USC-SIPI REPORT #363

Understanding Speech-to-Speech Communication in Cross-Lingual, Face-to-Face, Computer-Mediated Contexts: The Social Technical Design of Transonics Solutions

by

Scott Millward

April 2005

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Viterbi School of Engineering
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Suite 400
Los Angeles, CA 90089-2564 U.S.A.

Understanding Speech-to-Speech Communication in Cross-Lingual, Face-To-Face, Computer-Mediated Contexts: The Social Technical Design of Transonics Solutions.

Scott Millward
University of Southern California

Abstract

To efficiently develop communication augmentation devices for cross-lingual, face to face exchanges design strategies must be guided by both social preferences and technological limits. This report describes the application of an iterative, social-technical design cycle to the development of Transonics Solutions, a speech-to-speech translation device. The inherent gap between user requirements and engineering capability is also discussed and several stages of Transonics' development are analyzed. Specific barriers are identified and recommendations given for overcoming them during the next phase of the project.

Key Words: Transonics, Speech-to-Speech Translation, Cross-Lingual Face-to-Face Mediated Communication, Social Technical Design

Table of Contents

List of Figures

1.0 Introduction: A New Communication Environment (Cross-Lingual, Face-to-Face, Computer-Mediated Communication).

The increasing integration of high technology into people's everyday lives has given rise to a generation of "post-human" individuals (Hayles, 1999) who employ technology to communicate more frequently and in more varied ways than ever before. For the average person the global reach of communication technology, and the increased availability of travel have vastly increased the need for communicating in multi cultural environments. McLuhan's concept of the global village has indeed become a reality (McLuhan, 1964). Such a rapid increase in communication proximity threatens to outpace advances in social science theory. Entirely novel communication contexts are more frequently becoming available, ranging from cell-phone television, to SMS, to Blogs, to text translation devices and more, with each new medium affecting the utility of existing ones. This makes the development of sound theory, which often occurs in stages where observation and exploration precede controlled testing and prediction, much more difficult.

One such novel context is that of cross-lingual, face-to-face, computer-mediated communication (CLF2FCMC). In this context, two people who do not speak the same language are able to communicate face-to-face, using a device that records their speech and translates it, in turn, into the native tongue of each communicator. As the capability of this technology advances applications will include business transactions, security operations, education and entertainment. One of the first genres to benefit from this technology will be healthcare. This report describes the development process for the "Transonics Solutions" system, a speech-to-speech translation device designed for medical examinations in linguistically heterogeneous contexts. As the title suggests the device is a part of a larger communication strategy for overcoming linguistic and cultural barriers.

Designing devices for CLF2FCMC environments is challenged by the absence of adequate social theory explaining human behavior in these contexts. In the past the term "mediated communication" has carried with it the assumption that the participants are physically dispersed (Cathcart & Gumpert, 1983). Even when participants are visible to each other (i.e. video conferencing) this type of communication is categorized as something other than face-to-face. Multiple comparative studies have been conducted illustrating behavioral differences in these two distinct communication settings (Adrianson & Hjelmquist, 1999; Chan & Cheng, 2004; Parks & Roberts, 1998; Williams, 1977). Researchers have examined a variety of factors including type of mediation (Straus, 1994), its social bandwidth (Trevino, Lengel & Daft, 1987), the transparency of the medium (Biocca, Harms & Burgoon, 2003) and more.

The simultaneous presence of face-to-face communication and mediated communication cannot be explained by simply combining theories from each context. The availability of a mediated channel clearly impacts the way in which people choose to employ face-to-face strategies. Thus, a comparison of CLF2FCMC contexts to other mediated contexts, or to face-to-face communication is not appropriate. Instead, new theoretical conclusions must be drawn. Thus, we find ourselves at the beginning stages of theory formation where "tinkering" (Ramachandran & Blakeslee, 1998) is the most prudent path toward discovery. As the growing capabilities of speech-to-speech translation continue to provide deployable products, a more fully adopted mediation venue will allow this theoretical growth to expand. While it may be appealing to simply wait until people are using such devices before making predictions and recommendations about behavior, we must be mindful that technical capability will not be the sole driving force behind this new communication genre. If social scientists wait until such devices are widespread the critical analysis will constitute little more than clever hindsight having missed their opportunity for constructive input into the design process. Understanding the social dynamics in mediated communication conditions at the present stage can help engineer "next generation" translation system capabilities that aim to include user preferences in achieving a total communication exchange, rather than providing mere information transfer.

2.0 Objective: Flexible, Integrated, Communication Augmentation

The goal of this research can be described generically as improved information, and meaning, transfer between people with incongruent languages. In the target application environment this improved communication is sought as a means to combat basic inequities in health outcomes for limited English speaking (LES) patients who visit English speaking doctors. Doctor-patient exchanges during clinical examinations are highly dependent on accurate information transfer as a basis for decision-making involving diagnosis and treatment. In intercultural exchanges with LES patients translation errors are frequent, often times with significant clinical consequences (Flores 2004a, b). The most common type of error is omission, however, other mistakes such as substitution, editorialization, and addition also frequently occur (Flores 2004c). Accurate speech-to-speech translation can effectively eliminate many such errors.

Accurate information transfer is not the only factor at work in improving patient health outcomes. A link is clearly established between *quality* of doctor-patient communication and patient satisfaction (Comstock, Hooper, Goodwin & Goodwin, 1982; Beck, Daughtridge, & Slone, 2002), patient compliance (Ong, DeHaes, Hoos, & Lammes, 1995; Teutsch, 2003) and patient recall of information (Levinson & Chaumeton, 1999). Quality communication requires the application of uncertainty reduction strategies that meet patient expectations by adhering to patient conceptions of authority, conversational style and myriad other cultural factors (Ayonrinde, 2000). Thus, the physician patient clinical exam provides a decision-making venue that requires more than standard speech-to-speech translation or decision support. Transonics seeks to broaden augmented communication research to include the addition of user state and decision modeling as a useful step toward addressing the problem of inequitable health care treatment for LES patients.

This research proceeds under the assumption that a link exists between better communication and better health outcomes. It bears repeating that communication, in this sense, does not just entail conveying information from one person to another, but is better defined as a "meeting of the minds" through facilitated meaning transfer and understanding within a given context. Hence, the more specific goals of this research require a better understanding of how humans will behave in CLF2FCMC environments. This effort focuses on the "communication maximizing" strategies available in face-to-face, and mediated environments, as well as those strategies which may be provided by this unique dual venue. Such novel uses (i.e. passing notes under the table, or out of the camera's view, at a video conference) are key to the ability of CLF2FCMC to offer value added interactions.

The search for such uses is an important part of the effort to inform the development of both the translation device and the appropriate way to use it (training) for medical decision-making situations. It also seems prudent to bring such uses to the attention of users, rather than forcing their discovery. We acknowledge the fact that there is a significant gap between what users would prefer (real-time translation through a transparent medium[i]) and what technology can provide. By implementing a human-centered, iterative design process known as the social-technical approach (O'Day, Bobrow & Shirley, 1996, 1998) we will better understand how these preferences can be met given the constraints of heterogeneous linguistic capabilities and cultural differences.

3.0 Theory

3.1 The Social-Technical Approach: Merging engineering and social science approaches to communication technology design

The process of using human needs as the motivating force behind technological development is often referred to as human-centered design. While this term has been interpreted in a variety of ways (Henderson & Kyng, 1991; Simonsen & Kensing, 1997; Kensing & Blomberg, 1998; Newell & Gregor, 2000) here it is meant to suggest that the capabilities of the device under development are driven by predicted improvements in human communication, and constantly refined with the needs of the specific communicators as the impetus for change. It is this cycle of refinement that maintains the link between the functionality of a product and its usability.

