SoftBridge in Action: The First Deaf Telephony Pilot

William D.Tucker, Meryl Glaser and John Lewis
Department of Computer Science, University of Cape Town, Private Bag Rondebosch 7700
Ph: 082 494 8884 Fax: 021 689 9465 {btucker,jlewis}@cs.uct.ac.za, mglaser@uctgsh1.uct.ac.za

Abstract—Following on several partial prototypes, we built an automated Deaf Telephony bridging application with the SoftBridge platform. The SoftBridge performs multi-modal bridging in real-time using Text-to-Speech (TTS) and Automatic Speech Recognition (ASR) utilities accessed via web services. To perform Action Research with the Deaf community in Cape Town, we invited a Deaf user (DU) to participate in a pilot study. We conducted three tests varying the input/output modalities of the hearing user (HU). The DU had a standard text in/text out Instant Messaging client. The HU client used the following specific modality combinations: Text & TTS in/Text out, TTS in/Text out and TTS in/Text & ASR out. The SoftBridge logged the conversations for subsequent analysis. The trial showed a largely successful conversation. Success factors include a) a text and computer literate DU who is familiar with research practise, b) using the system to explain the research as we conducted it and c) that the multi-modal bridging capabilities overcame the expected shortcomings of TTS and especially ASR. The lessons learned from this trial will be applied to the next trial once the necessary modifications have been implemented.

Index Terms—Action Research, Deaf Telephony, Multi-modal interface, Synchronous Messaging.

I. Deaf telephony

Without audio communication capabilities, Deaf individuals in South Africa are restricted to text-based telecommunication options, e.g. Teldem, SMS, Instant Messaging (IM) or fax. Of these options, only the Teldem currently provides synchronous communications without requiring a computer with Internet connectivity. However, the Teldem connectivity circle is extremely small, due the absence of a relay service to the Public Switched Telephone Network (PSTN). Our research activity seeks to eventually provide an automated relay system to enable Deaf users to communicate with all users, Deaf or hearing.

II. Background activity

Research began with Teldem trials in the Deaf community [1]. These trials identified shortcomings of a Teldem-only approach. The most significant need was to relay between text and voice to bridge the Deaf telecommunication to the hearing world. We proposed a series of bridges from human-based relay to fully automated relay with Text-to-Speech (TTS) and Automatic Speech Recognition (ASR) [2]. Human relay centres are established in many first world countries. Thus, the focus of this research is to examine automated systems. The first fully automated prototype was Telgo323, a system that bridged a PSTN telephone to a Teldem [3][4]. Because of poor ASR performance, Telgo323 was only implemented in one direction - to transform the text from the Teldem to speech for the telephone. There was no mechanism to get text back to the Teldem. This limitation prohibited trials with the Deaf community. We also ported this system to the Session Initiation Protocol (SIP).

III. SoftBridge

To provide two-way communication to the Deaf user (DU), another arm of the research involved the development of the SoftBridge [5]. This system provides real-time bridging services between multiple communication modalities. SoftBridge is based on the Jabber Instant Messaging protocol, therefore even audio communication is treated as a series of short messages. The SoftBridge employs web services to perform adaptation between modalities, e.g. voice to text with ASR or text to voice with TTS. The web services approach easily accommodates improvements in adaptation technologies. The SoftBridge has the additional benefit of connecting disparate endpoints, e.g. telephone and PC. These characteristics make the SoftBridge an ideal platform for building Deaf telephony applications. Recognising that ASR technology is not yet conducive to general purpose communication, we can build variants of Deaf bridging applications without ASR. However, these applications can still employ TTS to good effect. This allows us to trial SoftBridge applications with Deaf users at this stage.

IV. Experimentation

The trial involved two computers in two nearby rooms. The computer for the DU ran an unmodified Exodus IM client which connected to the SoftBridge over Internet Protocol (IP).The DU client only performs Text in and Text out. The other computer housed the SoftBridge, web services and a client for the hearing user (HU) (in this case, the researcher) that allows one to toggle text and audio capabilities. We conducted three tests, varying the capabilities of the HU client.

A. Text & TTS in/Text out

This test mimics an IM chat session, as well as a text telephone, but with the added feature of being able to "hear" the DU's text.

Experiences. Instructions were initially interpreted from spoken English into South African Sign Language (SASL) by a researcher. However, we quickly shifted to use the SoftBridge system to inform the DU about the tests.

