Factors Affecting Seniors' Perceptions of Voice-enabled User Interfaces

Randall Ziman
Digital Whimsy Lab
University of Baltimore
Baltimore, Maryland 21201, USA
randall.ziman@ubalt.edu

Greg Walsh
Digital Whimsy Lab
University of Baltimore
Baltimore, Maryland 21201, USA
gwalsh@ubalt.edu

Abstract
This paper explores how seniors perceive Voice Enabled User Interfaces (VUIs) and the factors that shape those perceptions. An experiment was administered to 15 seniors (over age of 65), in which the participants searched for information using a traditional keyboard/mouse interface and an experimental voice/touch interface. An analysis of the data collected showed that seniors perceive meaningful differences between the two interfaces in terms of learnability, usability, ease of understanding and helpfulness.

Author Keywords
Seniors; older adults; voice user interfaces; search

ACM Classification Keywords
H.5.2 [Information interfaces and presentation (e.g., HCI)]: User Interfaces

Introduction
We are currently experiencing a dramatic aging of the population. In the United States, life-expectancy has improved from 47.3 in 1900 to 78.9 in 2015 [1]. For some groups, like white women, the changes are even more dramatic — increasing from 48.7 in 1900 to 81.3 in 2015. As people are living longer and the number of seniors (people over the age of 65) is growing rapidly. Data [5] shows that in 2016, there were 49.2 million people over the age of 65 and that

Gender	Age Range
F	65-69
F	75-79
F	75-79
F	75-79
M	80-85
F	80-85
M	75-79
F	70-74
F	80-85
M	80-85
M	86 or older
F	86 or older
M	80-85
F	86 or older
M	80-85

Table 1: List of participants.

they comprised 15.2% of the population By 2060, the number of people is expected to grow to 98.2 million, and comprise nearly 25% of the population [5].

Among the technologies available, Voice Enabled User Interfaces, VUIs, may hold potential for increasing usability for seniors. Many voice systems are efficient, intuitive and do not require the fine motor skills that older users can find challenging. However, research in this field is still limited and more work is required to identify the forces that shape user's perceptions of VUIs [3]. The goal of this paper is to gain a deeper understanding of how seniors interact with VUIs by identifying the major factors that affect user perceptions of VUIs.

Background

Multiple authors have recognized significance differences between individual seniors in attitude and aptitude regarding technology [7, 8, 2]. Although trends in the use of technology by seniors can be identified, research shows that after accounting for other characteristics, age alone is not an ideal predictor of technology use after accounting for other characteristics [6]. To illustrate the breadth of senior experiences, Morrell [10] points out that a 66-year-old baby boomer who has been working in IT since the 1990s may experience technology in a dramatically different way than an 85-year-old, blue collar worker who has been retired for 20 years.

Methods

To investigate the factors affecting seniors' perceptions of voice enabled user interfaces (VUIs), an experiment was conducted that compared seniors' experiences using a multimodal voice and touch interface (VTI) to a traditional keyboard/mouse interface (KMI). Fifteen seniors (over the age of 65) were invited to search for information online using each system while being observed by an investigator. At multiple points through the testing session, the participants were interviewed about their experience. The observations and interview responses from each session were analyzed to identify trends in how seniors interact with VUIs; the differences between VTIs and KMIs; and the major factors that affect user perceptions of each system.

Design & Participants

The core of the experiment consisted of instructing participants to search for information using each interface and then discuss their experiences with the interface. To control for system order, eight of the participants used the KMI first and seven participants used the VTI first. To control for question order and combinations, the questions and question order were varied for each participant.

Fifteen seniors were recruited to participate in the study, 3 from the Denver area and 12 from the South San Francisco Bay Area. Nine participants were female and six were male. All participants were over the age of 65 and volunteered to participate without compensation. See **Table 1**.

When asked to describe their technological expertise and experience on a background questionnaire, the participants varied between 1-4 hours of computer and/or tablet use. Participants also tended to identify as 'intermediate' users. However, in terms of both experience and efficacy there was representation at both ends of the spectrums.

