Presenting a speech-based mobile reminder system

Arjan Geven[1] Manfred Tscheligi[1] Alex Sorin[2] Hagai Aronowitz[2]

[1] CURE – Center for Usability Research and Engineering
[2] IBM Research Laboratory in Haifa, Israel

geven@cure.at

ABSTRACT

Aging is often accompanied by various forms of physical and cognitive decline. In this paper, we present a mobile guide that is aimed at people suffering from this cognitive decline – a memory aid at home and in mobile settings – called HERMES. A mobile guide allows the user to record events for later re-call (episodic memory) as well as capture information about future events and remind the user about when he or she might forget these otherwise (prospective memory). In this paper, we focus on the mobile environment of the technology. The mobile guide is similar to a memo-recorder: the PDA or mobile phone records speech both at home and in other environments. At a home server, the captured speech data is processed into text and information is extracted from the speech data and entered in the user's calendar. Depending on how the user wants to be remembered, the system provides the user with a reminder at a certain time or at a certain place.

Keywords

user experience, methodology, conjoint, mobile device evaluation

1. INTRODUCTION

Aging is often accompanied by various forms of physical and cognitive decline. Particularly, cognitive decline in aging brings along with it the reduced capabilities in working memory and information processing, a reduced ability to encode new memories of events and facts, and particularly remembering events (episodic memory), planning of events in the future (prospective memory) or the source of specific information (who told me about it?). This results in a loss of detail of memories and, more severe, the inability to plan one's own life and the accompanying reduced quality of life.

The main target group includes elderly people without any cognitive impairments other than AAMI (Age Associated Memory Impairment), that is defined as a specific impairment in memory, not affecting other cognitive areas and that does not have effects on the Activities of Daily Living (self-care, hygiene, dressing, etc), but affects their normal functioning because it interferes on the remembering of daily information. This impairment is often self-reported by the person him/herself and sometimes confirmed by another informant.

In some cases, patients with Mild Cognitive Impairment (MCI) will also benefit from the HERMES solution. The MCI is a defined condition in which a person has problems with memory, language, or another mental function severe enough to be noticeable to other people and to show up on tests, but not serious enough to interfere with daily life. Because the problems do not interfere with daily activities, the person does not meet criteria for being diagnosed with dementia. The best-studied type of MCI involves a memory problem and is called "amnestic MCI."

The HERMES project develops a system to support these age-related problems, specifically aiming at supporting aging people, who are suffering from normal, age-related, cognitive decline.

The cognitive support in HERMES relies on efficient and effective capturing of information, in an in-between form of capture that we call *semi-automatic capture*, combining the advantages of manual and automatic recording. HERMES is aimed to be *context-aware*, recording not only the important moment, but also a set of parameters that can be helpful in later retrieving such moments, as effective capturing has to deal with ways to later also effectively being able to retrieve this information. To this end, the technology relies strongly on the storing of *contextual information*.

Besides cognitive support in the form of a memory aid, HERMES also provides the user with cognitive training through games that function based on the content of the units. It thereby offers memory training, but also speed and reasoning capabilities can be trained through the game-like cognitive training program that is offered.

The system is based on a mobile part and a stationary part. The mobile part is represented by software running on an off-the-shelf mobile phone (Figure 3) running Windows Mobile, which captures mainly audio data. The stationary part is installed in the user's home and captures audio and video with a limited set of recording technologies (omni-directional microphones and video cameras). The system then processes the information and stores this in a journal which the user can access: a "memory of the past" or episodic memory. In addition to this memory of the past, the HERMES mobile part provides the user with a "memory of the future" or prospective memory aid.

2. RELATED WORK

An example of a simple Assistive Technology for Cognition is Neuropage, a wearable system in which the reminders are

