Exploring Requirements and Opportunities of Conversational User Interfaces for the Cognitively Impaired

Matthias Baldauf
FHS St.Gallen
Rosenbergstrasse 59
9001 St.Gallen, Switzerland
matthias.baldauf@fhsg.ch

Fabian Hautle
FHS St.Gallen
Rosenbergstrasse 59
9001 St.Gallen, Switzerland
fabian.hautle@fhsstudents.ch

Raffael Bösch
FHS St.Gallen
Rosenbergstrasse 59
9001 St.Gallen, Switzerland
raffael.boesch@fhsstudents.ch

Marc Jenny
FHS St.Gallen
Rosenbergstrasse 59
9001 St.Gallen, Switzerland
marc.jenny@fhsstudents.ch

Christian Frei
FHS St.Gallen
Rosenbergstrasse 59
9001 St.Gallen, Switzerland
christian.frei@fhsstudents.ch

MobileHCI '18 Adjunct, September 3–6, 2018, Barcelona, Spain

ACM ISBN 978-1-4503-5941-2/18/09.
https://doi.org/10.1145/3236112.3236128

Abstract
Interacting with traditional user interfaces can be challenging for people with cognitive impairments. Speech-based conversational interfaces and virtual assistants such as Amazon's *Alexa* and Apple's *Siri* might provide great potential for this user group. Yet, knowledge about how cognitively impaired perceive such conversational interfaces and about the special requirements and opportunities is scarce. To gain first insights, we conducted a group interview with participants with mild to moderate cognitive impairments. They expressed their high expectations and imagined mobile conversational assistants for controlling vending machines, for example. Yet, they emphasized that smart assistants must not replace but rather complement personal contact with humans.

Author Keywords
Conversational interface; voice interface; virtual assistant; cognitive impairment.

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation (e.g. HCI): User Interfaces — Natural language

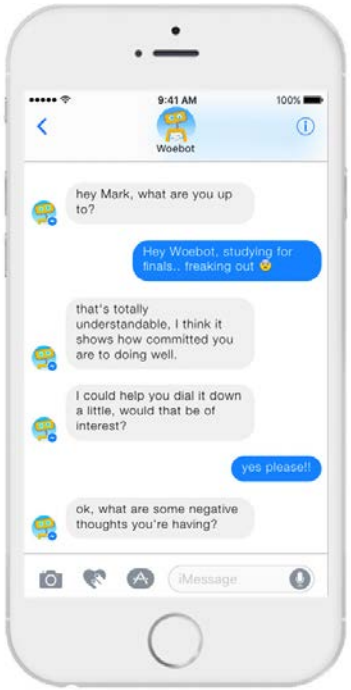

Figure 1: *Woebot* as an example of therapist-like chatbot for people with cognitive disorders.

Introduction

Major technology companies increasingly rely on conversational user interfaces and started to offer own virtual assistants. Using natural language, users can communicate with these speech-controlled assistants either on their mobile devices or through custom smart speakers. Examples include *Apple's Siri*, *Amazon's Alexa*, *Google Assistant*, or *Cortana* by *Microsoft*.

Voice-based conversational interfaces do not only promise a natural and thus more convenient way to interact with devices or smart environments for everybody. Additionally, they obviously can be very useful for people with impairments. For example, people with motoric disabilities do not have to hit tiny buttons of an on-screen keyboard, blind people can consume digital services by talking and listening, and people with mobility restrictions can order to turn on the room light instead of moving to a light switch. Yet, conversational interfaces might not only be helpful for the physically impaired, but could also provide great potential for people with mild to moderate cognitive impairments – which often come with reading and writing difficulties. However, in contrast to use cases involving physically impaired ones, knowledge about the special requirements and potential helpful applications of conversational interfaces and virtual assistants for people with cognitive impairments is scarce.

