Alexa, I Wanna See You: Envisioning Smart Home Assistants for the Deaf and Hard-of-Hearing

Tyrone Justin Sta. Maria, Jordan Aiko Deja

¹De La Salle University, Manila, Philippines

Abstract
Smart Home Assistants (SHAs) have become ubiquitous in modern households, offering convenience and efficiency through its voice interface. However, for Deaf and Hard-of-Hearing (DHH) individuals, the reliance on auditory and textual feedback through a screen poses significant challenges. Existing solutions primarily focus on sign language input but overlook the need for seamless interaction and feedback modalities. This paper envisions SHAs designed specifically for DHH users, focusing on accessibility and inclusion. We discuss integrating augmented reality (AR) for visual feedback, support for multimodal input, including sign language and gestural commands, and context awareness through sound detection. Our vision highlights the importance of considering the diverse communication needs of the DHH community in developing SHA to ensure equitable access to smart home technology.

Keywords
Deaf and Hard-of-Hearing, smart home assistants, augmented reality, multimodal input, context-aware

1. Introduction

Smart Home Assistants (SHAs), such as Amazon Alexa and Google Home, have become increasingly integrated into our daily routines, offering seamless augmentation of various activities. For instance, users can ask the SHA a question rather than manually searching for information on a device and receive an immediate, voice-based response. This interaction, typically conducted through voice commands, allows users to focus on their tasks without diverting their attention to a screen.

However, this convenience is not equally accessible to Deaf and Hard-of-Hearing (DHH) individuals, who could otherwise benefit from the efficiencies offered by SHAs. The primary mode of interaction—voice commands—poses significant challenges for DHH users. Due to their inability to hear their speech, DHH individuals often experience difficulties with speech production, leading to reduced speech intelligibility and quality [1]. This challenge is compounded by the fact that speech recognition software is typically trained on the speech patterns of hearing individuals [2]. As a result, DHH users encounter substantial barriers when attempting to interact with SHAs. For example, research by Glasser et al. [2, 3] found that DHH users experienced a 77% word error rate when using automatic speech recognition software, compared to an 18% error rate for hearing users. These disparities highlight the need for more inclusive design approaches that enable DHH individuals to fully engage with SHAs.
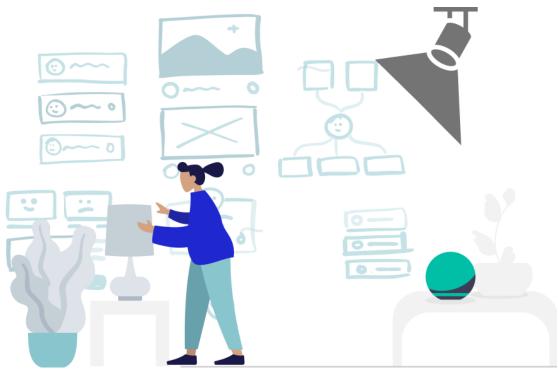
Figure 1: The vision of a smart home assistant designed for deaf and hard-of-hearing individuals that utilizes augmented reality for visual output, has multiple modes of input, and has context awareness.

SHAs predominantly provide feedback through spoken responses and visual indicators, such as lights, which create significant accessibility challenges for DHH users [4]. Some devices, like the Amazon Echo Show[1], attempt to address these issues by incorporating a screen, allowing DHH users to interact with SHAs via touch gestures and consume information visually. However, this solution disrupts the seamless interaction typically facilitated by SHAs' voice-based conversational outputs, as users must shift their attention to read the feedback on the screen [4].

To address these problems, researchers have opted to explore sign language as an input modality for SHAs [5, 6]. DHH individuals interacted with SHAs that understand sign language using sign language and finger spelling [6]. This approach is typical when communicating proper nouns or words without corresponding signs, reflecting a common practice within the community. Common commands used by these individuals involved looking up information, setting alarms, connecting with other smart devices, and adding captions to videos [5, 6].

For wake-words (e.g. *Alexa*, *Hey Google*), they typically waved in front of the device as it is the common form of getting the attention of others when they want to communicate [7, 6] or they finger-spelled the exact word. For showing feedback, DHH individuals prefer to get results through text output from a screen [5].

Existing solutions primarily focused on providing input to SHAs using sign language. Current

---

[1]https://www.amazon.com/echo-show-10/dp/B07VHZ41L8

feedback modalities that can be used comfortably by DHH individuals are text output from a screen. This interrupts the seamless interaction facilitated by SHAs' conversational outputs, as these individuals must divert their attention in reading the feedback from the screen [4]. Thus, in this paper, we envision ways in which DHH users can interact seamlessly with SHAs. We argue that such design approaches may provide a direction for considering the needs of DHH users in future iterations of SHAs.

2. Smart Home Assistant for DHH

We present three approaches to make better Smart Home Assistants for DHH. Namely, (i) using AR, (ii) integrating multimodal input and (iii) having context-aware spaces.

2.1. Using Augmented Reality

While seeing users in their homes typically navigate their environment easily, designing Smart Home Assistants (SHAs) for Deaf and Hard-of-Hearing (DHH) individuals requires special consideration to ensure they can multitask effectively. The goal is to enable these users to receive and process information without diverting their attention to a specific device. One approach is incorporating augmented reality (AR), which provides visual feedback directly within the user's physical space. This can be achieved by integrating mobile steerable projectors, as described in [8, 9, 10]. Using projectors eliminates the need for additional wearable accessories, which can be uncomfortable for DHH users, particularly those who wear hearing aids [10]. By embedding AR into SHAs, DHH individuals can benefit from visual feedback that remains within their field of view, allowing them to stay engaged with their surroundings without being tethered to the SHA device. For instance, ingredient lists could be projected nearby when cooking, or a timer notification could be displayed even if the user has stepped away from the stove. Such enhancements would significantly improve their ability to consume information and complete tasks seamlessly and efficiently.