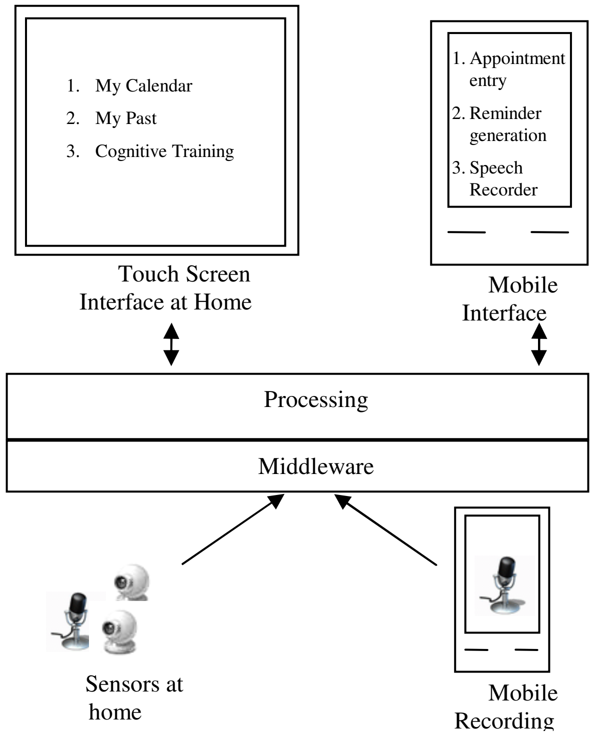

Figure 1: Overview of the HERMES system

inserted for each person and are sent to the beeper of the user in the appropriate moment (Wilson1997). A bit more advanced is MEMOJOG (Szymkowiak2004), a system that provides to the user memory predetermined messages at the suitable moment about the next activities of the day. The reminders are announced by an auditory alarm. The users, or their caregivers through a PC with Internet, can insert the reminder into the device.

When we direct our attention specifically to daily activity support and the remembering of tasks, the systems developed for planning and task management seem most relevant to discuss. A number of devices have been developed to assist older people through more or less advanced scheduling of tasks and provision of reminders. Nowadays the systems use pagers, cell phones and palmtops, but many of them can be described as alarm clocks with some extra bells and whistles. One of the first systems to employ AI planning techniques for greater flexibility was PEAT (Levinson 1997; Levinson 2004).

One of the more advanced systems of this kind at the moment is Autominder (Pollack et al. 2003). It employs Bayesian inference techniques for scheduling. The plans it creates are described in the language of disjunctive temporal problems (Tsamardinos and Pollack 2003), which make it possible to model flexible time constraints. This is important because rigid scheduling can lead to user dissatisfaction or excessive reliance on the scheduler (since it may not give the user the time to remember tasks by itself). In some situations, the reminders are perceived for the elderly such an intrusion. The elderly people feel that the device can make them more dependent and not force them to use his memory abilities. For this reason it's very important to give the announcements in a transparent non-obtrusive manner: only at appropriate moments and when absolutely required. The system should know when it is time to be quiet and when it can speak up to the user.

All schedulers and management systems to support older people however have been unsuccessful to provide real context support and require relatively many steps in their interaction to enter a new event.

3. VISION AND CHALLENGES

Many studies have found that despite older people's awareness of new technology and its potential to support independence, uptake is often fairly low. It is very common to find among elderly people rejection feelings toward new technologies. One of the most frequent matter commented by them is the worry about causing any damage to the technological system.

If devices and interfaces for computational and communicative equipment are going to be easy to use, the age-related decline in physical and cognitive functions must be taken into consideration. Today many interfaces are difficult to use, and they do not take older users presuppositions into account. It has for example been shown that older users face larger difficulties in learning and using different interfaces and computer applications and that they have more difficulties in handling large information spaces and sorting out task relevant information (Hirsch et al., 2000). Some of these age differences can be explained by the age-related cognitive decline.

Knowledge about the cognitive decline and aging process in general must be applied to the design of interfaces for older users. There is need to understand what age-sensitive cognitive functions play the largest role in different situations when conducting computer-related tasks, and also how to compensate for this age-related cognitive decline when designing interfaces.

Many older people downplay the need for the technology, and might even chose a radical change in their lifestyle and behaviour, even at the expense of independence or social interaction, rather than rely on a device that makes them feel embarrassed or incapable (Hirsch et al., 2000). Assuming that older people will use an assistive technology simply because they need it is therefore misguided. HERMES therefore addresses emotional and societal needs and actively supports the current lifestyles and behavior of the older person.

To support the user in performing everyday tasks and to assist independent living, HERMES is designed to give advanced activity reminders to support the prospective memory. The vision and the ultimate goal of HERMES project is to establish a personal external memory modelled after human associative memory. Contextual cues remind the user automatically and non-disruptively, facilitating remembering at the right place, at the right time. The reminders are automatically deduced as far as possible from the captured audiovisual information including the different kinds of context that are recorded to reduce the burden on the user.