Expanding from more generalized experiences with technology, participants were also asked to indicate their experience with and use of voice activated technology. Eleven out of 15 users indicated that they used a voice activated technology such as voice-to-text, Siri, Amazon Echo, Google Home, etc. Of the participants who used voice activated technologies, most participants started to use them 6 months

Scenarios

Factual
Find out which state had the highest voter turnout for the 2016 presidential election. Save the document(s) where you have found the information required.

Interpretive
Find out about the advantages and disadvantages of each type of school (charter, public, private). Save the document(s) where you have found the information required.

Exploratory
Put together two thirty-minute low-impact exercise programs that she could alternate between during the week. Save the document(s) where you have found the information required.

to 3 years ago, but have only used them a few times a week or a few times a month. Only one user indicated using voice activated technologies every day.

Search Tasks
Each search task was comprised of a short scenario and a related task that required the user to search for information. The structure of these tasks was based on [3, 4, 11, 12]. Previous work suggests search tasks are an ideal activity in this regard because the activity is highly familiar (the second most popular online activity for users of all generations), and the specifics of the task can be easily varied.

Utilizing [9], search tasks fell into three categories: factual tasks, interpretive tasks and exploratory tasks. Factual tasks require the user to find specific, concrete information that is not open to interpretation. Interpretive tasks require the user to make inferences and evaluations in the pursuit of information, while still looking for a specific answer or answers. Exploratory tasks are the most open-ended of the three and require the searcher to make significant judgments about whether information is helpful and correct or not.

For the purpose of this study, a total of twelve search tasks were used. Three of the questions came directly from the examples listed [3]. (See **Scenarios Sidebar** for examples from each category of search task).

Each testing session included five key touch points: setup, introduction, search system 1, search system 2, and an exit interview.

Technology
For both the KMI and VTI search system, all searches were conducted using the Google search engine via the chrome browser.

For the KMI search tasks, participants were asked to use a laptop with an external mouse. By using a new chrome account for each participant, the search history from previous participants did not affect later experiences.

Making use of the KMI interface required participants to place their cursor appropriately in a search field, type in search queries and press the enter button (or click the search button) to run the search. Users conducted iterative searches by clicking into a search field, editing the query and re-running the search. When participants found information that they deemed valuable, they were instructed to bookmark valuable information by using the mouse to click the star icon in the upper, right-hand of the address bar.

The VTI interface prototype utilized an Apple iPad and required participants to touch a microphone icon with their finger. Once users touched the microphone icon, the system chimed and the screen a voice input screen took over the interface. The voice input screen included instructions to “Speak now” until the user started talking. After a brief pause, the “Speak now” instruction would update to “Listening?”. Once the system detected sound (i.e. voice input), the screen would display the user’s voice input as it was being recorded. As the user continued to talk, the recorded input continued to update (the user’s additional input provides more context for the system to recognize the word most likely intended by the user; e.g.: pair and pear may sound similar, but pear is a better fit in the search query “when does pear/pair season start”). When the system detects a break in voice input, the search query (as recorded at that time) is run and the user is directed to a search engine results page.

Analysis
After all of the interviews were conducted, the audio recordings of the exit interviews were transcribed and a qualitative

Code	Factor
FMLK	Familiarity
USBT	Usability—Speech
USBK	Usability—Keyboard
HBTK	Habit—Keyboard
TRSK	Trust—Spoken Query
NVFT	Novelty Factor
SPDK	Speed
CMFK	Comfort Level
FNFT	Fun Factor

Table 2: Begany's [3] coding scheme for factors influencing users' perception.

content analysis was conducted. All transcripts were coded using an open coding approach to apply the codes established by [3]. The codes were applied to the entirety of the exit interview transcript. See **Table 2** for a list of codes and short description.

Results & Discussion

A comparison of post-system interview data shows that participants tended to generally look at both search systems favorablely. By quantifying the participant's responses on a five-point scale (1—very difficult, 2—difficult, 3—neutral, 4—easy, 5—very easy), the mean response was above 3 (neutral) for each question.

