Accessibility of voice-activated agents for people who are deaf or hard of hearing

REU Accessible Information and Communication Technologies Site, Gallaudet University

Evan Gambill, Mercer University
Jason Rodolitz, Pomona College
Brittany Willis, Gallaudet University

Mentor: Dr. Christian Vogler

1. Background and Objectives

- **Voice Interfaces are becoming ubiquitous**
- **Voice Interfaces are largely inaccessible to a Deaf and Hard of Hearing (DHH) audience**
  - Speech Recognition fails on Deaf speech [1, 2]
  - Text to Speech is slow and relies on other devices to generate synthesized speech [1]
- **We explored interest within the DHH community for alternative inputs to voice interfaces**
  - We used the Amazon Echo Show, a voice assistant with a screen
  - **We tested three inputs**
    - **ASL -** Similar to speech due to being natural language input, the user signs to the interface and that is translated into commands.
    - **Text to Speech -** Current technology, users type their commands and they're voiced to the interface
    - **Gesture -** A set of non-linguistic gestures that perform a variety of common tasks. Developed by the research team, based on suggestions received for a proximity sensor with five degrees of resolution.
- **Research Questions:**
  1. Is there interest in the DHH community for other ways to use with voice interfaces?
  2. What qualities do DHH users desire in smart home control systems?
  3. How do the three input methods tested compare in usability?
- **Hypothesis**
  - The most usable interfaces will use natural language
  - TTS will be significantly less usable than ASL

2. Methods

Technology Used

- Amazon Echo Show with closed captions on
- Nexus Tablet for TTS input and gesture reference
- Logitech ConferenceCam for gesture & ASL input
- FaceTime for gesture and ASL input

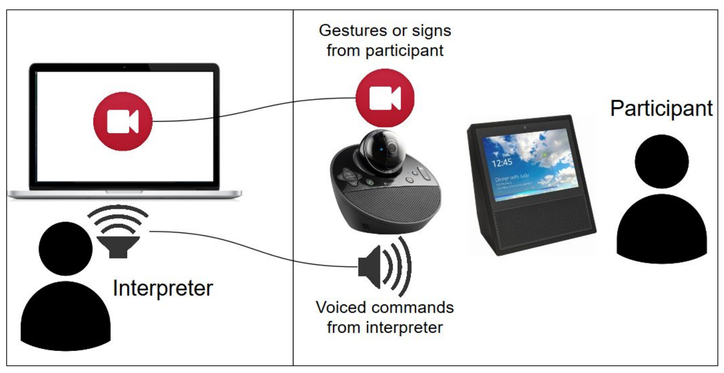


Figure 1: A diagram of our Wizard-of-Oz gesture and ASL interpretation system

To see our gesture set in action, please visit http://aict.gallaudet.edu/research/gestureset/ or scan the QR code to the right

3. Procedure

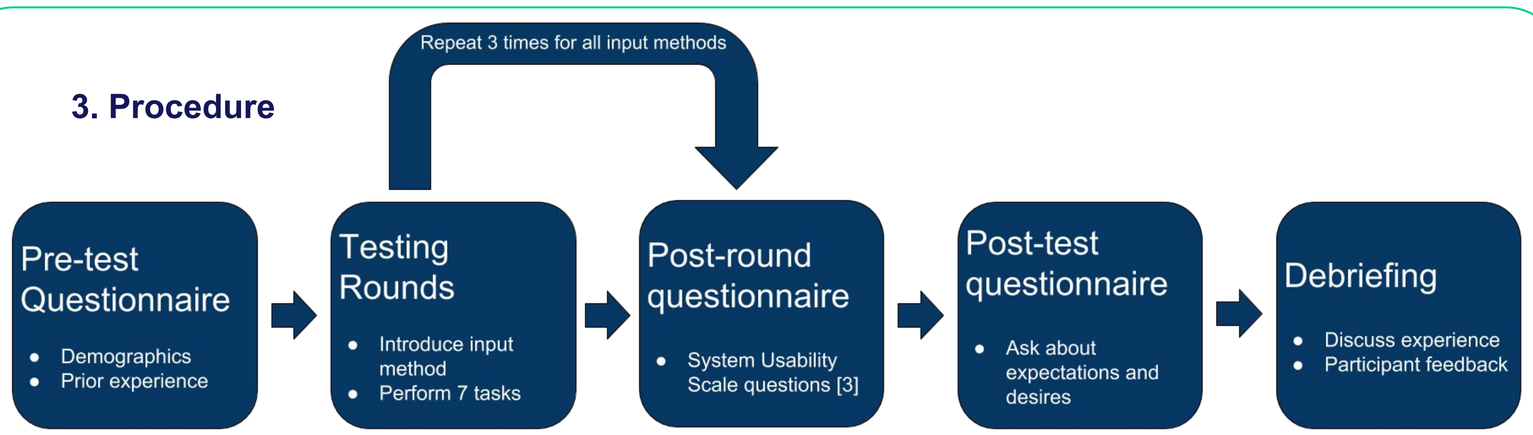


Figure 2: Input methods that participants would use to control smart home assistants, assessed after the experiment (N=12)