Typically device development has been approached in two distinct ways, the technical approach and the social approach (Grudin, 1989, 1990). The technical approach is guided by what can and cannot be done within the current limits of engineering and is driven by what will, and will not, function. The social approach is driven by human needs and preferences and is guided by what will and will not be used (Ackerman, 2000). The classic example is, of course, the VCR. Many such devices were purchased with the allure of gaining control over one's entertainment choices (e.g. cutting ties to scheduled programming or preserving home movies). The irony is that owners quickly found themselves unable to program this new device leaving them feeling more helpless than ever (Mark, 1991; Niemeyer, Rai & Kean, 1997).

The perfect merger between social requirement and technological capability is difficult to reach in settings that require human interaction. The unpredictable nature of user behavior creates a need for a system that can adapt in flexible ways that are beyond current engineering abilities. Tradeoffs must occur between what people want or need and what technology can provide. Ackerman (Ackerman, 2000) refers to this as the "social-technical gap", citing it as the most important problem in computer supported collaborative work environments.

In the case of Transonics, acknowledgment of this inherent gap allows research to focus on the examination of the communication environment as a hybrid model that is neither technically dictated, nor socially driven, but a little of both. We assert that translation devices, like Transonics, should focus on facilitating a whole human interaction as the primary goal. We claimed earlier that communication exchanges are more than simple information transfer, they are inherently relational, and as such may appear to be primarily a social problem. However, human solutions to this problem have limits. The integration of technological solutions can ease the achievement of social goals if such technology is allowed to integrate into the social process instead of dictate it. Our approach to studying CLF2FCMC follows a social-technical design template, by integrating the human and machine development processes into an ongoing iterative cycle (O'Day, et al., 1996, 1998).

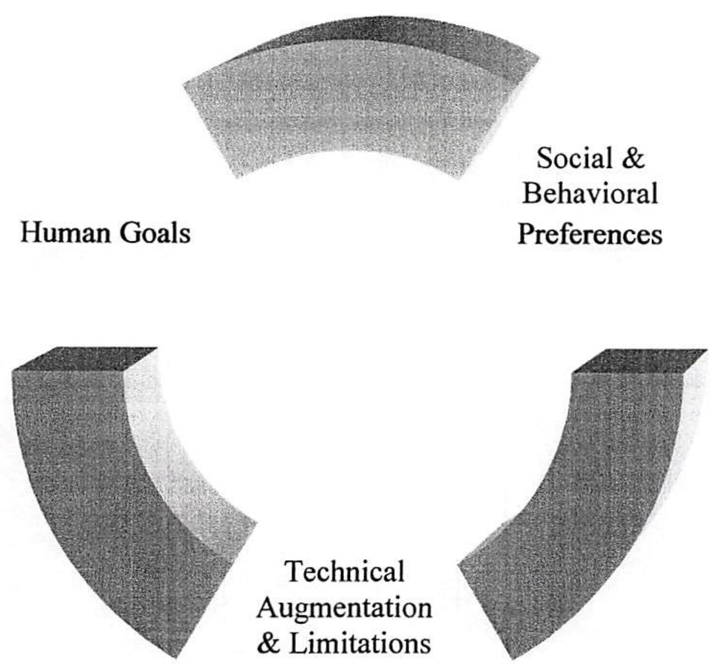

Figure 1. The social-technical design process

This cycle is bidirectional and rapidly repeating. The process helps to highlight the changes in each stage that are revealed by advances in another stage. This interconnectedness between goals, preferences and technology is apparent in many venues[ii].

Understanding social and behavioral preferences has traditionally been the prevue of social scientists while engineers consider the technical limits. Both seem interested in helping humans achieve their goals, but neither approach seems predisposed to routing their efforts through the findings of the other. Social science has provided a neat set of theories describing the way people interact via communication media (Fulk & Collins, 2001). They are:

- *Media capacity theories*: Media have different capacities to carry communicative cues. Simple cues such as words, can be carried almost any medium (e.g. telephone or e-mail), but more complex cues, such as gesture, need higher bandwidths (e.g. video conference).
- *Input Process Output Theories*: These theories analyze process gains and losses due to media limitations. For example, a teleconference may limit the availability of certain non-verbal gesture cues (process loss), but it creates an opportunity for persons on one side of the conference to communicate privately during the conversation by passing notes under the table (process gain). This channel is not typically available in face-to-face meetings.
- *Structuration theories*: These theories describe the usage of media as faithful or ironic. Faithful uses are those intended by the device designers while ironic uses are adapted or user constructed. The development of emoticons (expressive symbols such as :) used during text-based computer mediated communication)

These theories look at communication media in terms of what people can and cannot do with it. The goals are described in terms such as equality of participation, socio-emotional expression, conflict and consensus, efficiency (time to complete task), decision quality, and user satisfaction.

Of these social scientific approaches, media capacity theories may be the closest to a "traditional engineering approach". For example, media richness theory (a media capacity theory) describes how people will select communication media based on the communicative goals they hope to achieve and the manner in which they prefer to communicate (Daft & Lengel, 1984). This selection is often dictated by the limitations imposed by each mode of communication. For example the bad news of firing an employee can be delivered via e-mail to save the manager emotional stress, or delivered in person to provide more comfort the employee. However, bandwidth limitations inherently affect a media's capability to transfer certain communication cues, thereby affecting the communication strategies available to users and the goals that they are able to reasonably achieve[iii]. In other words, it would probably be difficult for the manager to deliver the message via email and also appear comforting.

What if the manager wants to deliver the message in person but simply cannot be there? A sound engineering response would be to create a communication medium that could, as closely as possible, approximated face-to-face communication. Perhaps a virtual reality chamber, or hologram projection, or even a high-end video conference would suffice. So long as all of the cues that are available in face-to-face communication are also available in this new communication medium would the two not be comparable? While no communication medium has been devised that approximates face-to-face communication, theory suggests that this is probably not the case, nor is it necessary. First, one cannot assume the implementation of new technology is capable of simply replacing a human social practice without affecting other aspects of the social-technical cycle (Cook & woods, 1996; Grudin, 1994). Second, the achievement of human communication goals is rarely limited to a specific course of action[iv]. Perhaps the employee would be just as comforted if the message were sent in an electronic card, or via singing telegram etc.

Both social scientists and engineers must agree on the importance of assisting users in achieving their goals. In the case of Transonics, we do not seek to achieve these goals by creating an exact replica of a human interpreter (this is technologically beyond our capabilities). Instead we will attempt to discover alternative means of facilitating interactions between linguistically heterogeneous doctors and patients. A social technical approach will aid in this discovery. A careful analysis of the entire development cycle at regular intervals provides the advantage of maintaining relative stability within the design process. This helps developers account for social factors within the product design, *during the product design*, and limits the need to reverse engineer social uses for a product after the fact.

3.2 Initial Goals

We have attempted to take the approach described above while developing the Transonics speech-to-speech translation device. Our initial assumptions about the goals of the actors were not formulated using literature review or expert interviews. Instead, lab members shared their personal beliefs merely in an effort to devise general starting points which held no special attachment to established sources. These original goals were:

1 - Needs of the patient
- Transfer information to doctor regarding disease state
- Reduce level of uncertainty regarding trust of doctor
- Discover how to fix what is wrong
- Discover how to prevent what is wrong from happening again

2 - Needs of the doctor
- Get information from patient about disease state
- Diagnose what is wrong
- Provide treatment recommendations
- Alleviate patient anxiety regarding trust

3 - Needs of the system
- Get clear, non-colloquial, speech input
- Minimize ambient noise