Some measures returned more pronounced differences between the two systems. When asked, "how easy/difficult did you find learning to use the [system]?", participants reviewing the voice system reported more favorable response. Although drawing on quotes from the exit interviews, the higher number of "neutral" responses regarding the keyboard system may be due to the fact that participants learned to use the keyboard system prior to the testing session (whereas the voice system truly did require learning on their part).

The second post-system interview question produced the greatest difference between the two systems. When participants were asked "how easy/difficult did you find using the [system] to search for information?", twelve participants described the voice system as "easy" or "very easy". On the other hand, only eight participants described the keyboard system as "easy" or "very easy".

When asked to compare the keyboard and voice systems in the exit interviews, eight participants preferred the keyboard system overall compared with only five who preferred the voice system (two did not have a preference). Interestingly, the participants overall judgements do not seem to be an aggregation of the sub-factors relating to their experience. For example, participants reported that the voice system was easier to learn (by a margin of six) and easier to use (by a margin of five) while still preferring the keyboard system overall.

The second most commonly identified factor was the usability of the voice system. This is intriguing because it contrasts with the overall preference for the keyboard system. Also, in the participants' discussion of the voice usability they also identified multiple potential usability problems for the keyboard system. For example, one participant said, "I am not a typist", presenting an interesting expression of generational perceptions about the role of typing. Other users mentioned the difficulties that they currently have with typing, some due to injury, some due to lack of experience. One user even discussed that he expected to lose his ability to type due to age related changes, at which point he would only be able to use a voice based system. Furthermore, multiple users pointed out that the voice based system removed the potential for spelling errors and typos, which seemed to be of significant concern to the seniors interviewed.

On the other hand, multiple users highlighted the advantages of the keyboards usability. The keyboard system was described as better for users to take their time entering a query and modify that query. In the words of one user, "when you're using the voice system, you really have to organize your thoughts before you commit to asking for information... not just organize your thoughts but organize them in such a way that you can verbalize them". This was particularly problematic for that user as she said "I think through my fingers.

As noted earlier, the habit of keyboard use stood out as a notable factor affecting users' perceptions. When users did explicitly mention their habits of using the keyboard system, they made it clear that they used these interfaces regularly and for a long time. For example, one user points out that she has been using a KMI regularly for "30 years".

The results of this study show that VUIs are a viable option for seniors. Despite the fact that seniors are often regarded as opposed to adopting new technologies, these findings build on the body of work showing that seniors can and will use voice based interactions. These findings seem to be particularly significant for working with novice users who are more likely than intermediate and advanced users to prefer a voice based interface. This implies that particularly when designers are working on untraditional or innovative interfaces for seniors, voice interactions should be considered as a viable option for lowering the perceived ease of use.

Every participant interviewed experienced at least one time out error. For the most part, these errors occurred when the users where pausing mid-thought while trying to construct their search query. This error was exacerbated by the fact that participants frequently looked up and/or away from the device when thinking, thereby missing any visual cues that an error had occurred. By the time the user looked back at the interface, either a search had been executed with only a segment of the users intended query; or the input had failed and none of the speech was input. Multiple participants expressed frustration at this point, although interestingly, participants did not seem to express as much frustration when struggling with the voice input as they did when struggling with the keyboard input (e.g. typos).

The primary usability challenge for seniors using the VTI was time-outs. Either the participants were not able to fully input their request; or they accidentally inserted unintended words into their query. Seniors may benefit significantly from a setting that prevents searches from automatically triggering after a pause/timeout. This interaction pattern is already prevalent voice interactions for sending text messages. For example, when sending a text message through Google Assistant, the interface will ask the user to input the message and then ask the user to confirm or edit before sending the message.

One of the key differences between the seniors' behavior when using the KMIs compared to the VTIs was their reliance on suggested search terms. When entering text into a Google search field, the interface automatically displays suggestions to complete the query or related searches. .

