasions the doctor made two or three attempts before the system accurately recognized her speech. While speech recognition was problematic, it provided ancillary benefits such as allowing the doctor to gesture while speaking and the patient to read the doctor's lips and facial expressions.

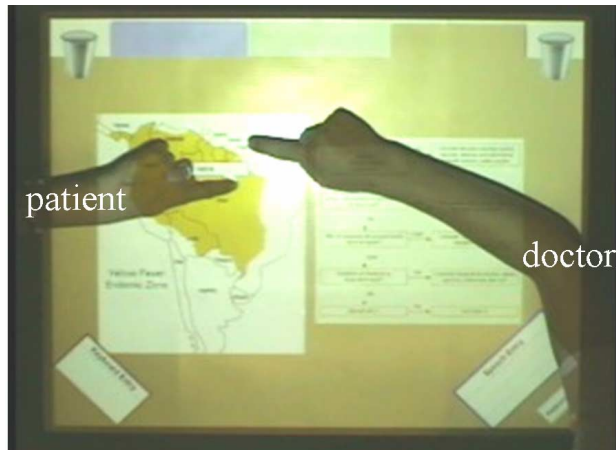


Figure 3: Digital Table condition, doctor circles region on the map and asks “will you be here,” patient responds by touching the “here” speech bubble and then Brazil.

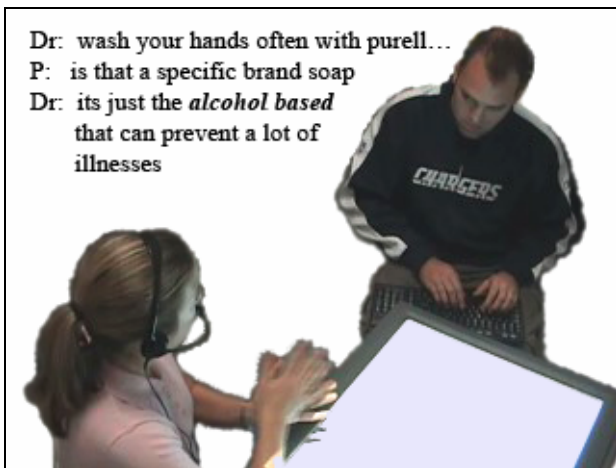


Figure 4: Digital Table condition, doctor and patient share gaze while doctor pantomimes hand washing. Text emphasis shows gesture timing.

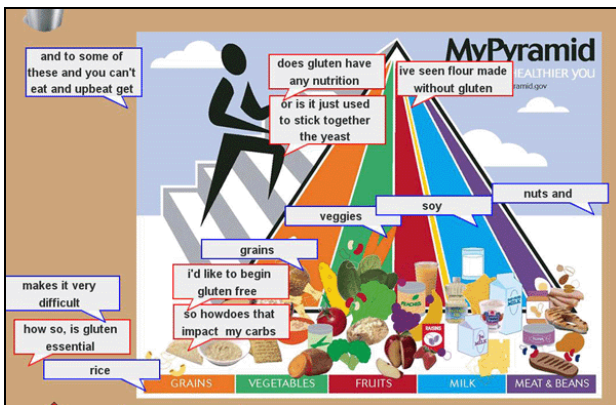


Figure 5: Screenshot from Digital Table condition, doctor and patient use speech bubble tails as pointing mechanism for a conversation about gluten. Gluten related speech bubbles are placed around the grains section of the pyramid.

Gestural and Nonverbal Communication. In the Digital Table condition, we observed that non-verbal and gestural communication played an important role in augmenting and ensuring successful communication. Importantly, the co-located, face-to-face nature of the digital table allowed participants to provide feedback to their partner about their state of understanding through deictic gesture (e.g., pointing), gaze sharing, and head nodding. Participants strategically moved speech bubbles in front of the other user to get their attention and pointed between bubbles to make a connection to previous speech [2]. We also noticed a pattern in which one participant moved or pointed to an object on the interface and then one or both participants nodded to confirm their understanding. This pattern occurred frequently in the Digital Table condition. Figure 3 illustrates an example where participants point together. Video of this instance reveals that the gesture is accompanied by gaze sharing and head nodding. Observations also revealed the use of iconic gestures to augment speech (in one case the doctor pantomimes hand washing, see Figure 4). The doctor, more so than the patient, used iconic or illustrative gestures to support her speech.