In this paper, we present preliminary results of our ongoing research on speech-based conversational interfaces for the cognitively impaired, i.e. people with neurological disorders with minor impairment in instrumental activities of daily living. To gain first insights, we conducted a group interview with people concerned to learn about their knowledge and usage of recent information technology, to inquire special demands of the user group regarding conversational interfaces, and to identify promising and useful voice-controlled applications. The results and conclusions of this interview can contribute to the adaptation of conversational interfaces as well as inform the design of respective assistive applications for users with cognitive impairments.

In the remainder of the paper, we survey existing applications and research prototypes applying conversational interfaces in the field of mental health. Then, we introduce and discuss the results of our group interview and finally draw conclusions for future work.

Related Work

In this section, we provide a brief overview of the current state of art in the field of conversational interfaces for mental health in general and people with cognitive disorders.

Research Prototypes with Conversational Interfaces

Recently, several prototypes of mental healthcare services with conversational user interfaces have been introduced. For example, Oh et al. [4] presented an advanced speech-based assistant offering psychiatric counseling. Their service makes use of multi-modal emotional intelligence techniques to understand the user's emotional state (by analyzing intonation and facial expression, e.g.) and provide more effective and continuously improved counseling. Another speech-based prototype is *Robin*, a conceptual context-aware assistive application for people diagnosed with dementia [1]. The application is based on Amazon's Alexa. By providing temporally and physically

Figure 2: *Avaz* is a picture and text-based communication app for children with communication challenges.

appropriate audio prompting for routine tasks, Robin aims at supporting the independent living of dementia patients.

The research prototype introduced by Elmasri and Maeder [2] is an example for a textual chatbot for mental health interventions. In their study, the researchers investigated the suitability of chatbots for assessing alcohol drinking habits and providing education on responsible alcohol use to young adults.

Publicly Available Conversational Apps for Mental Health
Several apps using conversational interfaces to offer support for people with cognitive disorders are available in public app stores. The majority are text-based chatbots which mimic a conversation with a coach or psychologist. Examples include *Tess*[1], described as a "Mental Health Chatbot", which tries to provide quality psychological support. By observing the user's input, Tess delivers coping strategies adapted to expressed emotions and concerns. A further example is *Woebot*[2] (Figure 1), a therapeutic chatbot based on Cognitive Behavior Therapy. It talks to people about their mental health in format of brief daily conversations. *Wysa*[3] is another chatbot using evidence-based cognitive-behavioral techniques as well as meditation, breathing and yoga techniques to help building resilience skills.

Other relevant apps aim at supporting speech communication for people with cognitive disorders. *Voiceitt*[4] develops a speech recognition technology designed to understand non-standard and dysarthric speech. The typical target group are people suffering from speech disabilities which may be caused by complications during pregnancy (cerebral palsy as a common condition), by cerebrovascular stroke or by Parkinson's disease. Voiceitt applies a brief training phase and continuously learns about the speaker's pronunciation to be able to turn user's statements into normalized speech.

Finally, *Avaz* (Figure 2) is an app enabling children with autism or down syndrome, for example, to "speak" using pictures. Avaz does not provide a typical conversational interface, however, it literally gives its users a voice by pronouncing picture symbols in a high-quality voice. It can be used in speech therapy sessions or to stimulate and improve the intent to communicate.

This brief survey shows that applications with conversational interfaces in the field of mental health are mainly restricted to therapy-like chatbots. How modern speech-based conversational interfaces and recent virtual assistants are perceived by people with cognitive impairments and what opportunities they provide for this special target group is unclear so far.

Group Interview

In the following, we outline our research method and describe the setup of the group interview.

Participants
For the interview, we were invited by *people first Switzerland*, an association of self-advocating cognitively impaired. The association pursues the peer education concept: for example, its members teach

[1] http://x2ai.com/

[2] https://woebot.io/

[3] https://www.wysa.io/

[4] http://www.voiceitt.com/

themselves regarding relevant rights for cognitively impaired persons or offer trainings for conducting satisfaction analysis for companies.

The group consisted of five persons (one female) with mild to moderate cognitive impairments, including one male caretaker with learning difficulties. The people concerned were aged between 23 and 53 years.