Unfortunately, once a user initiates a voice search by clicking the microphone icon, the full screen is taken over by a voice input modal. This modal is completely dedicated to the voice input and does not display suggested searches. To make things worse, the full screen take-over prevents seniors from referencing their previous search term, which increases the load on their working memory by simultaneously holding both their previous and newly desired search term.

Integrating the voice search with the traditional search interface stands to pose three key benefits to seniors: 1) seniors would be able to rely on the visual cues and free up working memory for thinking about their searches: 2) seniors would be able to use the suggested searches to form their queries, and 3) by giving the seniors useful visual information, they may be less likely to look up and away from their interfaces (which could in turn reduce the time out errors discussed above).

Conclusion

This study provide a preliminary understanding of how seniors interact with VUIs and the factors that shape their perceptions of those systems. The findings of this preliminary study connect strongly to pre-existing work. Future work is required to test the key findings of the study and provide further insight into how these findings can best be applied by designers to support seniors.

References

[1] 2017. Health, United States, 2016 - Individual Charts and Tables: Spreadsheet, PDF, and PowerPoint files. (July 2017). https://www.cdc.gov/nchs/hus/contents2016.htm

[2] Andrew Arch. 2008. Web Accessibility for Older Users: A Literature Review. (May 2008). https://www.w3.org/TR/wai-age-literature/

[3] Grace M. Begany, Ning Sa, and Xiaojun Yuan. 2016. Factors Affecting User Perception of a Spoken Language vs. Textual Search Interface: A Content Analysis. *Interacting with Computers* 28, 2 (March 2016), 170–180. DOI: http://dx.doi.org/10.1093/iwc/iwv029

[4] Pia Bourland. 2003. The IIR evaluation model: a framework for evaluation of interactive information retrieval systems. *Information Research* 8, 3 (April 2003). http://www.informationr.net/ir/8-3/paper152.html

[5] US Census Bureau. 2017. The NationâĂŹs Older Population Is Still Growing, Census Bureau Reports. (June 2017). https://www.census.gov/newsroom/press-releases/2017/cb17-100.html

[6] Neil Charness and Walter R Root. 2009. Aging and Information Technology Use. (2009). https://www.researchgate.net/publication/228633458_Aging_and_Information_Technology_Use

[7] Dana Chisnell and Janice Redish. 2004. Designing Web Sites for Older Adults: A Review of Recent Research. (Dec. 2004). https://assets.aarp.org/www.aarp.org_/articles/research/oww/AARP-LitReview2004.pdf

[8] Peter Gregor, Alan F. Newell, and Mary Zajicek. 2002. Designing for Dynamic Diversity: Interfaces for Older People. In *Proceedings of the Fifth International ACM Conference on Assistive Technologies (Assets '02)*. ACM, New York, NY, USA, 151–156. DOI: http://dx.doi.org/10.1145/638249.638277

[9] Jeonghyun Kim. 2006. Task as a Predictable Indicator for Information Seeking Behavior on the Web. (Jan. 2006).

[10] Roger W. Morrell. 2005. http://www.nihseniorhealth.gov: the process of construction and revision in the development of a model web site for use by older adults. *Universal Access in the Information Society* 4, 1 (Sept. 2005), 24–38. DOI: http://dx.doi.org/10.1007/s10209-003-0085-3

[11] 1615 L. Street, NW, Suite 800 Washington, and DC 20036 USA202 419 4300 | Main202 419 4349 | Fax202 419 4372 | Media Inquiries. 2010. Generations 2010: What different generations do online. (Dec. 2010). http://www.pewinternet.org/2010/12/16/generations-2010-what-different-generations-do-online/

[12] Barbara M. Wildemuth and Luanne Freund. 2012. Assigning Search Tasks Designed to Elicit Exploratory Search Behaviors. In *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval (HCIR '12)*. ACM, New York, NY, USA, 4:1–4:10. DOI: http://dx.doi.org/10.1145/2391224.2391228