On How Deaf and Hard of Hearing Users Might Use Sign Language Conversational User Interfaces

Abraham Glasser, Vaishnavi Mande, and Matt Huenerfauth

Rochester Institute of Technology, Rochester NY 14623, USA
{atg2036,vm7801, matt.huenerfauth}@rit.edu

Abstract. With the proliferation of voice-based conversational user interfaces (CUIs) comes accessibility barriers for Deaf and Hard of Hearing (DHH) users. There has not been significant prior research on sign-language conversational interactions with technology. In this paper, we motivate research on this topic and identify open questions and challenges in this space, including DHH users' interests in this technology, the types of commands they may use, and the open design questions in how to structure the conversational interaction in this sign-language modality. We also describe our current research methods for addressing these questions, including how we engage with the DHH community.

Keywords: Deaf and Hard of Hearing; Accessibility; Personal Assistants; Sign Language.

1 INTRODUCTION AND MOTIVATION

The recent proliferation of voice-based personal assistant technologies poses new accessibility barriers for many Deaf and Hard of Hearing (DHH) users. As the trend of ubiquitous voice-control interfaces continue, the urgency of addressing accessibility challenges in this technology increases. Prior research has established that many DHH users are concerned about accessing this new technology, and DHH users would prefer sign-language interaction with tools like Alexa, rather than using text input or non-ASL limited gestures [10]. Since conversational user interface (CUI) systems are often based in smart speakers that may be shared across multiple users in a household, these technologies are appearing in the homes of people who are DHH, e.g. when hearing members of the household purchase these devices.

Automatic Speech Recognition (ASR) is an underlying technology that supports users' speech-based interaction with personal assistant devices. ASR automatically transcribes verbal commands into text, which is then processed by the device. The DHH population is very diverse, with the level of hearing and speaking skill varying widely among individuals [5, 9]. In a prior study [8], we found that even among the voices of DHH individuals whom professional speech pathologists and naive hearing listeners agreed were very understandable, modern ASR technology was unsuccessful at understanding the speech. This was

a concerning finding, since it indicated that our human instincts about which voices among DHH individuals may be easy to understand may not be predictive of whether ASR technology will work successfully.

As a workaround for the speech-based interaction, some modern voice-based personal assistant devices offer a text-based input option. However, this alternative text-input option is not a complete solution for personal assistant devices. There are many settings and scenarios in which text-input would be undesirable by the DHH user, such as spontaneous usage in the home. Also, there are many DHH individuals who prefer communication in ASL, and some may have difficulty with an English text-based interface, e.g. due to literacy concerns.

From a universal design perspective, since CUIs support speech-based or text-chat interaction, many DHH users will expect for these devices to also support input and output in sign language. Despite some prior misleading media reports, no CUI is currently able to accurately understand sign-language input commands. There have been claims of ASL-input capability among personal assistant devices, but these demos are generally not robust, with the technology only working for a small set of fixed commands or when the sign language message is performed in an unnatural way [2, 3, 7].

There has been recent excitement among the DHH community and researchers in the area of sign-language technologies, as evidenced by research projects, hackathons, and workshops regarding in this area [1, 4, 6, 11]. While artificial intelligence researchers and developers are still making progress in the area of sign language recognition technologies, it is important for HCI researchers to begin investigating the future interaction potential of this technology. In particular, there is a need to understand what users may want from this technology and how to best design the interaction.

As discussed in the best-paper-award winning research study at the ACM ASSETS'19 conference [6], a major bottleneck for artificial intelligence researchers working on sign language recognition is data. Currently available sign language datasets are very expensive to produce, due to the significant cost in annotating video of human signing. While these datasets may support linguistic research, when considering the complexity and diversity of the language within each, they are not large enough to support modern deep-learning methods for sign recognition.

1.1 Open HCI Research Questions for Sign-Language CUI

Several CUI-based HCI questions have arisen in recent CUI research on the needs and interests of DHH users. Rodolitz et al. called for HCI researchers to continue exploring interaction methods for DHH with CUIs before they become ubiquitous in daily lives [10]. It is currently unknown which sets of commands DHH users are most interested in when using personal assistant devices. Fundamental research is needed to investigate DHH users' interest in this technology and to understand what they want to do with it.

Many aspects of the interaction with these devices are yet to be determined: For instance, it is unknown how DHH users may want to "wake up" a CUI system

so that it is expecting a command, how the system should visually acknowledge the command from a sign-language user, what types of vocabulary or linguistic structures sign-language users prefer to use when interacting with a system, how the system should show the results to the users (e.g., as sign-language animation or written text, etc.).