3.3 The Utility of Transonics in a Medical Context

Making predictions about the utility of a device like Transonics presumes some research evidence that suggests how improved communication (as communication was defined earlier) has affected cognition, collaboration, and performance in other environments. It may seem obvious that sharing information in accurate and meaningful ways is a vital part of the clinical decision-making process. Indeed, a litany of physician-patient communication studies concludes that a need exists to improve the process. It is also clear that lack of good communication can lead to catastrophic consequences (Helmreich, 2000; Meshkati, 1996; Smedley, Stith & Nelson, 2003; Tajima, 2002). This risk is particularly high in cross cultural communication environments (Taveras & Flores, 2004; Flores & Pachter, 2004; Flores, 2004). Although a need is well established, and a wealth of task specific information is readily available to both doctor and patient, even the best critical thinking skills and training may be influenced by seemingly uncontrollable factors ranging from unconscious personal bias (Johnson, Kurtz, Tomlinson & Howe, 1986) to stress (Lazarus & Folkman, 1984) to interpersonal conflict (Thomas, 1992). In short, it is one thing to claim better communication is needed it is another to facilitate it. Studies suggest that making the "right" clinical decision for a patient depends heavily on patient characteristics such as cultural conceptions regarding healthcare and treatment (Fitzgerald, Mullavey-O'Byrne, Twible & Kinebanian, 1995; Kohlase, 1995; Badia, Roset, Herdman, & Kind, 2001), attitude, (Rosen, Tsai & Downs, 2003) and patient emotional state (Ubel, 2005). By understanding that the communication process is an entire experience systems can be designed to maximize the aspects of this experience that are essential for the completion of a given task. Determining what these essential aspects are, requires the implementation of the social technical iterative process through user studies.

4.0 Methodology: Applying the Social-Technical Method

The design cycle of Transonics was initiated by the goals in section 3.2 and the technical requirements of the original research request. These requirements were general and included such design criteria as the capacity for a speech-in / speech-out interface, low word-error rate (high degree translation accuracy), and a rugged, portable housing. The initial prototype met each obligation to a degree, providing a starting point for the integration of social requirements, and the resulting technological modifications. The application of the iterative design process is described here in several steps, culminating in the first formal user experiment. Each step is briefly explained along with social lessons learned how these informed the technical development of Transonics

4.1 Prototype Development

Transonics was conceived of in 2002 as a response to a DARPA research solicitation (DARPA Contract No. N66001-02-C-6023). The original mandate was to create a device with speech-to-speech translation capabilities for immediate use by military troops conducting medical, safety and administrative missions abroad with non-English speaking populations. The original concept was portrayed as a conduit model of communication where information transfer through a mediated verbal channel was the only requirement. An early prototype of Transonics was designed specifically for medical interaction between English speaking doctors and Persian speaking patients. The system operated using a Java-based graphical user interface (GUI) which enabled the doctor to act as the primary user (figure 2).

To begin using the system, both the doctor and patient would put on headset microphones. Next, the mouse was used to maneuver the cursor over the box labeled English/Click2Talk. Clicking this box once began the system recording feature and the user could then speak a phrase or question into the microphone. Clicking the button again caused the device to stop recording and begin translating what the automatic speech recognition software was able to record. A five second limit was built into the system to avoid excessive recordings and to stop recording if the user forgot to press the button twice.

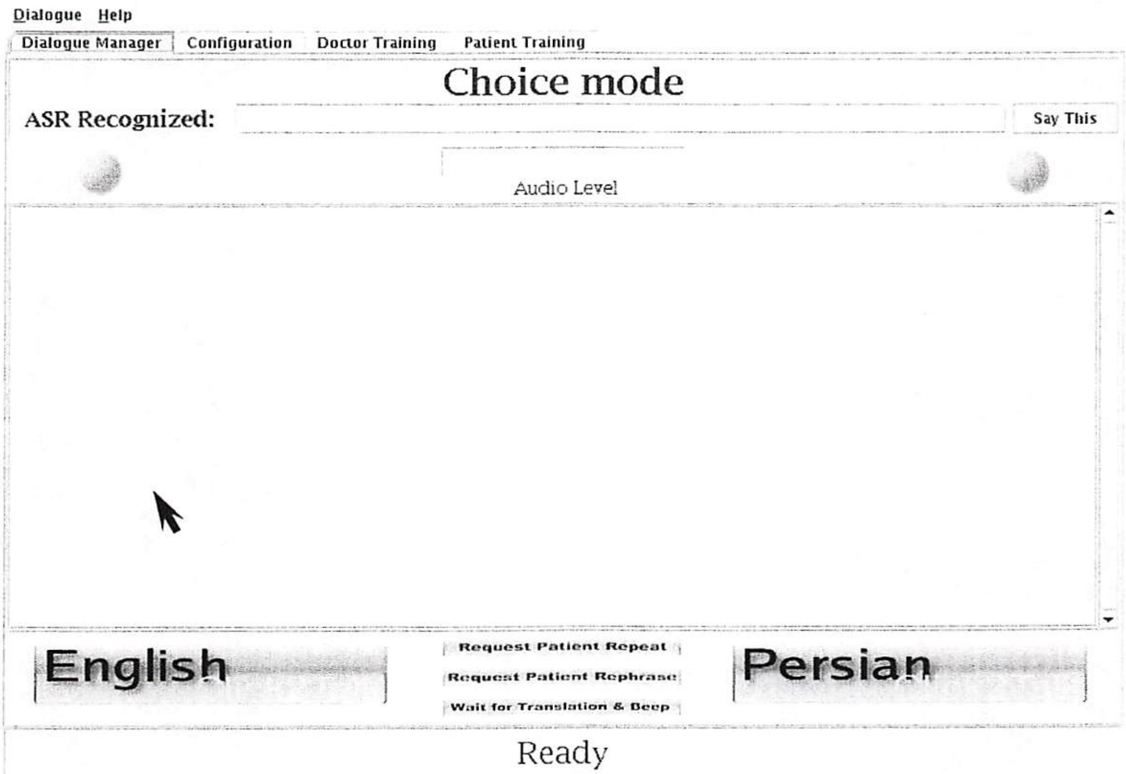

Figure 2. An image of the original interface design for Transonics.

Once the spoken phrase was processed, Transonics would display a list of translations which are then displayed for the user. The "I will try to translate these" pane shows translations that will be spoken word for word. This is more flexible but less accurate. The "I can definitely translate these" pane shows translations that will use similar phrases in Persian. This is less flexible but more accurate. If none of the choices match what the user wanted to say selecting "None of the above" allows the user to re-record his or her voice.

