Exploring Seniors Interacting with Voice User Interfaces

Stephan Schlögl
Institut Mines-Télécom
Télécom ParisTech
Paris, France
schlogl@telecom-paristech.fr

Markus Garschall
CURE Center for Usability
Research and Engineering
Vienna, Austria
garschall@cure.at

Grégory Legouverneur
LUSAGE Living Lab
Broca Hospital
Paris, France
gregory.le-gouverneur@brc.aphp.fr

Gérard Chollet
Institut Mines-Télécom
Télécom ParisTech
Paris, France
chollet@telecom-paristech.fr

Manfred Tscheligi
CURE Center for Usability
Research and Engineering
Vienna, Austria
tscheligi@cure.at

CHI'13, April 27 – May 2, 2013, Paris, France.
ACM 978-1-XXXX-XXXX-X/XX/XX.

Abstract
Recent product releases such as Apple's *Siri* and Google's *Voice Search* have strongly emphasized the use of voice as a modern interaction modality. This possibility of talking to a computing device may not only increase its overall user experience but might also have the potential to offer significant support to people who struggle with physical interaction channels. Seniors, in particular, would often appreciate an alternative to small mobile phone keypads, laptop touchpads and computer mice. This paper presents initial explorations of how seniors may interact with language-technology-driven interfaces, how these interactions measure up against traditional physical interaction channels, and what features they require to satisfy the needs of this specific user group.

Author Keywords
Ambient Assisted Living; Voice User Interfaces

ACM Classification Keywords
H.5.2 [User Interfaces]: Evaluation/methodology; H.5.2 [User Interfaces]: User-centered design; H.5.2 [User Interfaces]: Voice I/O

General Terms
Design, Human Factors

Introduction

We are living in a world that increasingly faces the challenge of technological divergence - a separation which is not only driven by people's social or economical background but also related to an increasingly aging society. Consequently, we may find people who want or have to operate technological devices and services while struggling with some type of physical or cognitive limitation. In 2010 17.38% of Europe's population was older than 65 years of age and soon we will have less than two people of working age for every person beyond 65 [4]. A better adaptation of products and services to the cognitive and physical abilities of this constantly increasing group of users seems therefore crucial.

Ever since the presentation of Apple's *Siri*[1] and Google's *Voice Search*[2] Voice User Interfaces (VUI) are booming. Companies increasingly provide us with potentially new application scenarios where spoken language is seen as the best way of operation. Efficient methods for designing these VUIs and adapting them to the particular interests of a specific user group are, however, missing.

The vAssist project aims at developing VUIs that are optimized for elderly people suffering from age-related restrictions. In order to gain a first impression of the potential design space and possible user requirements for this type of interfaces we conducted two rounds of focus groups in which participants in Austria and France were asked to use both physical and (prototypical) voice-based interaction channels. The following sections describe our approach in more detail and report on initial feedback and some general impressions which we would like to use as a basis to start a discussion about the challenges of designing technology for seniors and people suffering from physical and/or cognitive restrictions.

The vAssist Services

Ambient Assisted Living (AAL) solutions aim at supporting seniors in their everyday life. One particular aspect that AAL is dealing with is the prevalence of fine-motor problems and/or chronic diseases. European statistics show that currently 1.2 million senior citizens suffer from Parkinson's[3] and 630.000 from multiple sclerosis[4]. Some of our future technologies should therefore not only offer advanced services that help mastering health-related difficulties, but also provide adapted interaction paradigms that allow for the compensation of physical and cognitive restrictions. The vAssist project wants to address this challenge by developing a set of voice-operated health and communication services. Planned applications include audio and video calling, messaging (text and email), contact management, calendar and information search, as well as a well-being diary and the possibility to play cognitive games [9].

While different services will be independent from each other, the vAssist platform as a whole should be experienced as an integrated product. In order to better convey the feeling of interacting with a single system, our goal is therefore to resemble something that is close to a virtual 'butler'; similar to what is described as the 'Memex' in Bush's famous article from 1945 [1]. While we generally see great potential in using natural language as a means to augment the life of seniors, the adaptation of

[1] http://www.apple.com/ios/siri/
[2] http://www.google.com/mobile/voice-search/
[3] http://www.parkinsonsawareness.eu.com/campaign-literature/prevalence-of-parkinsons-disease/
[4] http://www.who.int/mental_health/neurology/Atlas_MS_WEB.pdf

the interaction style to the needs and cognitive skills of this specific user group, however, needs to be taken into account. For example, we required a simplified and adapted language that is free of technical terms. In addition effective feedback and guiding mechanisms should be in place so that users easily find their way. Finally, alternative interaction channels, which cover use cases where voice-based service operation is insufficient or inappropriate, may also be necessary.

