THE SYNFACE PROJECT: DEVELOPMENT AND EVALUATION OF A TALKING FACE TELEPHONE

ABSTRACT

During face to face communication, many hard of hearing people use lip reading to support their residual hearing. On the telephone this is not possible, and communication can be difficult. The Synface project is developing a talking face; it will use the sound from the phone line to generate realistic lip movements in near real time to support telephone communication. The prototype that will be used for evaluation trials, and concepts for the future product, are in development. The preliminary prototype interface has been subjected to a heuristic evaluation, the findings of which are summarised here. In addition, extensive user evaluation work using a number of different methods is in progress. Laboratory trials involving both hearing and hard of hearing people have been conducted. The final phase of testing will involve a broad range of methods for home, work and laboratory based trials. This paper gives an overview of the project, presents the work to date and discusses the future possibilities for this technology.

Keywords

automatic speech recognition, usability, telephone, lip reading, hard of hearing, avatar

1. INTRODUCTION TO SYNFACE

The Synface project aims to improve the accessibility of voice telephones for hard of hearing people. This group of people is handicapped in telephone use because of the lack of visual information; they are not able to use lip reading to support the auditory information they receive.

Synface uses automatic speech recognition of the sound received down the telephone line. The speech is recognised at the phoneme level, and visemes are generated on a synthetic face. The face lip speaks the callers' words in near real time. This provides the hard of hearing telephone user with additional visual information to facilitate their conversation.

There are advantages of this system over the existing options of video and text telephony. Videophones require both users to have expensive, specialist equipment. Text telephony may be used via an operator relay service, such as RNID Typetalk, run in conjunction with British Telecom in the UK, but the user may feel a loss of intimacy. To have a conversation using Synface, only the hard of hearing user, not the person they are talking with, must have specialist equipment. The equipment can be used via a normal telephone line, and the caller need not know that the hard of hearing person is using an assistive technology.

The concept was originally developed during the Teleface project [1], using the Swedish language. The present project aims to add real time speech recognition technology, and the ability to work with English and Dutch.

2. PROJECT PARTNERS

Five partners are involved in the Synface project. The Department of Speech, Music and Hearing at KTH (Royal Technical University, Stockholm) is responsible for the automatic recognition of speech, for the extraction of articulatory movements and for the development of the synthetic face. The Department of Phonetics at University College London (UCL) is working on the definition and evaluation of the visual speech information required for auditory-visual communication. Babel-Infovox AB, Stockholm, is the commercial partner. Their responsibilities include carrying out market surveys and building a prototype for evaluation.

Two user organisations are involved; Viataal, Netherlands, and the RNID (Royal National Institute for Deaf People). The responsibility of these organisations is principally to assist in the definition of the specification for the system design and carry out user evaluations of the prototypes.

2.1 Why is RNID Involved?

RNID is the largest charity representing the 9 million deaf and hard of hearing people in the UK. The work that the

Technology Directorate undertakes is wide-ranging, At present there is involvement in three European Union (EU) funded projects, in addition to undertaking research and development work in the fields of Information and Communication Technology and Product Development. The majority of this work is collaborative, working with academic and industrial partners all over the UK and Europe.

In this project our role is to ensure that the needs of hard of hearing people, the intended end users of the product, are considered throughout the research and development process. This involvement has included a market research data collection exercise, and taking a lead role in planning for the prototype, final product and evaluation trials.

3. THE CHALLENGES OF THE PROJECT

The Synface project must result in a system that takes arbitrary acoustic speech signals from a telephone lines and controls a 3D model of a talking face. The system must generate prominent information-bearing oral movements derived from the speech signals in near real time. There are a number of technological tasks involved in this. Towards the end of the project, a prototype will be assembled and evaluated in the UK, Netherlands and Sweden. It is hoped that the product could be available on the market within five years.

In a project of this type, where the technological challenges are so great, it is of particular importance to focus on the needs of the end users, and upon high quality evaluations of the product. Key to this task is the design of the prototype, and the methods used for the evaluations.

4. PROTOTYPE AND PRODUCT DESIGN

The prototype to be used in the final evaluation studies will run on a laptop. The final product as it would reach the market may not be computer based, rather it could be a dedicated device with a screen and telephone as the only (visible) parts of the system.