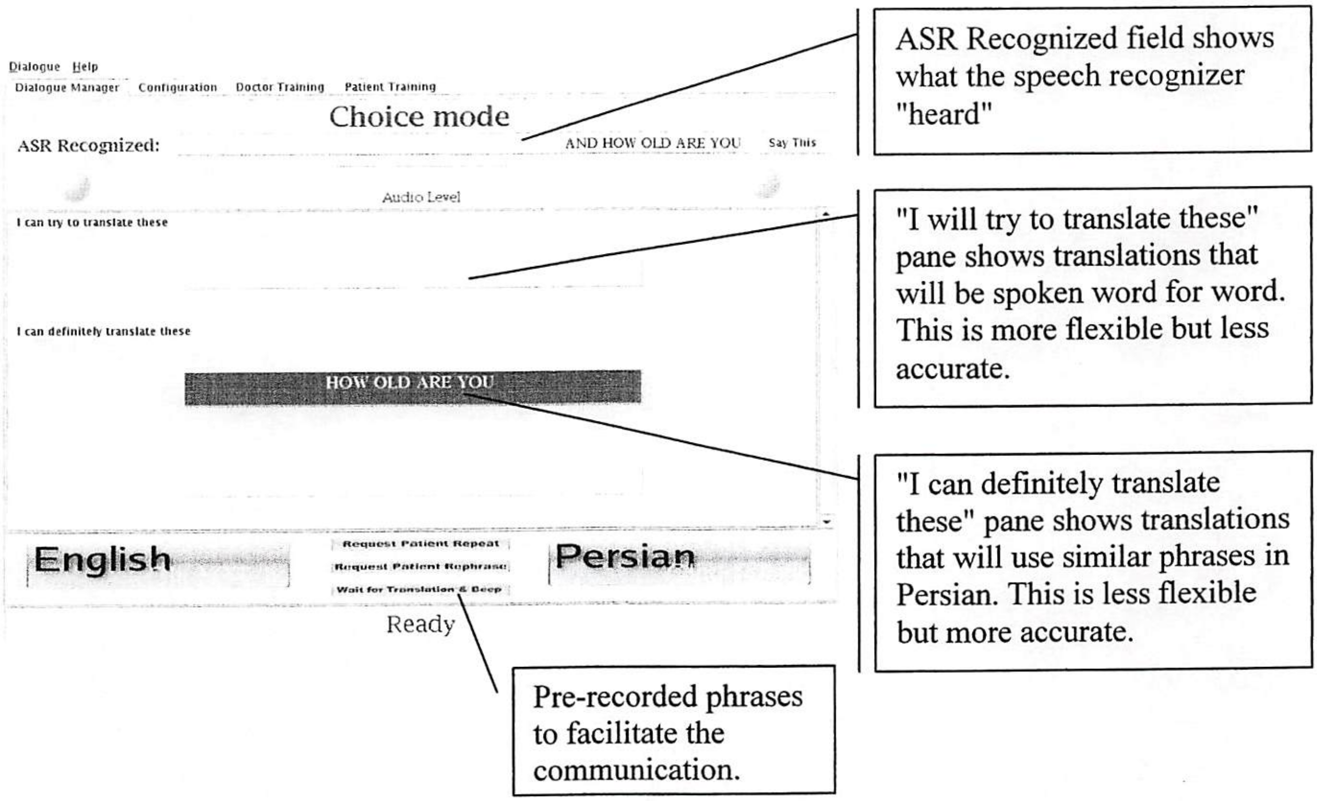

Figure 3. An explanation of some features found in the original Transonics interface

To allow the patient to utilize the Transonics System the doctor must use the mouse to click on the "Persian/Click2Talk" box at the bottom right of the screen. The patient is prompted to speak by a mechanical beep and is given five seconds to record a response. This time the phrase is automatically translated into English and several possible translations are provided for the doctor. The interface also provides buttons that read "Request patient rephrase", "Request patient repeat" and "Wait for translation beep". These buttons trigger pre-recorded instructions in Persian that assist the doctor in facilitating the communication with the patient.

<u>Social Lessons Learned</u>

The impact on user communication preferences and abilities was not considered at this point beyond providing the capability to translate basic spoken phrases between languages. The overarching assumption was that that literal translation would improve information and meaning transfer. It was not know how users would react to communication with such a device, or how the device would otherwise operate under real-world circumstances. Several months were invested in the development of the basic technological requirements of the research grant and the first opportunity to test a functional version of the device did not arise until much later.

4.2 MITRE Evaluation

In August 2004 several development teams participated in an evaluation for DARPA's CAST speech translation program at the MITRE Corporation in McLean, VA. The evaluation was not comparative, but rather the intent was to provide a structured analysis of system performance while mediating between two speakers discussing a medical condition and possible diagnosis. Two monolingual, English-speaking, military medical Subject Matter Experts (SME) and two monolingual, Persian speaking, Foreign Language Expert (FLE) used the Transonics device. The evaluation was compiled in a final report and distributed to each development team in January of 2005, however, several areas for improvement were immediately apparent based on the experience.

Social Lessons Learned:
All SMEs and FLEs received brief training on how to operate the Transonics device, however, several usability barriers clearly remained. Despite instructions covering the required turn-taking behavior FLEs consistently spoke before being prompted by the system, resulting in partial, or no, translation. Unfamiliarity with communication in a mediated face-to-face environment proved to be stressful for both SMEs and FLEs, with each showing visible signs of frustration. This was exacerbated by the system's difficulty recognizing the regional accent of one FLE. When speech was recognized accurately often times the SME would cancel the interaction because of a lack of satisfaction with the translation. Other errors included the presence of "double meaning" words such as the Persian word "***" which, depending on context, can be translated as 'yes' or 'why' in English. This was the cause of consternation on the part of one FLE when he was told that a medical procedure would be performed on him, and his repeated requests as to "why" were translated as repeated affirmations of the procedure. The following social lessons were taken from the evaluation.

- Turn-taking norms were not inherently clear in the system operation
- Interactions through the device caused social discomfort because they are too slow.
- The device must learn to recognize varied accents and differing speech patterns.
- Users must be instructed to accept translations with similar but not exact meanings to expedite conversation.
- The system must be made aware of "double meanings".

Technical Implementation :
Technological limitations dictate the necessity for the device to have a clearly delineated segment of speech to analyze, hence the need for well established turn-taking norms during use. The resulting communication pattern seemed restricted and unnatural compared to free-flowing face-to-face exchanges in which interruptions and talking in tandem are commonplace. To accommodate the social preference for a more natural exchange the indicators for patient speech were changed from impersonal "beeps" to a human voice saying "start speaking" and "stop" (in Persian).

Long awkward pauses often led to users believing that the machine was somehow "broken". To some of this uncertainty a more clearly delineated flashing light was added to indicate when the system was processing information, recording it, or standing by.

4.3 Subject Matter Expert Interviews

A key element in medical device design is the integration of the device into the existing routines of the physician. To facilitate this, input should be gathered from the eventual end user at the earliest stages of the iterative cycle (Rector, Horan, Fitter, Kay, Newton, Nowlan, Robinson, & Wilson 1992). For Transonics, interviews were conducted to determine the necessary capabilities of a speech-to-speech translation device, for both office settings and trauma centers. When possible, physicians were interviewed along with an office staff member who acted as an in-house translator for LES patients within his or her language expertise. Six interviews were conducted, four with doctors who practiced in office settings and two with emergency room doctors. Three doctors were interviewed with bilingual staff, two of which were Spanish speakers and one who spoke Farsi. The doctors specialized as follows: two dermatologists, two general practitioners, two emergency room surgeons. Interviews were approximately fifteen minutes, unrecorded and followed a set list of questions in each case.

Social Lessons Learned:
There was a sharp contrast between the perceived need for translation services of any kind between emergency room surgeons and office based physicians. ER surgeons rarely spoke with their patients, instead relying on visual inspections and test results when making diagnoses and rendering treatment. The need for translation did often arise when discussion with family members was required, however, ER doctors claimed that in most cases the human translation services that were provided to them were sufficient. This was not the case with office based doctors who saw a more pressing need for translation during examinations. These physicians were often required to make diagnoses based on verbal input from the patient and regularly had difficulty conveying treatment options and requirements to their patients. The presence of bilingual staff or family members made this communication easier, however, all four physicians associated this with other problems. Common complaints were crowded waiting rooms, usurping staff resources and poor translation quality. Lack of cultural understanding was also an issue. For example some patients attempt to negotiate the bill and feel slighted if the doctor's staff does not comply, or, certain cultures associate a specific minor disease state with very bad luck and have a tendency to express a greater emotional need for immediate treatment.

All four physicians agreed that overcoming the linguistic barriers to information transfer was paramount, and that cultural sensitivity is important, but secondary to making the right diagnosis and treatment. Doctors seemed to feel that the "right" diagnosis and treatment meant that which was most likely to lead to elimination, or control, of the disease state.

When pressed for specific capabilities that they would like a translation device to have several different answers were provided. One doctor wanted to have visual cues for patients to follow, such as a chart of the human body that could act as a common reference tool between languages. Other doctors wanted to be able to print out treatment recommendations in the target language, or use the device to print out medical history forms in various languages. All physicians expressed concern over the idea that using such a device would increase the time required for an examination, none of the physicians expressed any concern about a lack of trust or rapport building with patients when using a translation device.

- A need for translation devices exists in the daily routine of office based doctors and speech-to-speech translation is more useful in these settings than in emergency settings.
- Basic information transfer is the primary goal of the device, however, cultural sensitivity is also important.
- Visual cues may improve information and meaning transfer
- Doctors are concerned with giving the patient a "takeaway" message that acts as a tangible reminder to follow the treatment plan. Printing out instructions in the target language would fill this role.
- Rapport and trust are perceived by doctors as being unaffected by a mediating device or person.