Deducing this information is a challenging task which is based on the automatic transcription of spontaneous speech recorded in various environments, semantic annotation of the transcript and decision making. Transcription of spontaneous speech is a challenging task by itself (Kingsbury et al 2003). Elderly speech (Byrne et al 2004) and far field recording complicate it even more. In this area we will explore capabilities stemming from the personal nature of HERMES system. Large volumes of speech data produced by a small number of known speakers and accumulated by HERMES system during the day-by-day use can be used for unsupervised adaptation of the ASR acoustic models. This way the speech transcription system can improve its performance as time goes on. The semantic annotation and reasoning can be underpinned by HEREMES *world model* implemented by means of a knowledgebase encompassing several domains of interest, e.g. health, shopping.

Acknowledging the challenge posed by the automatic reminders deduction, we start with an intentional spoken appointments generated by the user. First, this establishes an initial phase for the exploration and development towards the automatic deduction of appointments from spontaneous speech. Secondly, this application by itself being properly designed is potentially valuable for the elderly.

As long as we know that the HERMES system is not infallible, we use certain fixed points that the user enters in addition to free speech, to support the information entered by speech. Namely, the appointment time will be entered by means of keyboard, touch-screen or TV remote control. If the trigger for the reminder is a location or a situation rather than a time moment then the time entered by the user will be interpreted as the moment when the appointment should be offered to the user for reconsideration if it is still pending. After making a decision with respect to the intended trigger of the appointment, the system asks the user to confirm the newly pending reminder.

Currently, we are investigating how users deal with this kind of intelligence and how we can design to accommodate for this.

In the next section we describe the process of placement, processing and issuing spoken reminders.

4. HERMES MOBILE REMINDERS

The goal of the HERMES mobile reminder system is to make these reminders as lightweight as possible. For this to happen, we employ a number of steps:

- First, the user inputs speech into the mobile device: the user clicks a "record event"-button, selects the right date/time for the event and then speaks freely any other related information into the system. At this point, the mobile device only records this information and stores it on the device.
- When the user is at home, the second step is initiated. The mobile device communicates with the home server and transfers the recorded data with the attached metadata (date, time), to the home server. This is a step that does not require any interaction from the user.
- In the third step, the home server transcribes the recorded speech and annotates the appointment in terms of *whith whom, where, triggering event*. The annotated appointment is indexed in order to provide a capability of fast retrieval of information from user's calendar.
- In the fourth step, the processed pending reminder is sent back to the mobile device which is able to display the reminder related information to the user and to play back the original audio record.
- Finally, the user receives a reminder at the relevant place and/or time. A few examples follow. The user is reminded to be at home on 5pm when his/her friend is expected to pay a visit. The user is reminded to buy coffee when he/she is near the grocery. The user is reminded to ask the cleaner to clean the window when the cleaning activity is going on at home.

Figure 2 illustrates the steps of the process described above.

This facilitates the following objectives that are defined in the project:

- HERMES issues audio-visual reminders, i.e. by synthesized voice messages, sounds and visual display. The audio and visual modalities are used together or separately depending on the user context and reminder's content.
- Combinational analysis of input information via text, speech, context, and visual processing technologies will produce inputs for reasoning.
- Context acquisition (Who, What, Where, When, Why), automated scene analysis, advanced context modelling and situation identification and conversation analysis lead to the recognition and identification of such an event. Reminders are generated and timed to help with management of daily activities as well as prospective memory aid for future 'to-dos'.
- The system facilitates the information previously captured only when is essential. The system avoids giving information that the older person can remind on his own.

5. STATUS AND NEXT STEPS

Currently, we are addressing both the interface design of the system and the speech recognition system. The interface has to empower the user, without embarrassing the user or making him or her feel incapable. Therefore, we chose a device that combines modern technology (touch screen, large display), with easy-to-configure interaction possibilities. The touch screen allows us to create buttons at any size that do not pose problems to our users. The interfaces on the mobile device and the home server will be based around the same touch-based interfaces for consistency. At this moment, we are developing the mobile interface concept for reminding. The rules for reminder generation and their interface as well as GPS-based location reminders based on specific POIs (grocery store, pharmacy, a friend's place) are of interest to us.