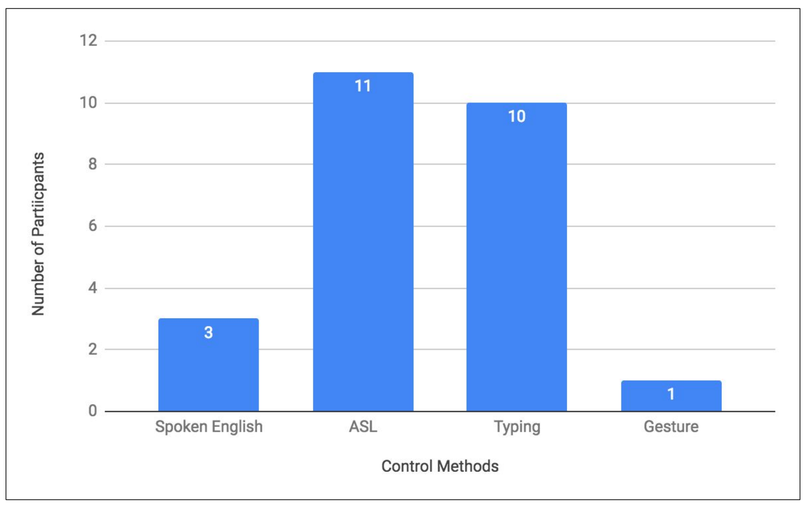
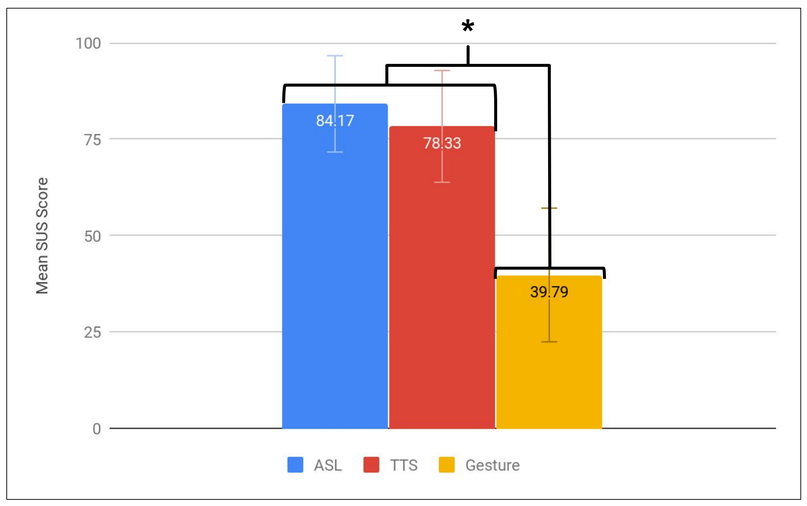


Figure 3: Average SUS Scores for each input method (N = 12).



4. Results

- Subjects (N=12) were satisfied with the ASL (SUS μ = 84.17) and TTS (SUS μ = 78.33) systems, but significantly less pleased ($p<.001$) with gesture (SUS μ = 39.79) (Figure 3)
- Subjects were similarly accurate with the ASL (μ = 6.09 out of 7) and TTS (μ = 6.27) systems, but less accurate ($p < .01$) with gesture (μ = 4.91)
- Subjects would prefer to use ASL or text based systems to control smarthome interfaces (Figure 2)
- In a survey of desired qualities of future smart home control systems, captions, the ability to input commands multiple ways, and efficiency were most important

5. Qualitative Feedback

- The TTS interface was generally satisfactory but participants wanted some changes
  - Participants couldn't hear when the TTS device made speech errors, and felt uncomfortable with it voicing for them at times. They would prefer the device to directly transmit commands
  - Apps designed for keyboard use that speak when you press enter, rather than needing a button
- Subjects would prefer a gesture set more closely related to ASL
- Subjects would like to know what signs the ASL system understands so they can use them

6. Conclusions

- Of the three input methods tested, there was no significant difference between ASL and TTS, while gesture was significantly worse than the other two
- Benefits of these natural language inputs are that they have little additional learning curve, and there has been a large amount of prior research and development on English natural language input
- However, ASL recognition and processing is still early in development
- More research is required before writing off gesture as an accessible alternative interface
- Testing on new interfaces should happen with actual use, in the home, over a long period of time

7. References

1. Bigham, J. P., Kushalnagar, R., Huang, T. K., Flores, J. P, & Savage, S. (2017). On how deaf people might use speech to control devices. *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '17).* 383-384 ACM. https://doi.org/10.1145/3132525.3134821

2. Fok, R. Kaur, H., Palani, S., Mott, M. E., Lasecki, W. S. (2018). Towards more robust speech interactions for Deaf and hard of hearing users. *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accesibility (ASSETS '18).*

3. Brooke, J. (1986). System usability scale (SUS): a quick-and-dirty method of system evaluation user information. *Reading, UK: Digital Equipment Co Ltd*, 43.

8. Acknowledgments

This work has been generously supported by an NSF REU Site Grant (#1757836) awarded to Dr. Raja Kushalnagar, PI and Dr. Christian Vogler, Co-PI.
Paula Tucker from the Gallaudet University Technology Access Program provided interpreting support for the ASL and gesture experiments. Norman Williams provided technical support.
Google provided helpful input on user interface considerations.
Person by mikicon from the Noun Project

Contact Gambill, Evan; evan.buckley.gambill@live.mercer.edu
Rodolitz, Jason; jason.rodolitz@pomona.edu,
Willis, Brittany; brittany.willis@gallaudet.edu