The role of gesture is especially relevant to cooperative work because it indicates that participants were iteratively refining and confirming their conversational understanding and engaging in highly coordinated activity through verbal and nonverbal channels. Our design enabled this interaction through the face-to-face design, horizontal form factor, moveable speech bubbles, and voice recognition, freeing the doctor’s hands and thereby enabling co-occurring gesture-speech.

Affordances of Digital Space. The digital table transformed the ephemeral nature of speech into a tangible and persistent form, thus creating affordances that are not available in traditional conversation. We observed interesting behaviors with the speech bubbles because of their form. When a phrase was added to the display that referred to a previous utterance, the “owner” of the speech bubble often moved the new phrase close to the previous utterance. In conversation, the speaker must help listeners understand a reference to a previous utterance through context and explicit referencing. The digital table allowed users to reference previous conversation by placing new speech near an existing speech bubble. Similarly, we observed the doctor and patient using the tail of the speech bubble as a pointing mechanism. That is, participants strategically placed speech bubbles around the display so that the tail of the speech bubble pointed to part of a background visual (see Figure 5). The persistent nature of speech with the digital table allowed participants to review their conversation. We observed both the doctor and patients looking back over their previous conversation. The doctor said “it was good to look back at what I had covered with that particular patient,” and explained that “it would be helpful because it is not uncommon in medicine to have very similar conversations with different patients throughout the

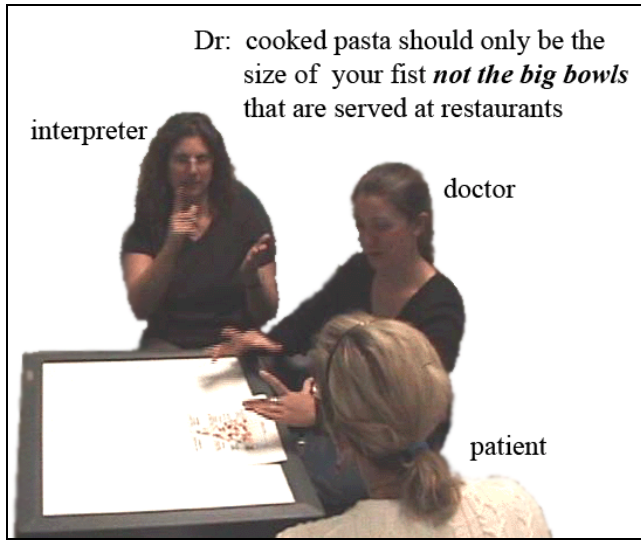


Figure 6: Interpreter condition, patient watches simultaneous gestures by doctor and interpreter.

day.” As Figure 5 illustrates, the speech bubbles occluded other objects on the display. Addressing issues of layering and transparency is an area for future work on SSI.

Presence of an Interpreter. We noticed differences in how patients attended to the doctor when the interpreter was present. Deaf participants looked at the doctor when they signed but then shifted their gaze to the interpreter when the doctor began speaking. In the Digital Table condition, participants typically looked at the doctor when she was speaking and then looked down at the display. In the Mixed condition, we observed a pattern in which the patient watched the interpreter sign and then looked down at the display to read the English version. Several participants explained that seeing the doctor’s speech in both ASL and English was helpful. The interpreter also found benefit in having the digital table present:

“What was nice for me as the interpreter was to have the printed word. When I didn’t know how to spell something, especially in medical situations, if the printed word is there...I can point to it...so that about the communication board is very attractive.”

The patient faced visual attention challenges when both the interpreter and doctor were gesturing at the same time. Figure 6 shows an example of this interaction.