Technical Implementation:
The physician interviews helped us begin to construct a coherent strategy for implementing the Transonics device into the routine of the doctor in a non-invasive manner. The technological requirement of recording a phrase in its entirety before attempting translation would clearly become a limiting factor. The inability to translate in real time would compromise the utility of the device in time critical domains. Thus emergency settings were ruled out as target environment. Within the office setting doctors still wanted to be able to achieve the primary goals of information exchange within a reasonable time and they wanted to do so in ways that did not require the an increased strain on staff resources. This meant designing a device that would be simple to use while still accounting for the cultural sensitivity needs of the patient.

Two technical solutions were proposed to facilitate these user preferences. First, the added emphasis on pre-formed questions was discussed. The device interface already had the capability of recalling a history of past questions, however, this was only accessible through a separate page that the doctor had to find by clicking a discrete tab under one of the menu options. Bringing this history to the foreground would allow doctors to avoid the two-step record/translate process each time they wanted to ask a generic question. In addition, it was concluded that the doctor could eliminate more interaction errors by expanding the menu of pre-recorded phrases intended to help regulate patient responses

A second solution involved the implementation of a cultural tutor to help the doctor interact with the patient. The tutor would be programmed to suggest methods of communicating with the patient based on pre-programmed data about cultural norms and annotations from previous conversations. At a later stage this concept grew to also include domain knowledge decision-support for doctors and conversational modeling by the device.

4.4 Exploratory Pilot Study

After the prototype evaluation of Transonics it was clear that the functionality of the device was in large part dictated by the manner in which the users interacted with it. The doctor interviews revealed certain key areas for which a device such as Transonics would likely be useful, however, the social requirements of the users, beyond specific medical contexts, were not obvious. To assist in identifying usability barriers for Transonics a pilot study was conducted using Ph.D. students from the Annenberg School for Communication acting as doctors, and a bilingual project member acting as the patient. Users were administered a brief survey before interacting with the device, given short oral instructions on its operation, and then asked to use it for ten to fifteen minutes to interact with the patient. After their interaction an informal focus group discussion took place to try and gather more feedback about the experiences of each user. Subjects were all about the same in terms of self perceived computer skill (average) but varied slightly concerning prior conceptions and use of speech interface devices.

<u>Social Lessons Learned</u>:
The device received low scores on whether or not the user could focus on both the patient and the device. Many comments suggested that the "doctor" did not feel as though he/she was speaking to the patient, but was instead speaking to the machine and the machine was responding. Despite this social discomfort, most people found the device easy to operate in general, however, several comments pointed to specific things that could have been improved. For example, multiple translations of Persian speech were provided for the doctor, however, these were deemed unnecessary because the doctor (who did not speak Persian) had no way of knowing which one was right. It was suggested that the patient be able to see the Persian speech recognition before translation by the machine. It was thought that this might help the interaction by allowing the patient to monitor the accuracy of his or her own translation. The delay between user speech, machine translation and output caused an uncomfortable silence for some users, and poor recognition seemed to exacerbate this feeling. Generally, users who were observed during multiple failed voice recordings reacted with nervous laughter and signs of frustration, such as turning their palms upward and looking around the room. At least one user felt the delay caused by poor speech recognition could be eliminated by using a typing interface, however, the task did not require the doctor to actually physically examine the patient, so perhaps these sentiments would be different if the users hands were required to do some other activity. There were also several problems with the device interface including the small size of buttons and the failure to click activation buttons at appropriate times. Some of the same issues revealed in the MITRE evaluation remained, however, the appropriate solutions had become clearer.

- The ability to focus on the patient and the device was low.

- The interface design was not intuitive, and as a result users failed to take advantage of several key device capabilities (such as the prerecorded phrases).
- The process of recording a phrase or question for translation is slow, often due to user failure to stop the recording manually when he/she is finished talking.
- There is no means of monitoring the accuracy of the Persian-English translation.
- Interactions through the device cause social discomfort because they are too slow.
- Users could be instructed to accept translations with similar but not exact meanings to expedite conversation.

Technical Implementation:
To address the interface design issues, a new interface was designed which primarily relied on an external keypad for controlling the device. The keypad was designed using a USB number pad with a mixture of medical symbols and device specific instructions (figure 4.). It was hoped that by replacing the two-step process of maneuvering the mouse to an area and then clicking with a one step button push, the user would free up cognitive resources for the task at hand. It was also hoped that the more tactile button push would reduce the number of errors caused by failure to stop the recording process. The keypad had the added benefit of providing the patient with a visual cue that would assist in turn taking. Finally, the keypad offered easier access to prerecorded phrases and effectively brought all of the control to the fingertips of the user, instead of spreading them throughout the screen. The mouse remained active, however, the user was now able to navigate through the screen and activate items using arrows on the keypad. This added a layer of redundancy to the interface and was an effort to allow users to figure out the way that worked best for them.

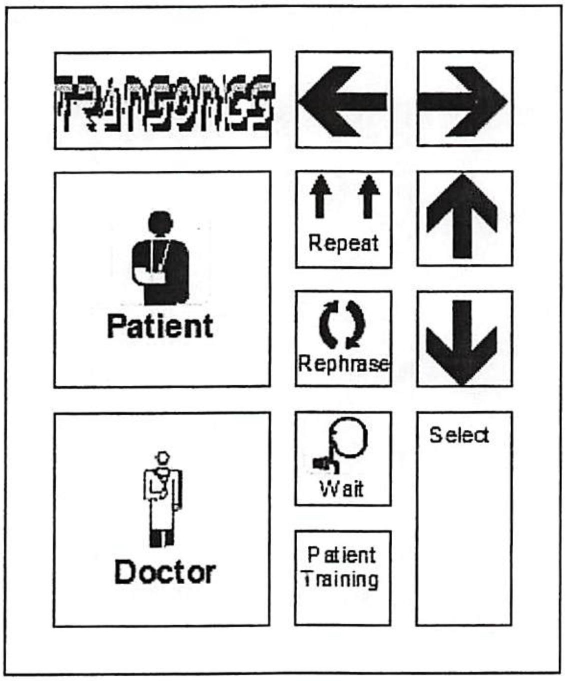

Figure 4. The second generation Transonics keypad interface

The screen itself was also redesigned to help address some of the social lessons learned. It was determined that some of the discomfort among users was caused by "dead time" during which nothing was being recorded, but the device was not indicating that it was processing information. With no update on the state of the device users sometimes felt as though they had done something wrong or still needed to perform an additional action in order to make the device function. To inform the user of the device state an icon window, placed at the top of the screen, changed as the system state changed. For example, a solid microphone (the same color as the background) would change to a red microphone with the words "recording" to indicate that Transonics was recording the user's voice. Other changes included the relocation of the English and Persian keys to the top of the screen and the addition of "record" in lieu of "Click2Talk". The color scheme was also redesigned to reflect a more polished appearance.

Figure 5. The second generation Transonics screen.

A further measure was taken to address the social behavioral concerns highlighted by the interviews and user studies. A six-minute training presentation was developed to highlight the recurring difficulties with the system and provide suggestions for overcoming them. The presentation covered the basic operation of the device and suggested some strategies for streamlining the interaction. This was meant to be used in conjunction with practice examinations using the device prior to any deployment.

4.5 Developing and Testing Device Training

At this point the social-technical cycle had clearly begun to pay dividends by identifying specific gaps between user preferences and technical capability. However, the changes had been primarily on the technical side with less thought regarding the necessity for users to try and adapt to the needs of the device. Nonetheless, the goal remained for users to discover how to interact with Transonics in ways that amplified their communicative ability. It was also increasingly clear that how this should be accomplished was not inherently obvious to users. It was determined previously that the user interface was not intuitive in this regard and that the current level of training received by users was lacking. Thus, the next round of experiments was designed with the objective of determining how to better train users to use the device in more communication enhancing ways. Specifically we sought to test our recent interface and training developments described above. Measurement metrics were aimed at the degree of accomplishment for three goals: information transfer, interpersonal relationship building, and collaborative decision-making regarding treatment.