User Feedback. The DU was concerned that the HU would not be able to understand "Deaf speech", e.g.
[11:47]<david> for example deaf people write like that - me come see u tonight

which includes both grammatical differences (from SASL) as well as spelling abbreviations and anomalies.

Lessons Learned. People who are familiar with the Teldem need to be instructed that a half-duplex protocol is not necessary for this system. The DU waited for a reply instead of typing at will. We also saw an instance where the DU habitually used the "GA" at the end of his turn, indicating the Go Ahead signal conventionally used in Teldem conversations, e.g.

[12:48]<david> i do not understand pls reply ga

We used the system to familiarise the DU to the trial as opposed to SASL interpreting. This immediately elicited a meaningful conversation between the researcher and the DU. We also noted that this level of conversation was possible because of the high English and computer literacy of this particular DU.

Having the 2 computers situated close together distracted from a realistic setup because one researcher could move between rooms to facilitate the conversation. We should therefore move to a long distance model for future trials.

B. TTS in/Text out

The Deaf client remains text in/out, but the hearing client toggles off the incoming text so that only the TTS is heard.

Experiences. The TTS messages arrive abruptly and are quickly delivered. Thus, the researcher had to concentrate more than for the previous test and would prefer to playback the voice messages again and/or visualise the conversation thread.

User Feedback. The DU is concerned about Deaf literacy. He asserts that Deaf users would be able to understand one another despite degraded English, but an unfamiliar HU would not. The DU also claims that this system is better than SMS because of the fast turn around time.

Lessons Learned. Presence indication is needed to alert participants that a message is being prepared. The DU was also not aware that text is not being sent character by character, but rather that pressing Enter sends the message.

C. TTS in/Text & ASR out

The DU client remains text in/out. The HU client keeps the incoming text hidden and also uses voice and ASR to send messages to the DU.

Experiences. We expected, and got, poor ASR performance. To improve the recognition rate, the researcher spoke with artificial pauses between words and overly careful articulation. However, the poor ASR often forced the researcher to employ text to clear up misunderstood output from the ASR service.

User Feedback. This particular DU demonstrated well-honed repair skills. He used effective requests for clarification when the ASR produced unintelligible messages, e.g.

[12:45]<david> start again
[12:45]<bill> And K.
[12:45]<david> what k?

Because we found the ASR output confusing, we marked all typed messages to differentiate the message source to keep the conversation flowing.

Lessons Learned. The various modes of communication need to be automatically annotated in the recorded conversation log. For example, text typed at the keyboard must be differentiated from text output from the ASR service.

V. NEXT STEPS

Both DU & HU clients require presence indicators. The research conversation log requires automated annotation for log analysis tools. Teldem-ese, eg. *pls call tmw*, needs to be inserted into the TTS dictionary.

Overall, most of the software manipulation is for the HU. The text in/out interface for the DU is unchanged from Teldem or fax. The HU, however, experiences TTS output and has to adjust to ASR input. The HU also has to deal with multi-modal input and output as opposed to only voice on a telephone handset, or only text for Internet IM. Our DU also repeatedly pointed out that it would be the HU who would struggle with this system due to poor language skills in the Deaf. This DU had a high level of both text and computer literacy. Therefore, we will conduct trials with more Deaf users who display more typical text and computer literacy.

REFERENCES

[1] M Glaser, "A Field Trial and Evaluation of Telkom's Teldem Terminal in a Deaf Community in the Western Cape", South African Telecommunication Networks and Applications Conference (SATNAC 2000), Stellenbosch, South Africa, 11-13 September 2000.

[2] M Glaser and WD Tucker. "Web-based Telephony Bridges for the Deaf", South African Telecommunications Networks & Applications Conference (SATNAC 2001), Wild Coast Sun, South Africa. Sep 3-5, 2001.

[3] J Penton, W Tucker and M Glaser. "Telgo323: An H.323 Bridge for Deaf Telephony". South African Telecommunications Networks & Applications Conference (SATNAC 2002), South Africa. Sep 1-4, 2002.

[4] W Tucker, M Glaser and J Penton. "A Bridge for the Problem of Deaf Telephony". Science in Africa, June 2002.

[5] J Lewis, W Tucker and E Blake. "SoftBridge: An Architecture for Building IP-based Bridges over the Digital Divide". South African Telecommunications Networks & Applications Conference (SATNAC 2002), Drakensberg, South Africa. Sep 1-4, 2002.

William Tucker is on study leave doing a PhD at the University of Cape Town. He leads the Broadband Applications and Networks Group at the University of the Western Cape, where he is a Senior Lecturer.