Communication Preferences

After experiencing the above conditions, each deaf participant completed a communications preferences survey. Figure 7 summarizes survey responses.

Privacy. Six of eight deaf participants reported that the digital table alone was best for private medical conversations. Cathy said, “the digital table is best for very private conversations, but using an interpreter in a private conversation depends on whether or not I know the interpreter.” Sharon explained, “for other meetings like a work situation or job interview, I

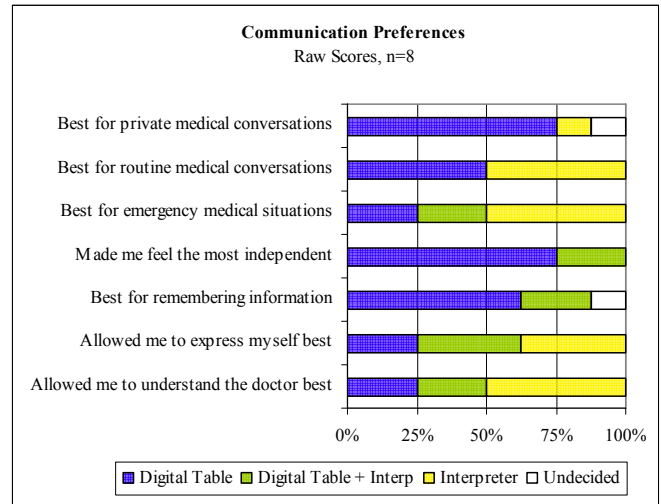


Figure 7: Survey of communication preferences.

would prefer to have an interpreter. But for personal meetings, like with a lawyer, doctor, or specialist, I prefer the digital table.” Jesse also said the digital table is good “if a client feels they can’t confide with an interpreter present.”

Independence. Six of eight participants reported that the digital table alone made them feel the most independent. Amber explained, “I don’t have to wait for an interpreter. It saves time.” Cathy, on the other hand, said, “it’s true, I did not need to rely on an interpreter, but sometimes independence isn’t exactly what I want—I value smoother conversation.” There is a tradeoff involved with using the digital table: conversation may be slower, but the patient has autonomy and privacy.

Remembering information. Although we do not have data to judge this, seven participants indicated that using the digital table would be helpful for remembering information. Participants said “it provides a record that I could go back and look at” and “it’s all documented.”

Understanding the doctor. Half of participants stated that including the digital table helped them understand the doctor best. Jesse explained, “the table could be used to clarify words that the interpreter may not understand or comprehend.” Similarly, Mark said, “the table showed the exact words the doctor used.” Alex said her preference “depends whether the interpreter is fluent and sharp. The table is better if the interpreter is bad. In that case I would prefer to type for myself.”

Speed of conversation. In follow-up discussion, several participants said they preferred the interpreter when speed of conversation was critical. Sharon said “the table will work only when both parties are patient.” Amber explained, “in an emergency I won’t have time or energy to type.” There are cost-benefit tradeoffs between communicating through a quicker channel (the interpreter) versus a private and independent channel (the digital table).

DISCUSSION

Based on our experience with SSI, we discuss cultural factors, design lessons regarding dictated speech on tabletop displays, and plans for extending this work.

Cultural Factors

ASL is the native and preferred language for many members of the deaf community. There are cultural implications involved in designing an English-based technology for the deaf. For general conversations, several participants preferred communication in ASL with an interpreter. Alex says that communicating through the interpreter allowed her to express herself best: “My identity is Deaf. I prefer interaction in ASL.” The tradeoff happens when a conversation is extremely personal or when one wants independence from an interpreter. In this case, our deaf participants indicated that communicating through the digital table promised sufficient increased benefit for them to use English instead of ASL.