Defining the User Group

An obvious starting point for building technology and services for seniors is a clear definition of potential users and their distinct requirements. Following this motivation we have separated future vAssist users into three classes:

- Primary Users (seniors)
- Secondary Users (formal/informal caregivers)
- Tertiary Users (service providers)

Primary users in vAssist are defined as persons aged 65 and older. These persons may show physical limitations as well as restrictions in their vision and hearing, and/or show types of AAMI (Age Associated Memory Impairment). Secondary users are divided into formal and informal caregivers. Formal caregivers are health professionals with a specific (academic) medical education who receive payment for their work and efforts. This includes general practitioners (GP), but also specialists such as neurologists, psychologists, gerontologists, nurses, nutrition or diet coaches or other medical personnel whose focus lies on treating seniors who show restrictions in their (fine-) motor skills and/or suffer from chronic diseases. In general these professionals supply diverse supportive (home) care and/or medical care services to senior citizens. Informal caregivers in vAssist are defined as family members, relatives and/or friends who voluntarily take care of elderly people without any contract or payment. They provide supportive services ranging from housekeeping and grocery shopping to helping with sanitary care and being a social companion. Finally, tertiary users in vAssist are service providers which are defined as third party companies and institutions that are involved in the provision, operation and maintenance of socio-technological solutions (like the one planned by vAssist).

With this definition of who our users are we are now able to move on and start with the exploration of relevant user requirements; initially focusing on primary users.

Gathering Requirements

The first round of requirements gathering, which focused only on primary users of vAssist, consisted of a mix of qualitative and quantitative analysis methods. Our goal was to obtain a general understanding of the domain, for which we used the following questions as a guidance:

- What tele-communication and tele-medical devices and services are mainly used by senior adults?
- What types of voice-operated services may have potential to add value to the daily life of senior users; in particular to those who suffer from (fine-) motor skill restrictions and/or chronic diseases?
- What preferences do senior users have when it comes to various voice-control design characteristics?

In order to explore these questions in more detail we used focus groups as a qualitative data collection instrument. The method has previously been employed by a variety of HCI researchers and proven to be a valid approach for initial domain exploration [8] [6] [2]. Most authors agree that the main advantage of focus groups lies is their ability to use group interaction as a means to generate fruitful user requirements data. In addition we also sent out questionnaires which provided some quantitative feedback on the technologies and services that are commonly used by seniors.

Methodology

Two rounds of focus groups were organized both of which let primary users discuss devices, services, interaction and business requirements of future tele-communication and tele-medicine applications. Prior to the start of a focus group participants were provided with some basic information about the vAssist project and asked to sign a consent form. The research team leading the the focus groups consisted of the moderator and one other researcher. Both of them participated in the discussion, took notes, and collected information on flip-charts. A set of pre-defined questions was used to guide the discussions. In addition to a written protocol all sessions were audio/video recorded and the collected data was interpreted following a structured content analysis [7].

The Wizard of Oz (WOZ) simulation technique was used to better convey the idea of voice-control. WOZ allows for the demonstration of novel interaction scenarios without requiring a fully functional prototype. Instead a human 'wizard' simulates the functions of the potential future system [5]. In addition existing voice-control features were demonstrated, using a recent smartphone and a tablet computer. The goal of these demonstrations was to give participants a general impressions of how a future vAssist product might work. As part of this, users were also invited to participate in group work sessions where GUI and VUI interactions could be compared (e.g. writing a short email using touch interaction vs. writing it using voice-based interaction). Doing this we were able to engage 'inexperienced' participants in using modern technologies, providing us with some insight into what works and what does not work.

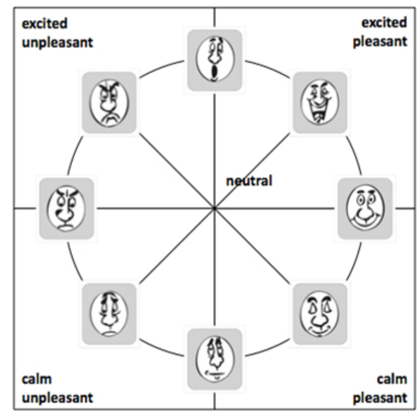

Figure 1: EmoCards use nine different representations to express the emotional state of interacting with a product.

In order to better structure this initial feedback we asked participants to choose EmoCards [3] as a way to express their emotional reactions to the two different interaction modalities. EmoCards help to determine the level of pleasantness and arousal of products and services by offering a selection of nine different cards, each of which representing a different emotional state (cf. Figure 1). A user is then asked to select the one state that best describes the interaction with the product. Finally, to expand upon the feedback coming from the EmoCards, we asked participants to also enumerate positive as well as negative aspects of using VUIs and GUIs for the different interaction scenarios. Following we briefly summarize and discuss some of the results we gained from this analysis.