In addition, the technical and performance requirements for sign recognition technologies have not yet been established: For instance, it is unknown what threshold of accuracy is needed in automatic sign recognition technology to create a usable experience for DHH users – or whether the current state-of-the-art in sign-language animation technology is sufficient for providing users with understandable output.

Since ASR technology has a much longer history than automatic sign-language recognition, there has been prior research on how hearing individuals speak when using ASR. However, there are still fundamental open questions as to how DHH individuals may linguistically interact with an inanimate device using sign language.

2 OUR RESEARCH METHODS

To address several of these open research questions, our research team has begun a research project to investigate the requirements of DHH users for conversational-based interfaces, with a particular focus on users of American Sign Language (ASL). The goal of this research is to engage with the DHH community on this topic, so that we can learn what they would want from such technologies, via interviews and a large online survey. We completed the initial interview-phase of our research, conducting interviews with 21 DHH users of ASL about their interest in using sign language to convey commands to personal assistant devices. From these interviews, we acquired a set of desired features or capabilities for the personal assistant system to understand whether the interests among this community in differs from other groups of users. These initial interviews informed the design of a questionnaire for an online survey we have also conducted with 86 DHH people across the U.S. From this survey, we identified a set of "scenarios" that users believe would be high-priority for interacting with such systems.

Currently we are working on a remote study in which DHH users can interact with an actual personal assistant device. From our initial interview and survey, we identified use-cases and commands that DHH users are interested in, which will be used to inform the creation of a set of scenarios or prompts, which may be useful during these remote, "lab-based" studies. In these sessions, DHH users will interact with a personal assistant device (with a screen for displaying output), using a Wizard-of-Oz recognition approach, in which DHH users interact with a device using sign-language commands which are "voiced" into spoken English by an interpreter. This study design will enable our team to investigate user's interests in sign-language-based interaction with these devices before automatic recognition technology is actually available.

These lab-based studies will enable us to investigate several of the open research questions outlined above, e.g. in regard to what users would actually try to do with this technology, how the interaction can best be structured, and how users would linguistically construct their commands to the device.

In addition to investigation of these HCI research questions, a side-effect of our project is that we will be collecting video recordings of the DHH users interacting in sign-language with the device. Our goal is to create a video dataset of a variety of DHH individuals interacting in ASL with such devices; such recordings will likely be of interest to computer-vision researchers interested in creating sign-recognition technology for this genre of ASL utterances.

3 Acknowledgements

This material is based upon work supported by the National Science Foundation under Award No. 1462280 and 1822747 and by Microsoft. We are grateful to our collaborator Becca Dingman for her contributions to this research.

References

1. 9th workshop on the representation and processing of sign languages: Sign language resources in the service of the language community, technological challenges and application perspectives, https://www.sign-lang.uni-hamburg.de/lrec2020/cfp.html
2. Augmented reality app can translate sign language into spoken english, and vice versa, https://bit.ly/2BfFwjB
3. Sign-language hack lets amazon alexa respond to gestures (2018), https://www.bbc.com/news/technology-44891054
4. Ai for accessibility hackathon 2019 (2019), https://blogs.partner.microsoft.com/mpn-apac/ai-for-accessibility-hackathon-2019/
5. Bigham, J.P., Kushalnagar, R., Huang, T.H.K., Flores, J.P., Savage, S.: On how deaf people might use speech to control devices (2017). https://doi.org/10.1145/3132525.3134821
6. Bragg, D., Koller, O., Bellard, M., Berke, L., Boudreault, P., Braffort, A., Caselli, N., Huenerfauth, M., Kacorri, H., Verhoef, T., Vogler, C., Ringel Morris, M.: Sign language recognition, generation, and translation: An interdisciplinary perspective (2019). https://doi.org/10.1145/3308561.3353774
7. Coldewey, D.: Signall is slowly but surely building a sign language translation platform (2018), https://tcrn.ch/39e9kcT
8. Glasser, A.: Automatic speech recognition services: Deaf and hard-of-hearing usability (2019). https://doi.org/10.1145/3290607.3308461
9. Glasser, A.T., Kushalnagar, K.R., Kushalnagar, R.S.: Feasibility of using automatic speech recognition with voices of deaf and hard-of-hearing individuals (2017). https://doi.org/10.1145/3132525.3134819
10. Rodolitz, J., Gambill, E., Willis, B., Vogler, C., Kushalnagar, R.: Accessibility of voice-activated agents for people who are deaf or hard of hearing (2019), http://hdl.handle.net/10211.3/210397
11. Ward, J.: Microsoft's windows and cortana absolutely must learn sign language (2018), https://www.windowscentral.com/microsoft-must-bring-sign-language-recognition-windows-and-cortana