Experimental design:

An exploratory study was designed to determine the usability of the interface, and the training needs of participants using Transonics in a clinical examination environment. English speaking doctors and Persian speaking patients participated in a clinical examination under two separate, comparative conditions (1) translation device (2) human translator. Subjects acting as doctors were recruited from the student population at the Keck School of Medicine. Subjects acting as patents were recruited from the Iranian Graduate Student Association at USC[v].

Doctors were allowed 15-minutes to collect a relevant health history from the patient, and based on this information provide a diagnosis and recommend a treatment plan. In addition, half of the participating doctors received training on the translation device in the form of a 6-minute instructional Power Point presentation and 10-minutes of practice with a Persian-speaking member of the experiment staff. The other half was only given an instruction sheet and a brief oral description of the device's basic operational functions. All patients were trained on the use of the device using pre-recorded instructions, which they listened to upon entering the exam room. Patients were also standardized by training them to exhibit symptoms of the same disease state.

In the translation device condition, doctors entered the exam room to find the patient waiting and Transonics running on a laptop, ready to use. In the second condition a Persian speaking interpreter was waiting with the patient. In both conditions doctors were able to read the patient's chart before entering. In the Transonics condition an instruction card was left next to the device which reiterated how to operate it. All sessions were video and audio taped, however, no experiment staff were in the exam room with the doctor and patient (and interpreter). At ten minutes a knock was mad on the door to indicate five minutes were remaining, and at fifteen minutes the exercise was stopped by experiment staff.

Social Lessons Learned:

The findings revealed that when difficulty using the device interface occurred it correlated with overall poor ratings of device satisfaction and less confidence by doctors that that their patients trusted and/or understood the diagnosis. Doctors who were selfrated as average or poor at using new technologies reported significantly more difficulty forming interpersonal relationships, gathering information and making collaborative decisions with their patients. Most of these struggles stemmed from failures in speech recognition due to irregular speech patterns, excessive use of space fillers or inflexible vocabulary usage.

Not surprisingly, users with previous experience using voice interfaces found Transonics easier to use and more satisfying. These experienced users also found it significantly easier to focus on the other participant while using the device, and doctors had more confidence in their patient's level of understanding, and trust, regarding their diagnoses.

The amount, and quality, of doctor training also had an impact on user satisfaction for both doctor and patient. Doctors who received more training prior to their examinations of patients increased their number of successful interactions (defined as asking a question-and getting a comprehensible answer) by nearly double from an average of 8.4 per 15 minutes to 15.4. This coincided with a rise in mean doctor satisfaction (untrained μ=2.86, trained μ=4.857,) however, it also correlated with a marked decrease in patient satisfaction. One explanation for this may be that the increases number of interactions was due to increased involvement by the doctor in the operation of the system and less acknowledgment of the patient on a social level. This finding highlights the need for "transparency" in the system.

A better prediction of patient satisfaction was positive immediacy. Immediacy is defined as "the expression of affection, inclusion, and intensity of interaction between two people"(Buller & Street, 1992, p. 123). Video tapes of the 15 mediated interactions were coded for some positive and negative immediacy behaviors. These behaviors included negative aspects such as: body orientation away from other, crossed arms, self-touching as well as positive aspects: mutual gaze, head nodding responses, complementary hand gestures. The analysis revealed that as doctor's positive immediacy behaviors increased so too did patient satisfaction, $r(15) = .544$, $p < .05$ ($p = .036$). This is consistent with other findings in face-to-face situations (for a review see Burgoon, Buller & Woodall, 1989).

A tendency toward positive nonverbal immediacy is likely an individual characteristic of communication style, however, these results argue for an effort to build a mediation device that does not restrict such tendencies, and if possible, supports them. The video analysis also suggested that those examinations which took on characteristics of two "human-machine" interactions were less satisfying than those that more resembled a single, mediated interaction. While better training may solve a portion of this problem by allowing users to focus less mental effort on the device, the fact that the interaction is neither "hands-free" nor "eyes-free" presents an obstacle. Limits on user attention resources are another argument supporting the need for a more transparent system.

These findings reinforce the interconnectedness between social and technical aspects of device development At the earliest stages user training will play a significant role in the implementation of the Transonics device. Most users are likely to be inexperienced with completing complex tasks in linguistically heterogeneous communication environments. Doing so with a translation device is even less likely to be familiar. Developing a device that allows users to change the interaction parameters as their capabilities grow will be imperative.

- A more transparent system design is needed.
- Better training is needed to teach users how to employ Transonics as a communication tool within a larger strategy.
- The system needs to be adjustable for varied user skill levels.
- The system must support relational communication aspects as well as information transfer.

Technical Implementation:
Lessons learned from these experiments are still being implemented. Improved training was developed to facilitate doctors' conception of the device as a single communication tool in a repertoire, rather than a general-purpose conduit through which all information must be forced. This training included an overview of how the system functions, how to interact with it and video examples of what to do, and not to do, to get the most out of it. A practice module is under development to allow doctors to interact with a Persian computer agent who will assist the doctor in learning useful communication strategies.

A second agent is under development to act as a conversational guide to the users. Chai, Pan, Zou & Houck (2002) have argued that most current research programs on multimodal interfaces have suffered from a neglect of the context in which user inputs are embedded, focusing rather narrowly on algorithms for multimodal fusion. Our recent research in speech to speech translation has resulted in a similar conclusion regarding the importance of discourse and pragmatic context. Seemingly perfect translations (when looked at out of context) have resulted in large cognitive disparities among participants when examined within context. For example, consider a word such as "bar", in English, which can refer to an establishment (pub), a counter, a law association, a long material (such as a bar of gold), a barrier, a musical notation etc. Asher and Lascarides (1995) present a linguistic analysis on how discourse affects the meanings of words, providing further evidence that the information contained in discourse could aid in the correct transfer of concept during the translation process.

Other areas of potential technical growth include the development of cultural agents for advising the doctor *how* to communicate with a certain patient, and a visual tracking/user state modeling component that will assist the doctor in monitoring the engagement and comprehension level of the patient.

5.0 Discussion

The implementation of the iterative social-technical design process into the Transonics development cycle revealed the need for constant, thoughtful change. The initial design stemmed from a conduit model of information transfer and was developed under the assumption that machine word-translation would fill the linguistic gap and be enough for communicative success. As the device was refined, with the entire communication process in mind, the model was changed so that Transonics more resembled a mediation tool for a single communicative channel (verbal) embedded in a larger meaning sharing system.

The most valuable outcome of our observations in these various contexts has been a greater understanding of user needs and preferences. For example, we have discovered that when lacking experience with the device users tend to try and force all communication through it rather than using other open channels. We have also noted that users who do use Transonics in conjunction with nonverbal gestures, and eye contact, tend to have more satisfying, and useful interactions. To this point the focus has been mostly on the needs of the primary user (the doctor) however, we have also gained valuable insight regarding how to make the entire interaction a success. Yet, as the needs of both users become clearer so too are the limits of the Transonics platform, and its software.

Even embedded within a larger system it has become increasingly apparent that the technological requirements of the Transonics device require communicators to make certain behavioral changes. Our observations of communication through the Transonics device revealed these social and behavioral effects that were hidden during the prototype construction stage. These included user's anxiety about their ability to work with the technology, a reduction in interpersonal rapport when the device was overused, and a variety of other important "lessons learned" that continue to guide the manner in which this device will be deployed. These lessons have made easier the task of identifying the specific nature of the social-technical gaps for this particular device.