Feedback and First Impressions

In Austria 10 potential primary users (4 male, 6 female; mean age: 70.20 years; SD: 3.19 years) and in France 8 (3 male, 5 female; mean age: 78.00; SD: 6.48) participated in the the first round of focus groups. All of them had at least some experience with modern Information and Communication Technology (ICT) and showed age-related restrictions in their (fine-) motor skills. We started with an exploration of technical devices and their physical location at people's homes and then followed with a discussion about the types of communication channels participants would use to interact with their family members and friends. Results show that mobile and land-line phones are preferred communication instruments for short distance, and PCs/laptops (e.g. Skype) for long-distance interactions with family and friends. Devices are usually situated in the living room, the kitchen or the the home office and often several of them are available (e.g. mobile phone, land-line and internet-enabled PC). In addition, participants highlighted the cost-effectiveness of text-based communication (i.e. email, text message) which (if necessary) can outweighs potential problems of controlling interaction devices (e.g. keyboard, mouse, touchscreen). In general primary users showed an affinity to text-based communication when using the PC/laptop. Text messaging on mobile phones is, however, only used if general input barriers can be overcome (i.e. mobile phone keyboards as well as font sizes are often too small to be operated). Spoken input was perceived as a good alternative to finger-operated input techniques.

The second focus group consisted of 8 potential primary users in Austria (4 male, 4 female; mean age: 70.22 years; SD: 8.84 years) and 8 of them (3 male, 5 female; mean age: 78.00; SD: 6.48) in France. Also here participants had some experience with modern ICT and showed relevant age-related restrictions. This time we started with a demonstration of different voice-controlled interaction scenarios using Apple's *Siri* application on a recent iPhone. Such was followed by a group work session where each participant had the chance to write an email using both the VUI and the GUI, respectively. The following discussion produced insights with respect to requirements of voice-controlled computing systems, and also highlighted some important characteristics of VUIs such as personalization and feedback mechanisms that should be taken into account when designing these types of technologies. From a GUI perspective the small size of the screen, the font and the keyboard were perceived as the main interaction barriers. The interaction with the VUI, however, was generally perceived as positive; in particular with respect to its input characteristics which do not require the operation of small controls.

Discussion

Even though these focus groups only provided us with some initial understanding about how people with age-related restriction may operate VUIs, they do already allow for the definition of some basic requirements for this distinct user group. From a basic hardware perspective we have learned that, in order to be able to physically interact with devices, they should at least have the size of a postcard. Also, while users are generally flexible with respect to the type of device they use (i.e. mobile phones, tablet computers and PC/laptops are accepted as interaction devices) they do require a specifically simple and clearly structured interface as well as an easy to understand written instruction manual. In addition, font sizes and manual input elements, if necessary, need to be adjustable to support users' physical restrictions.

Looking more specifically at the aspect of voice-control our explorations showed that our potential users generally enjoyed using this new way of interaction. They preferred a natural input language over commands although sentences tended to be short and rather precise. In cases where problems with the voice recognition or language understanding occurred users expected from the system to pro-actively provide a solution or offer them an alternative interaction channel. Also, our participants consistently preferred a female voice over a male one as they found such was easier to understand. Generally we found that the voice-based interaction increased the overall engagement with the device. This impression was supported by participants' demands of giving the system a name so that it better resembles the characteristics of a (virtual) human and their preference for a 'friend-like' interaction style. That is, most participants would prefer a personal form of communication (e.g. "Hi Paul") over a formal one (e.g. "Hello Mr. Smith").

Conclusion

In summary our first explorations of seniors interacting with (potential) VUIs indicate that voice can be seen as an efficient and rather engaging input modality; in particular if users suffer from age-related physical restrictions which limit their possibilities of interacting with physical input channels. Technical problems and insufficient component quality (e.g. low recognition rate) may however drastically reduce this positive experience and quickly lead to frustration. Hence, efficient error recovery strategies, sufficient feedback, and an additional fall-back modality that may be used in cases where voice-based interaction becomes too tedious or simply remains stuck, represent 'must-have' features that any application targeting this specific user group needs to offer. With respect to the applied design and research methodology the combination of focus groups with WOZ simulation, extended by some non-verbal emotional feedback (i.e. EmoCards) and quantitative questionnaire data, has proven to be an efficient approach for initial domain exploration.

References

[1] Bush, V. As we may think. The Atlantic, 1945.

[2] D., M. Focus groups. *Annual Review Sociology 22* (1996), 129–152.

[3] Desmet, P. M. A., Overbeeke, C. J., and Tax, S. J. E. T. Designing products with added emotional value: Development and application of an approach for research through design. *The Design Journal 4*, 1 (2001), 32–47.

[4] European Union. Active ageing and solidarity between generations. In *Eurostat, Statistical books* (2011).

[5] Kelley, J. F. An empirical methodology for writing user-friendly natural language computer applications. In *Proceedings of CHI* (1983), 193–196.

[6] Kitzinger, J. *Qualitative Research in Health Care*. P. M. J. Publising Group, 1996, ch. Introducing Focus Groups.

[7] Mayring, P. *Texte versthen: Konzepte, Methoden, Werkzeuge*. Universitäsverlag, 1994, ch. Qualitative Inhaltsanalyse.

[8] Merton, R. K., Fiske, M., and Kendall, P. L. *The Focused Interview: A Manual of Problems and Procedures*, second ed. Free Press, 1990.

[9] Woods, B., Spector, A. E., Prendergast, L., and Orrell, M. Cognitive stimulation to improve cognitive functioning in people with dementia. *Cochrane Database of Systematic Reviews 4* (2005).