Conducting research with a deaf population presents specific challenges. For non-signing researchers, an interpreter must be onsite to help facilitate interviews and usability studies. We found that traditional means of recruiting such as online postings and email were inadequate. Involving community members and domain experts early on in our process led to a partnership with our city’s deaf community services center. One of their members created a video blog posting in ASL that advertised our study. This was highly effective and illustrates the importance of reaching a population of interest through their preferred language and communication medium.

Speech Recognition and Digital Tables

Speech recognition is a promising technology to support natural forms of interaction around digital tables. Compared to previous work on spoken commands [25][26], dictated speech presents new interaction challenges for both deaf and hearing populations. We identify design principles that increased interface usability and speech recognition:

(1) *The system should limit the impact of ambient noise and the speaking user should not have to turn the microphone on and off.* An on/off button adds an extra layer of unnecessary complexity and work for the user.

(2) *The speaking user should have control over the speech that is added to the display.* Our system presents the user with three best guesses from the recognizer. This gives control to the speaker, allowing them to select only accurately detected words and phrases. While this design requires an additional step, it greatly improves fidelity.

(3) *The interface should enable conversational repair.* Mistakes in recognition and conversation will happen, so it is important to provide users with a mechanism for repairing their speech. SSI enables this through trashcans, intentionally placed in the corners of the display to increase situation awareness by others.

(4) *The interface should provide an auxiliary way to enter speech.* There will always be new words or phrases that stump the recognizer. In our design, we provided a virtual keyboard for the speaking user and found that this alternative worked well.

(5) *Application designers should take into account the tolerance of their user population and pace of conversation in the domain of interest.* The doctor in our study said that speech input was useful, but that some doctors would not have the time or patience for voice recognition software and might prefer a second physical keyboard.

Extensions of this Work

SSI currently works for only two users interacting with the table. We anticipate that other scenarios would involve additional users and are exploring designs that allow various user configurations and input modalities. Using a more capable speech recognizer is also an obvious next step (e.g., Dragon NaturallySpeaking Medical edition [5]). With improvements to the system, SSI could be useful for a variety of group work tasks of interest to the deaf community. Deaf participants mentioned wanting this application for counseling sessions, financial services, design projects, classrooms, and even in retail stores.

While SSI focuses on one subset of the deaf community, we believe it would also benefit people who are Hard of Hearing or have late onset deafness. Interpreters are not used as widely with these populations, so the assistance of a digital table may be even more desirable. We are also considering ways of integrating ASL video interpretations into the interface so that users who are less proficient in English may watch a video interpretation of the speech.

This work also has implications for hearing populations. Medical conversations are challenging for a number of reasons, including forgetting questions and instructions. The affordances of SSI such as preloading questions and referencing past conversation stand to benefit hearing populations as well.

CONCLUSION

This paper presented the design and analysis of SSI. Our work demonstrates that tabletop displays can reduce communication barriers for deaf users in a way that maintains face-to-face interaction and enables privacy and independence. Research on SSI contributes to the growing interest in multimodal multitouch collaborative systems and complements previous work in the field. Finally, this work promises to extend to cooperative computing scenarios for hearing populations to enhance communication through multimodal input.

ACKNOWLEDGMENTS

Research is supported by a NSF Graduate Research Fellowship, NSF Grant 0729013, and a Chancellor’s Interdisciplinary Grant. We thank our study participants,

faculty and staff from UCSD Medical School, and MERL for donating a DiamondTouch table.