For example, face-to-face environments produce an expectation of immediate message production and feedback. In the case of Transonics this is not an option, producing a climate of uncertainty among the users. Those with limited experience in cross-lingual communication environments may feel this to an even greater degree. In addition to precluding real-time communication, the record-translate-playback functionality of Transonics increases the length and number of "awkward silence" moments. Social discomfort is associated with such moments during verbal exchanges as well (Jaworski, 1993; Johannesen, 1974). This inability to produce real-time translation, and eliminate pauses, is less a function of poor technological capability and more function of how language is structured. Because phrases often derive their meaning from the words at the end of sentences, any translation device must consider the entire phrase as the unit of analysis instead of trying to translate at the word level. Thus, there are inherent limits to how quickly such devices can be made to operate.

Adding to the gap between social preference and technological capability is the fact that translation devices are attempting to replicate a process which has innate flaws. The transfer of meaning across cultural boundaries occurs even when participants speak the same language. Indeed, what constitutes meaning has been the subject of philosophical debate for thousands of years. In the clinical exam setting alone communication struggles to create shared conceptions of humor, time sensitivity, emotional reaction and a host of others. Can technology really make this process better?

Eventually this gap between what users would like, and what modern engineering can provide, may be closed with the advent of some unforeseen technology. However, what can be done in the mean time to get the most out of the capabilities we do have? Is it the abject lesson of the social-technical design process to merely reveal what is wrong without providing a means of fixing it?

6.0 Recommendations

The development cycle for Transonics is far from complete. To date the process has revealed many areas which require improvement if the device is to provide a true, added-value, to real world scenarios. Continuing the iterative cycle of social discovery and technical refinement is still a promising path. Yet exploratory study, as valuable as it is, resembles a rudderless ship, bobbing in the ocean. The following recommendations are meant to serve as guiding principles to expedite the journey.

Ackerman (2000) argues for the focus on computer supported cooperative work (CSCW) to shift toward a "science of the artificial". This position, that the social-technical gap is an inherent aspect of CSCW, has been supported during the Transonics development process. This is especially apparent in the face-to-face setting of Transonics where communication through a computer lends itself to talking to another person without needing to physically acknowledge their presence. If progress is to be made in the utility of CLF2FCMC devices the context itself must be recognized as artificial, and new paradigms developed. This means that a device like Transonics should not necessarily endeavor to replicate face-to-face human interaction, but rather should seek to accomplish the same, or similar, goals in a different way.

To illustrate, consider the previously discussed idea of real-time communication. In the current design, during moments when the machine is processing input an awkward silence occurs. When this happens during face-to-face communication the effect is often to produce feelings of discomfort among the communicators. If we were to approach this problem by comparing CLF2FCMC and face-to-face communication we would see that the one thing differing between the two contexts is the availability, or lack of, real-time communication. This comparison does not treat CLF2FCMC as its own, separate, context and ignores the fact that it is not simply a lack of real-time communication that creates the problem, but rather the presence of silence.

A new communication context requires a new approach and the restrictions of existing communication paradigms simply will not allow for progress in this case. Instead, a new model must strike a balance between user goals and technological capabilities while trying to account for user preferences as much as possible. The extent to which these preferences are met may indeed dictate the final utility of the device, however, the added ease of attaining a communication goal is strong motivation to slightly alter these preferences.

The question remains, how do we develop an approach to novel communication contexts? Arias et al. (Arias, Eden, Fischer, Gorman, & Scharff, 2000) suggest that a key guideline for collaborative design in human-computer interaction (HCI) contexts is the necessity for systems to evolve rather than be constructed. There is a long history of communication devices that were built for specific purposes and developed through use into tools for achieving unrelated goals[vi]. Indeed, users' need to discover utility within a device often fuels the discovery of unintended applications, and hence designs innovation. In the case of Transonics we have already seen such innovation to a degree. For example, in one practice session the Persian speaking member of the staff was delayed while working with another trainee. While waiting, two of the trainees began to have a conversation with each other in English using the Transonics device. This allowed them to more fully explore the device's limits and familiarize themselves with the speech interface.

The more users are exposed to Transonics in realistic settings the more we will learn about how it can be used to achieve communication goals. To facilitate this process the device itself must be designed as an open-use system. Attempting to account for every possible communication scenario is likely to restrict the development of unforeseen uses for the system. Instead, an architecture which allows for users to ad utility to the device as it is constructed is more likely to benefit the actual problem solvers (the users) in the long run. In the case of Transonics several opportunities exist for open design such as the ability to customize hotkey phrases or to allow greater patient involvement.

Approaching CLF2FCMC as a completely new communication context, employing open design and encouraging user innovation are key recommendations for the development of Transonics. These actions must still take place within an iterative cycle which forces the development team to continually to retrace our steps, making sure that changes are made for the better. In this case "for the better" means that the goals of users are facilitated, and their preferences maintained to the highest degree allowable by technology.

Several issues remain to be resolved including the difficulty of meaning transfer, the limitations of domain knowledge, the incorporation of cultural and emotional content, and the impact of user technical skill. As each component is addressed we must remember to consider the impact on the entire communication experience. It is only by taking this holistic approach that we can develop a device for overall improved achievement of communication goals, and the resulting positive impacts on user quality of life.

References

Ackerman, M. (2000). The intellectual challenge of CSCW: The gap between social requirements and technical feasibility. *Human-Computer Interaction, 15*(2), 179–203.

Adrianson, L., & Hjelmquist, E. (1999). Group processes in solving two problems: Face-to-face and computer-mediated communication. *Behaviour & Information Technology, 18,* 179-198.

Arias, E., Eden, H., Fischer, G., Gorman, A. & Scharff, E. (2000). Transcending the individual human mind: Creating shared understanding through collaborative design. *ACM Transactions on Computer-Human Interaction,* 7(1), 84–113.

Asher, N. & Lascarides, A. (1995) Lexical disambiguation in a discourse context. *Journal of Semantics, 12,* (69-108).

Ayonrinde, O. (2003). Importance of cultural sensitivity in therapeutic transactions: Considerations for healthcare providers. *Disease Management & Health Outcomes, 11*(4), 233-248.

Badia, X., Roset, M., Herdman, M. & Kind, P. (2001). A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. *Medical Decision Making, 21,* 7-16.

Beck, R., Daughtridge, R., & Slone, P. (2002). Physician-patient communication in the primary care office visit: A systematic review. *The Journal of the American Board of Family Practice, 15*(1), 25-38.

Biocca, F., Harms, C., & Burgoon, J. (2003). Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence-Teleoperators and Virtual Environments, 12*(5), 456-480.

Buller, D. & Street, R. (1992). Physician-patient relationships. In Robert Feldman (Ed.), *Applications of Nonverbal Behavioral Theories and Research,* (pp 119-141). Hillsdale, N.J: Lawrence Erlbaum.

Burgoon, J. K., Buller, D. B., & Woodall, W. G. (1989). *Nonverbal communication : the unspoken dialogue*. New York : Harper & Row.

Cathcart, R., & Gumpert, G. (1983). Mediated interpersonal communication: Toward a new typology. *Quarterly Journal of Speech, 69*, 267-277.

Chai, S. Pan, M., Zou, X. & Houck, K. (2002) Context-based multimodal input understanding in conversational systems. In *Proceedings of the Fourth IEEE International Conference on Multimodal Interfaces.*

Chan, D. K. S., & Cheng, G. H. L (2004). A comparison of offline and online friendship qualities at different stages of relationship development. *Journal of Social and Personal Relationships, 21,* 305-320.

Comstock, L., Hooper, E., Goodwin, J.M., & Goodwin J.S., (1982). Physician behaviors that correlate with patient satisfaction. *Journal of Medical Education, 57*(2),105-112.

Cook, R. & Woods, D. (1996). Adapting to new technology in the operating room. *Human Factors, 38*(4), 593-613.

