Designing Accessible Conversational Interfaces for Older Adults: The Case for New Usability Guidelines

Christine Murad

University of Toronto, TAGlab
Toronto, ON, Canada
cmurad@taglab.ca

Cosmin Munteanu

University of Toronto, TAGlab and
University of Toronto Mississauga,
ICCIT
Toronto, ON, Canada
cosmin.munteanu@utoronto.ca

Abstract

Conversational interfaces are becoming more prevalent and are appearing more in the commercial market, such as Siri, Google Home, Amazon Echo, etc. These interfaces have great potential to improve the experience of interacting with technology through using one's voice. These interfaces can be very useful for older adults in particular, as they can help address digital accessibility barriers such as loss of vision, mobility impairments, and cognitive impairments. However, many usability issues still currently exist in these conversational interfaces, both in learnability and usability. In order to help solve these usability issues, our work explores the development of design guidelines and heuristics that will help improve the design of voice interfaces for older adults, of which currently none exist.

Author Keywords

Speech interaction; Design Guidelines

ACM Classification Keywords

H.5.2. Information interfaces and presentation: User Interfaces;

Introduction

There has been an emergence of speech interfaces that have been introduced into the commercial market, especially personal speech assistants such as Siri, Google Home, Amazon Echo, etc. Many of these interfaces allow users to interact with technology using one's voice, without a graphical interface. It's been shown that people have a natural preference for using speech to interact with technology, since we use speech to interact in our everyday lives [12]. These interfaces can also be very beneficial for people who are limited by physical accessibility barriers, such as visual impairments and mobility impairments [12]. There is great potential for the adoption of these commercial speech interfaces.

Older adults are especially able to benefit from these speech interfaces, as they face many of these physical accessibility barriers as they age [12]. Research has been done exploring older adults' attitudes about speech interfaces. Older adults have been shown to be very receptive to conversational interfaces. They've been viewed as a natural, simplistic way to be able to interact with technology [2,3]. Bickmore [2] performed a study to explore how older adults used and accepted relational agents. He built a relational agent called FitTrack that acted as an exercise advisor for older adults. He found that older adults were very satisfied with the agent, and recognized that the agent was beneficial and useful for them. However, as mentioned earlier, older adults will only be comfortable to adopt these technologies if they see a benefit in them, versus the cost of adopting a new, unfamiliar technology into their life [2,6]. Therefore, any barriers to usability and adoption can make these interfaces not worth investing time into.

However, current speech interfaces contain barriers to adoption by older adults. It is difficult for someone to learn and remember how to use a speech interface, especially if the interface does not contain a visual display [4,5]. Much cognitive effort is required to use a speech interface. Based on the technology acceptance model, older adults are more likely to adopt new technology if it is easy to use and it is useful for them [2,6]. Therefore, these barriers risk digitally marginalizing older adults from using and adopting speech interfaces.

Design Heuristics for Voice Interfaces for Older Adults

There are many established user interface guidelines and techniques for designing interaction for graphical interfaces. Some of the most notable and established sets of design heuristics for graphical user interfaces are the ones created by Nielsen [9] and Shneiderman [11]. They both outline guidelines that are necessary to ensure user interfaces are usable and easy for users to interact with. For example, Nielsen [9] states that user interfaces should be designed so that, rather than having to recall previous information, a user should be able to recognize previous information through the interface at any point in time. He also outlines that user interfaces should allow users to easily recognize and recover from errors. Shneiderman [11] highlights that user interfaces should support an "internal locus of control" – that is, users should feel that they have control of the system and the actions initiated during interaction.

However, there is a lack of equivalent principles for the design of speech interfaces. Attempts to directly map interaction techniques from the domain of graphical

Figure 1: An older user attempting to interact with Google Home (screengrab from https://www.youtube.com/watch?v=e2R0NSKtVA0, copyright Ben Actis, retrieved on January 26, 2018).

interfaces to voice-based interfaces were not successful. Sherwani [10] developed a speech interface called VoicePedia, that allows people to search through Wikipedia using a command-based voice interface that was meant to mimic the graphical version of the same system. However, they found that it is difficult and not user-friendly to directly map the interaction of graphical interfaces to speech interfaces. Yankelovich [13] also states that GUI interaction does not map well to speech interfaces . Therefore, more work has to be done in order to develop guidelines for speech interfaces, especially taking into consideration the needs of older adults in these interfaces.