The target market for Synface is hard of hearing people, most of whom are older adults [2]. While it is of course vital that the system works, that the lip movements are generated accurately and in near real time, it is also important to make prototypes and products that are user friendly and that people will be comfortable using. Trials of a good technology would be compromised if the system were presented poorly. As such, effort has been invested in consideration of the prototypes appearance and interface.

4.1 Prototype: Hardware

Many older adults do not regularly use a computer [3]. It may therefore impede the user trials to present Synface as a computer. At present, though, the technical requirements of the software are such that a computer is essential.

One possible solution to this is to use the type of set up illustrated below (Figure 1). Here a laptop with a custom made metal plate over the keypad is shown. This simplifies the appearance of the computer, avoiding confusion and enabling the user to access only those functions that are necessary to use the prototype. A similar set up has been successfully used before by RNID, in trials of prototypes of a new type of textphone aimed at older, hard of hearing people. It has the advantage of enabling the user to keep their old telephone, with a simple add-on providing the extra facility.

Figure 1 The Prototype concept for Synface home based user trials

4.2 Prototype: Interface

As part of this project, an interface must be designed for use in evaluation trials. This, like the appearance of the prototype hardware, will be critical to the success of the testing. A graphical user interface has been developed [4], incorporating the talking face in addition to telephone controls and a personal phone book (Figure 2). This interface was subjected to a heuristic evaluation by RNID, to ensure that it conforms to general usability guidelines. It is at present being redesigned to take into account the findings of this evaluation.
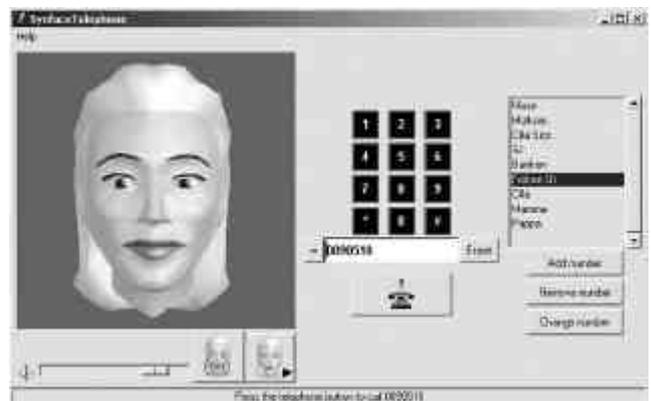

Figure 2 The initial prototype interface

Four members of the Technology Directorate at RNID carried out the heuristic evaluation. All had expertise in product technology or ergonomics, and a high level of familiarity with the needs of the target population. A number of specific usability problems were indicated. By compiling these some more general areas for development were highlighted;

- the layout was not designed to occupy a full computer screen (see Figure 2) – it is to be resized to conform to standard screen dimensions;

- the layout afforded no clear information on workflow – the new layout should take this into account;

- the components (face, phone controls, phone book) were small – components to be rearranged to make better use of space;

- delineation of individual buttons and components poor – consider outlining buttons and components more boldly;

- text was small, and presented in black on a mid grey background – font, style and colours to be reviewed;

- some of the icons used were ambiguous – alternative icons to be designed and evaluated, and text labels on buttons as an alternative to icons to be considered;

- the system status information was present but was small and peripherally located – means of drawing the users attention to this information, e.g. highlighting, to be considered.

4.3 Product: Rationale

It is likely that different user groups will have different requirements of the Synface product. Someone using it to support them in an office job may find that installing it on their desktop computer is most convenient for them. It may be possible to run the software on a handheld computer so that the device can be used on the move with a mobile phone. Alternatively, an older adult using the technology at home may prefer to have a small screen next to their existing telephone on which the face would be viewed.

The final product should therefore be available in a number of different formats to suit the needs of different user groups.

5. EVALUATION METHODOLOGIES

Various methods have been and will be used to evaluate the Synface prototype. These have been planned according to stage of development of the system. In addition, they are based upon the need for a broad range of information and feedback. The data collected looks at a number of factors. These include general usability, the degree to which the lipreading support helps with communication, and general acceptability of the Synface system. Both qualitative and quantitative, subjective and objective measures will be taken.