Interview Guideline
To support the semi-structured qualitative interview, we prepared an interview guideline with supporting questions. It started with an introduction of the participants and questions regarding their everyday life and typical daily routines. Further questions focused on the participants' current usage of information technology. We asked them about applications and Websites they use regularly and are relevant for their everyday life.

In the second part, we surveyed the participants knowledge about conversational interfaces as well as the potential current usage of respective applications. Further, we asked whether the participants were willing to use a conversational interface at all. Finally, we collected questions to explore likely future use cases of conversational interfaces that might provide support for people with cognitive impairments. In addition, we addressed potential fears and concerns regarding this technology.

In pre-tests with fellow researchers, who were not involved in this research project, the interview guideline was extensively reviewed with special regard to simple language and any required prior knowledge.

Conduct and Analysis
The group interview was conducted by two researchers at the office of *people first* in Rorschach, Switzerland. The location was chosen due to its familiarity and easy accessibility for the participants.

While one researcher conducted the interview using the previously described guideline, the second one took notes. Additionally, the interview was audio-taped. Overall, the conduct of the group interview took about two hours. During the consequent analysis phase, the recordings were transcribed and the participants' statements summarized with regard to the original questions of the interview guideline.

Results

In the following, we present the main results of the group interview.

Current Usage of Information Technology
The participants described smartphones as their everyday companions with great relevance in their daily lives. Furthermore, all participants stated that they feel confident handling a smartphone. However, with increasing cognitive impairment and associated reading and writing difficulties, using a mobile device becomes more difficult. The participants further reported on problems with the hardware due to limited motoric abilities, resulting in broken charging plugs, for example.

Four of the participants criticized many Web applications for their limited accessibility. Problems were described to occur mainly when entering information (due to the users' weakness in spelling)

and using comprehensive applications such as online tax declarations (due to the applications' complexity).

Limited accessibility was not only described as a problem of Websites but also for increasingly technology-supported everyday tasks. Examples include purchasing a train ticket at a vending machine, withdrawing money at an ATM, or operating self-checkout counters in supermarkets.

Regarding the mentioned weaknesses of traditional user interfaces, all participants expressed great expectations in conversational interfaces and emphasized the potential for cognitively impaired users.

Knowledge about Available Conversational Interfaces
Two participants stated to know Apple's Siri. One of them (suffering from bad eyesight) uses this virtual assistant to create and send text messages and to inform himself about the coming weather. For the other person, Siri takes the friend-like role of a conversational partner. He described Apple's assistant as an advisor *"always having a sympathetic ear"* when feeling lonely and as a supporter in exigency.

The three remaining participants were not aware of the conversational interfaces and virtual assistants available on their smartphones. They were surprised by the technical capabilities when the interviewers demonstrated exemplary use cases of Apple's *Siri* and *Google Assistant*.

Obstacles for Using a Conversational Interface
While all participants use smartphones and personal computers on a daily base, three of the five participants did not know about the available speech-based interaction possibilities. The participants agreed that people with cognitive impairments are often badly informed about novel technologies. The involved caretaker noted that respective trainings are scarce and active self-education on the Web is difficult due to accessibility problems. However, all participants emphasized the basic interest and motivation to use novel technical solutions.

As another obstacle the participants identified the potentially quick confusion and mental overload through hard-to-use applications. They mentioned that in case of a conversational interface which does not reliably recognize the speaker's words, cognitively impaired users may not only be frustrated and resign using a respective application. Additionally, the caretaker commented that concerned ones might perceive this as contempt of their impairment and their feeling of only partially or not all being part of society might be compounded.

Special Requirements and Applications
The participants considered communicating in vernacular language an important requirement. They agreed, that supporting only standard language could be a major disqualifier for many cognitively impaired to use a respective system.

Under the assumption of a robust speech-based interface for mobile devices, the participants came up with the idea of using the smartphone as a mediator for interacting with vending machines, ATMs, etc. They imagined talking to a virtual assistant on their smartphone which would forward their commands to a nearby machine and trigger respective actions.