REFERENCES

- [1] Cavender, A., Ladner, R., Riskin, E. MobileASL: intelligibility of sign language video as constrained by mobile phone technology. Proceedings of ASSETS 2006, 71-78.
- [2] Clark, H. (2003). *Pointing and Placing*. In "Pointing: Where Language, Culture, and Cognition Meet." Edited by Sotaro Kita. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 243-68.
- [3] CloudGarden. 2008. <http://www.cloudgarden.com/>
- [4] Dietz, P. and Leigh, D. DiamondTouch: A Multi-User Touch Technology. Proceedings of UIST 2001, 219-226.
- [5] Dragon NaturallySpeaking Medical. 2008. www.nuance.com/dictaphone/emr
- [6] Fischer, G. and Sullivan, J. Human-centered public transportation systems for persons with cognitive disabilities – Challenges and insights for participatory design. Proceedings of the Participatory Design Conference 2002, 194-198.
- [7] iCommunicator. 2008. www.mycommunicator.com
- [8] Java Speech API. 2008. java.sun.com/products/java-media/speech
- [9] Jefferson, G. (2004). *Glossary of Transcript Symbols with an Introduction*. In "Conversation Analysis: Studies from the First Generation." Jesse Benjamins Publishing Company, pp. 13-32.
- [10] Matthews, T., Carter, S., Pai, C., Fong, J., Mankoff, J. Scribe4Me: Evaluating a mobile sound transcription tool for the deaf. Proceedings of UbiComp 2006, 159-176.
- [11] Matthews, M., Fong, J., Mankoff, J. Visualizing non-speech sounds for the deaf. Proceedings of ASSETS 2005, 52-59.
- [12] McNeill, D. (1992). *Hand & Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- [13] Miller, D., Gyllstrom, K., Stotts, D., Culp, J. Semi-transparent video interfaces to assist deaf persons in meetings. ACM Southeast Regional Conference 2007: 501-506.
- [14] Moffatt, K., McGrenere, J., Purves, B., Klawe, M., The participatory design of a sound and image enhanced daily planner for people with aphasia. Proceedings of CHI 2004, 407-414.
- [15] Morris, M.R., Morris, D., and Winograd, T. Individual Audio Channels with Single Display Groupware: Effects on Communication and Task Strategy. Proceedings of CSCW 2004, 242-251.
- [16] Morris, M.R., Ryall, K., Shen, C., Forlines, C., and Vernier, F. Beyond "Social Protocols": Multi-User Coordination Policies for Co-located Groupware. Proceedings of CSCW 2004, 262-265.
- [17] NIDCD. 2008. www.nidcd.nih.gov/health/hearing
- [18] Piper, A.M., O'Brien, E., Morris, M.R., and Winograd, T. SIDES: A Cooperative Tabletop Computer Game for Social Skills Development. Proceedings of CSCW 2006, 1-10.
- [19] Ringel, M., Ryall, K., Shen, C., Forlines, C., and Vernier, F. Release, Relocate, Reorient, Resize: Fluid Techniques for Document Sharing on Multi-User Interactive Tables. Proceedings of CHI 2004, 1441-1444.
- [20] Schuler, D., and Namioka, A. (Eds.) (1993). *Participatory Design: Principles and Practices*. Hillsdale, NJ: Lawrence Erlbaum Assoc.
- [21] Schull, J. An extensible, scalable browser-based architecture for synchronous and asynchronous communication and collaboration systems for deaf and hearing individuals. Proceedings of ASSETS 2006, 285-286.
- [22] Scott, S.D., Carpendale, M.S.T., and Inkpen, K. Territoriality in Collaborative Tabletop Workspaces. Proceedings of CSCW 2004, 294-303.
- [23] Shen, C., Vernier, F., Forlines, C., and Ringel, M. DiamondSpin: An Extensible Toolkit for Around-the-Table Interaction. Proceedings of CHI 2004, 167-174.
- [24] Starner, T., J. Weaver, and A. Pentland, A Wearable Computing Based American Sign Language Recognizer. In *Personal and Ubiquitous Computing*, Volume 1 (4), 1997.
- [25] Tse, E., Greenberg, S., Shen, C. and Forlines, C. Multimodal Multiplayer Tabletop Gaming. Proceedings of PerGames 2006, 139-148
- [26] Tse, E., Shen, C., Greenberg, S. and Forlines, C. Enabling Interaction with Single User Applications through Speech and Gestures on a Multi-User Tabletop. Proceedings of AVI 2006, 336-343.
- [27] Turoff, M. 1975. Computerized Conferencing for the Deaf and Handicapped. ACM SIGCAPH Newsletter, Number 16, 1975, pp. 4-11.