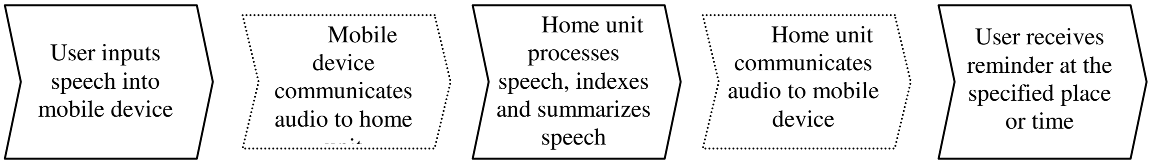

Figure 2: Steps that an appointment goes through

On the speech side, we are collecting the necessary speech corpora to be able to fine-tune the speech processing algorithms to the specific vocabulary and voice characteristics of older people to improve speech recognition and reduce the need for fallback mechanisms and improving the user's acceptance of the system through technological development

A close collaboration is required between the user interface design and the speech processing developments in this project. The user's acceptance of the system relies on making sure that the user is supported in his or her daily life activities. This also means managing expectations correctly with regards to the necessary recognition errors that will occur and how the system and the user deal with this.

Figure 3: Mobile touch screen device

6. ACKNOWLEDGEMENTS

The research presented in this work has been supported by the EU ICT project HERMES (Ref: FP7-ICT 216709). The authors acknowledge help and contributions from all partners of the project.

7. REFERENCES

LoPresti, E.; Mihailidis, A. & Kirsch, N. (2004). Assistive technology for cognitive rehabilitation: State of the art. *Neuropsychological Rehabilitation, 14*, 5-39

Wright, P.; McCarthy, J. & Meekison, L. (2005), *Making sense of experience*. In: Blythe, M.A., Overbeeke, K., Monk, A.F. and Wright, P. (2004). Funology: from usability to enjoyment Kluwer Academic Publishers, pp. 43--53

Wilson, B.A. and Evans, J.J. and Emslie, H. and Malinek, V. (1997). Evaluation of NeuroPage: a new memory aid. Journal of Neurology, Neurosurgery, and Psychiatry, 1997, 63:113-115.

Szymkowiak, A., Morrison, K., Inglis, E. A., Gregor, P., Shah, P., Evans, J. J., & Wilson, B. A. Memojog - an interactive memory aid with remote communication. , Paper presented at the Workshop on Universal Access and Assistive Technology (CWUAAT) 22nd-24th March 2004, Cambridge, UK.

Levinson, R. 2004, A Custom-Fitting Cognitive Orthotic That Provides Automatic Planning And Cueing Assistance, Proc. Technology And Persons With Disabilities Conference 2004, CSU Northridge,

Levinson, R. 1997, The Planning and Execution Assistant and Trainer. Journal of Head Trauma Rehabilitation, April, Aspen Press.

Tsamardinos, I. & Pollack, M. (2003), 'Efficient solution techniques for disjunctive temporal reasoning problems', *Artificial Intelligence* **151**(1-2), 43—89.

Pollack, M.; Brown, L.; Colbry, D.; McCarthy, C.; Orosz, C.; Peintner, B.; Ramakrishnan, S. & Tsamardinos, I. (2003). Autominder: an intelligent cognitive orthotic system for people with memory impairment. *Robotics and Autonomous Systems, 44*, 273-282.

Hirsch, T.; Forlizzi, J.; Hyder, E.; Goetz, J.; Kurtz, C. & Stroback, J. (2000), The ELDer project: social, emotional, and environmental factors in the design of eldercare technologies, in 'Proceedings on the 2000 conference on Universal Usability', ACM Press New York, NY, USA, , pp. 72--79.

Kingsbury, L. Mangu, G. Saon, G. Zweig, S. Axelrod, V. Goel, K. Visweswariah, and M. Picheny (2003). Towards domain-inependent conversational speech recognition, In Eurospeech,, Geneva, Switzerland, September 2003.

W. Byrne, D. Doermann, M. Franz, S. Gustman, J. Hajic, D.W. Oard, M. Picheny, J. Psutka, B. Ramabhadran, D. Soergel, T. Ward, and Wei-Jing Zhu, (2004) Automatic recognition of spontaneous speech for access to multilingual oral history archives, IEEE Transactions on Speech and Audio Processing, Special Issue on Spontaneous Speech Processing, 12(4):420-435, July 2004.