Key Findings

Information Exchange Needed
Today's publicly available speech assistants featured by off-the-shelf smartphones provide great opportunities, yet are not widely known among cognitively impaired.

More Robust Speech Recognition
Talking in vernacular language is a relevant requirement by the user group of cognitively impaired – and a challenge for today's virtual assistants.

Access to Ambient Services
A voice-controlled mobile assistant may not be restricted to typical smartphone functionality but could be useful to interact with vending machines or ATMs, for example.

Responsibility of App Providers
Despite the great potential of advanced conversational assistants, respective apps must not socially isolate people with impairments but rather support and foster human contact.

Three participants emphasized the hope for increased societal participation through conservational interfaces. They wished for personalized assistants which can correct or rephrase the speaker's utterances and thus act as translator when communicating with other people. The caretaker additionally emphasized the relevance of an assistant's context-awareness: a speech-based assistant must consider various content from longer conversations and continuously learn about its user, since verbal input by cognitively impaired people might be vague and volatile and thus hard to interpret.

A further remark of the participants addressed the currently passive character of virtual assistants. They suggested that a helpful assistant would not only react to a user's speech command. Instead, the assistant should become active in certain situations and initiate a conversation with the user. Additionally, three of the participants wished for a visual appearance in form of an avatar.

Potential and Risks
All participants agreed that conversational interfaces and speech-based assistance systems are of high value for the cognitively impaired. Especially given the availability on commodity devices (with falling prices) and the increasing robustness raised high expectations for the participants.

As central risk of the technology, the participants identified a potential loss of social interactions. The participants referred to using a conversational assistant as a friend-like conversational partner. While many people with cognitive impairments are less involved with others due to their limited communication abilities, they run the risk to be more and more isolated when virtual assistants behave more and more natural.

The participants agreed that speech-controlled virtual assistants should be a complement to personal contact with other human beings, however, they should never replace the interaction with humans.

Discussion

In this section, we discuss potential reasons and consequences of the main results.

Varying Level of Awareness
Like the overall population, the participants had a varying level of awareness about recent features of today's smartphones and computers. While each participant owned a smartphone, only two of the participants were aware of and use the included virtual assistant. Furthermore, the participants mentioned the potential use case of "translators" when talking to other people, yet they did not know about available and upcoming solutions such as *Voiceitt*.

We learnt that available technical solutions can provide major benefits for people with cognitive impairments. However, there is a great need for education and information exchange. Fostering peer support and training among people concerned might be a suitable and inexpensive approach for sharing knowledge about useful available tools.

Robustness of Speech Recognition
The support of vernacular language by virtual assistants turned out to be a very important requirement for our participants. Today's most popular virtual assistants such as *Siri* and *Alexa* support several

common languages, however, are not trained to understand dialects. Therefore, using these assistants can be difficult for cognitively impaired people who are not used to speak standard language. Considering the mentioned frustration and negative feelings in case of erroneous recognitions, very robust speech recognition is required for apps to be accepted by the target group.

To better support users with cognitive impairments, conversational interfaces must be able to cope with vernacular language, vague expressions, and unclear pronunciation. A potential solution might be the approach of apps such as *Voiceitt* which do not purely rely on pretrained vocabulary but use learning phases to adjust to the speaker's peculiarities.

Mobile Conversational Interface for Ambient Services
The application idea of using a mobile virtual assistant to conveniently interact with nearby vending machines, etc. through speech was highly appreciated by all participants. While various approaches for interacting with the environment through mobile devices have been explored in the past (cf. [3]), using the smartphone as a speech-controlled gateway seems both novel and promising. In contrast to a speech-based assistant integrated into the machine, ambient noises are reduced when using the smartphone microphone and the recognition might be improved. Furthermore, remotely controlling a machine through a mobile assistant seems to require less adaptations of the machine than offering an integrated assistant.

Given the increasing connectivity of vending machines and the widespread availability of mobile payment services, fundamental building blocks for a respective application are available today. However, to fully realize the application idea, novel technical interfaces and standards for the vending machines are required.