Daft, R. & Lengel, R. (1984). Information richness: a new approach to managerial behavior and organizational design. In Cummings, L. & Staw, B. (Eds.), *Research in organizational behavior 6,* (pp. 191-233). Homewood, IL: JAI Press.

Fitzgerald, M., Mullavey-O'Byrne, C., Twible, R. & Kinebanian, A. (1995). Exploring cultural diversity: A workshop manual for occupational therapists. *School of Occupational Therapy*, The University of Sydney.

Flores, G. (2004). The impact of medical interpreter services on the quality of health care: A systematic review. *Medical Care Research and Review*; In Press

Flores, G. & Pachter, L. (2004). Culture, ethnicity and linguistic issues in pediatric care: Urgent priorities, unanswered questions, and a research agenda. *Ambulatory Pediatrics, 4*(4), 276-282.

Fulk, J. & Collins-Jarvis, L., (2001). Wired meetings: Technological mediation of organizational gatherings, In F. Jablin & L. Putnam (Eds.), *The New Handbook of Organizational Communication: Advances in Theory Research and Methods* (pp.624-663). Sage: Thousand Oaks.

Grudin, J. (1989). Why groupware applications fail: problems in design and evaluation. *Office Technology and People, 4*(3), 245-264.

Grudin, J. (1990). The computer reaches out: The historical continuity of interface design. In *Proceedings of CHI* (pp. 261-268). Seattle, WA: ACM Press.

Grudin, J. (1994). Groupware and social dynamics: Eight challenges for developers. *Communications of the ACM, 37*(1): 92-105.

Hayles, N.K. (1999). *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: Chicago UP.

Helmreich, R.L. (2000). Culture and error in space: Implications from analog environments. *Aviation, Space, and Environmental Medicine, 71*(9-11), 133-139.

Henderson, A., & Kyng, M. (1991). "There's no place like home: Continuing design in use". In J. Greenbaum & M. Kyng (Eds.). *Design At Work* (pp. 219-240). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Jaworski, A. (1993). *The Power of Silence: Social and Pragmatic Perspectives.* Newberry Park, CA: Sage.

Johannesen, R. (1974). The functions of silence: A plea for communication research. *Western Speech, 38,* 25-35.

Johnson, S., Kurtz, M., Tomlinson, T. & Howe, K. (1986). Students' stereotypes of patients as barriers to clinical decision-making. *Journal of Medical Education, 61*(9), 727-735.

Kensing, F. & Blomberg, J. (1998). Participatory design: issues and concerns. *Computer Supported Cooperative Work, 7,* 167-185.

Kohlase, C. (1995). The influence of culture and religious orientations on Latino attitudes toward illness, suffering, and death: Implications for medical decision-making, (Doctoral dissertation, University of Southern California, 1995).

Lazarus, R.S. & Folkman, S. (1984). Coping and adaptation. In W.D. Gentry (Ed.), *The handbook of behavioral medicine* (pp. 282-325). New York: Guilford.

Levinson, W. & Chaumeton, N. (1999). Communication between surgeons and patients in routine office visits, *Surgery, 125*(2), pp.127-134.

Mark, M.A. (1991). The VCR Tutor: Design and evaluation of an intelligent tutoring system. (Master's Thesis, The University of Saskatchewan, 1991).

McLuhan, M. (1964). *Understanding Media.* New York: Mentor.

Meshkati, N. (1996). Cultural factors influencing safety need to be addressed in design and operation of technology. *International Civil Aviation Organization (ICAO) Journal, 51*(8), 17-18, 27-28.

Newell, A. F. & Gregor, P. (2000). User sensitive inclusive design: In search of a new paradigm. *Proceedings on the 2000 conference on Universal Usability*, Arlington, Virginia; November, 39-44.

Niemeyer, S., Rai, V.K. & Kean, R. (1997). Older females' opinions about functionality of home products and equipment. *Journal of Family and Consumer Sciences, 89*(1).

O'Day, V. L., Bobrow, D. G. & Shirley, M. (1998). Network community design: a social-technical design circle. *Computer Supported Cooperative Work: The Journal of Collaborative Computing,* 7(3-4), 315-337.

O'Day, V. L., Bobrow, D. G. & Shirley, M. (1996). The social-technical design circle. In M.S. Ackerman (Ed.), *Proceedings of the ACM 1996 Conference on Computer Supported Cooperative Work* (pp. 160-169). November 16-20; Boston, MA. New York: ACM.

Ong, L.M., DeHaes, J.C., Hoos, A.M. & Lammes, F.B. (1995). Doctor-patient communication: A review of the literature, *Social Science and Medicine, 40*(7), 903-918.

Parks, M. R., & Roberts, L. D. (1998). 'Making MOOsic': The development of personal relationships on line and a comparison to their off-line counterparts. *Journal of Social and Personal Relationships, 15*, 517-537.

Ramachandran, V.S. & Blakeslee, S. (1998). *Phantoms in the Brain: Probing the Mysteries of the Human Mind,* New York: William Morrow.

Rector, A., Horan, B., Fitter, M., Kay, S., Newton, P., Nowlan, W., Robinson, D. & Wilson, A. (1992). User centered development of a general practice medical workstation: The pen & pad experience. In P.Bauersfeld, J.Bennett and G.Lynch (Eds.), *Proceedings of ACM Conference on Human Factors in Computing Systems, CHI 92* (pp.447-453), May 3-7, Monterey, CA. New York: ACM Press.

Rosen, A., Tsai, J., & Downs, S. (2003). Variations in risk attitude across race, gender and education, *Medical Decision Making, 23*, (511-517).

Simonsen, J. & Kensing, F. (1997). Using ethnography in contextual design. *Communications of the ACM, 40*(7), 82-88.

Smedley, Brian D., Adrienne Y. Stith, & Alan R. Nelson, Editors. (2003). Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. *Institute of Medicine Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care*. The National Academies Press.

Tajima, A. (2002). Fatal miscommunication: English in aviation safety. *World Englishes, 23*(3), 451-470.

Taveras, E. & Flores, G. (2004) Why culture and language matter: the clinical consequences of providing culturally and linguistically appropriate services to children in the emergency department. *Clinical Pediatric Emergency Medicine,* 5, 76-84.

Thomas, M.B. (1992). *An introduction to marital and family therapy.* Englewood Cliffs, N.J.: Macmillan Publishing Company.

Teutsch, C. (2003). Patient-doctor communication, *Medical Clinics of North America, 87*(5), 1115-1145.

Trevino, L. K., Lengel, R., & Daft, R. L. (1987). Media symbolism, media richness, and media choice in organizations: A symbolic interactionist perspective. *Communication Research, 14*, 553-574.

Ubel, P. (2005). Emotions, decisions, and the limits of rationality: Symposium introduction, *Medical Decision Making, 25*, 95-96.

Notes

[i] At a recent demonstration of Transonics using military personnel several users requested that the system be converted to an earpiece, resembling a telephone headset, that could translate words as they were spoken. This has also been a common request by those who have seen the device in other venues.

[ii] The intersection between technological determination and social determinism, or what can be created technologically and what should be created, is a conceptual debate that has been sparked in recent years by the claims of Ray Kurzweil (Age of Spiritual Machines). Kurzweil argues that within the next hundred years humans' flesh bodies will become outmoded, replaced by more efficient machine housings.

[iii] Media richness theory refers to a medium as "lean" or "rich" based on its ability to transfer information cues. E-mail would constitute a lean medium since it is asynchronous and primarily limited to text and pictures, while video conferencing would constitute a rich medium since real-time gestures and other nonverbal cues can be transmitted.

[iv] Systems theory (Ludwig von Bertalanffy) describes the concept of equifinality, or multiple paths to the same end.

[v] This study was conducted in accordance with the guidelines of the University Park Institutional Review Board. It was approved under USC UPIRB # 04-11-341.

[vi] In modern history the best examples may be the telephone and the desktop computer. Both of these devices were originally developed as business tools but quickly evolved to become forms of entertainment and means of developing interpersonal relationships.