Moving Forward

As it stands, we have been designing these smart speaker devices without principles or guidelines to direct the design of the conversational voice interfaces embedded inside them. They have been advertised as a natural way to interact with technology, as speech is a natural form of interaction [1,7,8]. However, without any guidance or understanding on how these smart devices are meant to be interacted with, these devices become unusable. This is visible in recent examples of digitally marginalized users such as older adults trying to interact with such smart objects (Figure 1) – a category of users which is often touted as the ones who could benefit the most from such devices.

The development of design heuristics for conversational voice interfaces is an important step for addressing current usability issues and improving the interaction of smart speech devices. The goal of our work, therefore, is to develop guidelines that can guide designers in making conversational voice interfaces more user-friendly for older adults.

Author's Biographies

Christine Murad is a graduate student at the Technologies for Aging Gracefully lab in the Department of Computer Science at the University of Toronto. Her research revolves around the usability and design of conversational voice interfaces, and the exploration and development of design heuristics to aid in creating intuitive and user-friendly conversational voice interactions. She completed her undergraduate studies in Computer Science at the University of Toronto (2017).

Cosmin Munteanu is an Assistant Professor at the Institute for Communication, Culture, Information, and Technology at University of Toronto Mississauga, and Co-Director of the Technologies for Ageing Gracefully lab at University of Toronto. His area of expertise is at the intersection of Human-Computer Interaction, Automatic Speech Recognition, Natural User Interfaces, Mobile Computing, Ethics, and Assistive Technologies. He has extensively studied the human factors of using imperfect speech recognition systems, and for the past two decades has designed and evaluated systems that improve humans' access to and interaction with information-rich media and technologies through natural language. Cosmin's multidisciplinary interests include speech and natural language interaction for mobile devices, mixed reality systems, learning technologies for marginalized users, usable privacy and cyber-safety, assistive technologies for older adults, and ethics in human-computer interaction research. URL: http://cosmin.taglab.ca

References

1. Matthew P. Aylett, Per Ola Kristensson, Steve Whittaker, and Yolanda Vazquez-Alvarez. 2014. None of a CHInd. *CHI EA '14*: 749–760.
2. Timothy W. Bickmore, Lisa Caruso, and Kerri Clough-Gorr. 2005. Acceptance and usability of a relational agent interface by urban older adults. *CHI '05*: 1212.
3. Timothy W. Bickmore, Lisa Caruso, Kerri Clough-Gorr, and Tim Heeren. 2005. "It's just like you talk to a friend" relational agents for older adults. *Interacting with Computers* 17, 6: 711–735.
4. Eric Corbett and Astrid Weber. 2016. What can I say? Addressing user experience challenges of a mobile voice user interface for accessibility. *MobileHCI '16*: 72–82.
5. Benjamin R Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. 2017. "What can i help you with?": Infrequent Users' Experiences of Intelligent Personal Assistants. *MobileHCI '17*: 1–12.
6. Marcel Heerink, Ben Kröse, Vanessa Evers, and Bob Wielinga. 2010. Assessing acceptance of assistive social agent technology by older adults: The almere model. *International Journal of Social Robotics* 2, 4: 361–375.
7. Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents. *CHI '16*: 5286–5297.
8. Cosmin Munteanu, Ben Cowan, Keisuke Nakamura, Pourang Irani, Sharon Oviatt, Matthew Aylett, Gerald Penn, Shimei Pan, Nikhil Sharma, Frank Rudzicz, and Randy Gomez. 2017. Designing Speech, Acoustic and Multimodal Interactions. *CHI EA '17*: 601–608.
9. Jakob Nielsen. 1994. Enhancing the explanatory power of usability heuristics. *CHI '94*: 152–158.
10. J Sherwani, Dong Yu, and Tim Paek. 2007. Voicepedia: towards speech-based access to unstructured information. *Interspeech*: 2–5.
11. Ben Shneiderman and Catherine Plaisant. 2010. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*.
12. Anjeli Singh, Andrea Johnson, Hanan Alnizami, and Juan E Gilbert. 2011. The Potential Benefits of Multi_Modal Social Interaction on the Web for Senior Users. *J. Comput. Sci. Coll.* 27, 2: 135–141.
13. Nicole Yankelovich, Gina-Anne Levow, and Matt Marx. 1995. Designing SpeechActs. *CHI '95*: 369–376.