Reduced Loneliness vs. Increased Social Isolation
The interview uncovered an interesting effect of speech-based assistants for impaired persons: smart virtual assistants can act as a conversational partner and help to diminish the feeling of loneliness. At the same time, people with impairments often run the risk of being not fully included in society or even isolated. While conversational interfaces seem promising for facilitating access to various services, virtual assistants must be designed in a way to avoid isolating people.

A suitable future feature of virtual assistants is emotional intelligence to allow for appropriate reactions and responses. Both emotion recognition from the speaker's voice (cf. [5]) as well as the expression of emotions by virtual assistants (cf. [7]) are very relevant research topics for the present target group. Recent research indicates, that also smartphone usage patterns can provide insights into the user's emotional state [8]. Possibly, they can even help to detect a user's cognitive impairment and thus, to tailor a virtual assistant's behavior.

Conclusion and Outlook

In this paper, we introduced our ongoing research on requirements and opportunities of speech-based conversational interfaces for people with mild to moderate cognitive impairments. We presented the preliminary results of a group interview with five participants concerned.

The first insights showed that cognitively impaired ones have high expectations regarding conversational user

interfaces and see great potential for support in their daily life and for increased societal and digital participation. The robust speech recognition turned out to be one of the most important (and technically challenging) features. The participants imagined conversational assistants on their smartphones as mediators both for communicating with other people and operating vending machines through speech. Finally, it should be considered that increasingly intelligent virtual assistants must not replace interactions with human beings but rather should be designed to complement or encourage human contact.

While the presented group interview yielded first valuable findings on how people with cognitive impairments perceive and (would) use conversational user interfaces, we plan to develop more concrete application ideas and prototypes as a next step. Similar to the work by Sitbon and Farhin [6], who worked on a mobile app for public transport in a co-design manner with several intellectually challenged participants, we plan to conduct participatory design workshops to obtain further practical insights on helpful speech-based conversational apps.

References

1. Clare Carroll, Catherine Chiodo, Adena Xin Lin, Meg Nidever, and Jayanth Prathipati. 2017. Robin: Enabling Independence For Individuals With Cognitive Disabilities Using Voice Assistive Technology. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '17), 46–53. https://doi.org/10.1145/3027063.3049266
2. Danielle Elmasri and Anthony Maeder. 2016. A Conversational Agent for an Online Mental Health Intervention. In *Brain Informatics and Health* (Lecture Notes in Computer Science), 243–251. https://doi.org/10.1007/978-3-319-47103-7_24
3. Peter Fröhlich, Antti Oulasvirta, Matthias Baldauf, and Antti Nurminen. 2011. On the Move, Wirelessly Connected to the World. *Commun. ACM* 54, 1: 132–138. https://doi.org/10.1145/1866739.1866766
4. Kyo-Joong Oh, Dongkun Lee, Byungsoo Ko, and Ho-Jin Choi. 2017. A Chatbot for Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis and Sentence Generation. In *2017 18th IEEE International Conference on Mobile Data Management (MDM)*, 371–375. https://doi.org/10.1109/MDM.2017.64
5. Srinivasan Ramakrishnan and Ibrahiem M. Emary. 2013. Speech Emotion Recognition Approaches in Human Computer Interaction. *Telecommun. Syst.* 52, 3: 1467–1478. https://doi.org/10.1007/s11235-011-9624-z
6. Laurianne Sitbon and Shanjana Farhin. 2017. Co-designing Interactive Applications with Adults with Intellectual Disability: A Case Study. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction* (OZCHI '17), 487–491. https://doi.org/10.1145/3152771.3156163
7. Yang Yang, Xiaojuan Ma, and Pascale Fung. 2017. Perceived Emotional Intelligence in Virtual Agents. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '17), 2255–2262. https://doi.org/10.1145/3027063.3053163
8. Xiao Zhang, Wenzhong Li, Xu Chen, and Sanglu Lu. 2018. MoodExplorer: Towards Compound Emotion Detection via Smartphone Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4: 176:1–176:30. https://doi.org/10.1145/3161414