5.1 Evaluations Using Hand Labelled Speech

Evaluations to date have been laboratory based, and focussed on the comprehension of speech when using Synface. The automatic speech recognition system and the lip speaking components of the system are being developed in parallel. In these trials, therefore, pre-recorded speech was hand labelled, with the phonemes being added manually.

Two types of trials have been conducted to assess the effect on comprehension of speech with the Synface system. Early trials used degraded speech, and normally hearing participants. Participants listened to the degraded speech accompanied by no visual support, by the synthetic face, or by a video of the talker [5].

Later trials in the UK tested the system with hard of hearing people, listening to speech that was not degraded, but that was telephone band limited. In addition to the measurement of sentence comprehension, post test questionnaires were completed by the participants. Here participants indicated whether they thought Synface was useful, and whether they would like to use it in the future.

It was demonstrated for both participant groups, and in all three languages, that speech comprehension was poorest with no visual information, and best when a video of the speaker was viewed in addition to receiving the auditory information. Performance with Synface was between these two levels; better than where no visual information was available, but not as good as with video of the speaker.

The post test questionnaires administered to the hard of hearing group provided valuable general information. Of the 13 respondents, six reported that the system would be very useful to them, and six said that it would be useful. One did not know if it would be useful or not. Eleven respondents indicated that if this system was available to them, they would use it at home, and seven that they would use it at work. Ten of the respondents indicated that they would rather use Synface than an operator relay service such as RNID Typetalk. The majority of respondents said that they would be willing to pay between £100 and £500 for the product. Respondents were asked whether they would consider buying the system if it was available as a special telephone unit or as a computer add-on; twelve said yes to both of these options, and one said maybe to both.

In addition to these structured questions, participants were asked to write about the good and bad points of the system as a means of communication, and about what they would like changed to improve it. Most respondents liked having the opportunity to lip read to support their residual hearing. The most common negative comment was that the face looked artificial, and that it lacked expression. While expression would be technically very hard to incorporate into this system, it was suggested that some simple additions like natural movement of the head and eye blinks may help to resolve this.

The feedback received in these trials has been generally positive. Synface was effective in facilitating hard of hearing peoples' understanding of auditory information. Participants also responded well to the technology, with most indicating that they would be keen to use it in the future, and would consider purchasing it for use at home or work.

5.2 Evaluations of the Prototype System

The final evaluations will test the full prototype system and will take place in the home, workplace and laboratory. In these trials users will have real time conversations over the telephone, thus using Synface for its intended purpose.

In the UK, it is planned that home based user trials will take place. Synface will be installed in peoples' homes and they will be able to use it for all their telephone calls. In addition, the trial coordinators will call all participants periodically to ensure that they use the system regularly and are becoming accustomed to it. This calling will also enable the coordinators to monitor user progress with the system, to help resolve any difficulties and to administer short structured interviews.

In Sweden the trials will be workplace based, using PC's set up at KTH. The Netherlands trials will be laboratory based. These evaluations will simulate home and workplace use of Synface, but the laboratory situation will enable more detailed observations of user performance.

6. SUMMARY

The Synface project is working towards providing lip reading support for hard of hearing voice telephone users. The evaluation work that is taking place within the project will ensure that the final product meets the requirements of the user group, and that it is usable and accessible. Studies carried out to date indicate that general acceptance of the system is high, although further consideration must be given to the interface and to the product itself.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] Beskow J., Dahlquist M., Lundeberg M., Spens K-E and Ohman T (1997). The Teleface project: Multimodal speech communication for the hearing impaired, *Proc of Eurospeech*.

[2] Davis A. (1995). *Hearing in Adults*. Whurr Publishers Ltd.

[3] Rousseau G. and Rogers W. (1998). Computer usage patterns of university faculty members across the lifespan, *Computers in Human Behavior*, **14**, 417-428.

[4] Ward K. (2002). User interface for the Synface project, *MSc thesis report, TMH-KTH (http://www.speech.kth.se/ctt/publications/papers/exjobb_kw.pdf)*.

[5] Siciliano C., Williams G., Beskow J. and Faulkner A. (2003). A multilingual synthetic face as a communication aid for the hearing impaired, *15th International Congress of the Phonetic Sciences*.