2.2. Multimodal Input

In designing smart home assistants (SHAs) for individuals who are deaf and hard-of-hearing (DHH), it is essential to consider their preferred mode of communication. Supporting multiple input modes can help cater to their needs, whether sign language [11] or verbal communication with the help of hearing aids [4, 3]. When interacting with SHAs, users typically begin with invoking a wake-word followed by a query, with responses delivered through auditory feedback, on-screen text, or device control. However, when using sign language, certain words or expressions lack direct translation, leading to finger spelling, which takes more effort than verbal communication. For example, DHH individuals used finger spelling to express the wake-word "Alexa" [6]. Using gestural inputs should be explored to lessen their efforts on such tasks [7]. We could extend this to menial tasks such as asking for the weather, setting alarms, and controlling other smart devices. For example, users can clap once to turn on a smart light or show their palm to ask for the weather. By considering a range of input modes, SHAs can be designed to be more accessible and user-friendly for DHH individuals.

2.3. Context Aware Spaces

Deaf and Hard-of-Hearing (DHH) individuals often rely on voice-to-text programs to aid in understanding speech, extending their auditory perception [3, 2]. This concept can be further expanded to Smart Home Assistants (SHAs) by enabling them to monitor and interpret household sounds, such as the whistle of boiling water, door knocks, or the activity of appliances [12, 13]. By understanding and responding to these auditory cues, SHAs can offer more context-aware assistance, extending their functionality beyond traditional tasks like internet searches, setting alarms, or providing weather updates. Additionally, integrating augmented reality (AR) to deliver visual alerts and information can significantly enrich the daily experiences of DHH individuals, especially in situations that depend on audio cues. For example, an SHA could notify a DHH user of a boiling kettle while they are occupied with another task or alert them to someone knocking at the door, thereby enhancing their awareness and interaction with their surroundings.

3. Discussion

We based the visions of a SHA designed for DHH individuals to make the technology accessible and augment their daily activities. SHAs have become ubiquitous and are becoming integrated into our daily lives [14]. We introduce ways for DHH individuals to interact with SHAs seamlessly, from simply an information provider and smart device controller to an extension of their hearing around their living space. We believe that the features we presented may represent that of a SHA designed with DHH users in mind, however this is subject to further exploration and validation with actual users. We also discuss some benefits and disadvantages that go with it as well.

SHAs usually have smart speakers and a mic that recognizes voice inputs and outputs audio feedback. However, these modalities are not accessible to DHH individuals [1, 2]. This is why the SHA designed for DHH users must attempt to become a companion that can extend their senses to facilitate seamless interaction with smart home technology alongside their overall home environment.

It will also be equally challenging to set up an AR environment within the home [15]. SHAs usually come out of the box with smart speakers that are sometimes accompanied by a display without any additional setup. Using a steerable projector (e.g. such as in the work of [8]) comes up with challenges in terms of its placement [16]. It would be ideal for the projection to capture every corner of the room without being obstructed. However, achieving the ideal placement requires additional equipment and expertise [17].

In the design of SHAs for DHH, we factored in several modalities and parameters, such as image data for sign language detection and sound data detecting voice and the home environment [18, 4]. Security and privacy must be strictly upheld in all these conditions, especially if it involves data on users' activity [19]. As we factor in more considerations in designing SHAs for DHH in the future, we introduce more benefits and issues with them. We posit that SHAs designed for DHH must always consider the unique communication needs and preferences of the DHH community [20] while prioritizing accessibility, inclusivity [21], and robust security measures to safeguard users' sensitive data.

4. Summary

We outlined the challenges that DHH individuals face when interacting with SHAs. We looked at existing solutions that address these problems. Through our examination of the existing challenges and current solutions, we have formulated a vision for SHAs designed for DHH users. Our vision prioritizes accessibility. Incorporating features such as AR integration, support for multimodal input, and context awareness. Furthermore, we emphasize the importance of considering the unique communication needs and preferences of the DHH community for both input and output modalities when developing SHAs aligned with DHH users.

References

[1] M. S. Thirumalai, S. Gayathri, et al., Speech of the hearing impaired, (No Title) (2004).

[2] A. T. Glasser, K. R. Kushalnagar, R. S. Kushalnagar, Feasibility of using automatic speech recognition with voices of deaf and hard-of-hearing individuals, in: Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility, 2017, pp. 373–374.

[3] A. Glasser, K. Kushalnagar, R. Kushalnagar, Deaf, hard of hearing, and hearing perspectives on using automatic speech recognition in conversation, in: Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility, 2017, pp. 427–432.

[4] J. Blair, S. Abdullah, It didn't sound good with my cochlear implants: Understanding the challenges of using smart assistants for deaf and hard of hearing users, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 4 (2020) 1–27.

[5] A. Glasser, V. Mande, M. Huenerfauth, Understanding deaf and hard-of-hearing users' interest in sign-language interaction with personal-assistant devices, in: Proceedings of the 18th International Web for All Conference, 2021, pp. 1–11.

[6] A. Glasser, M. Watkins, K. Hart, S. Lee, M. Huenerfauth, Analyzing deaf and hard-of-hearing users' behavior, usage, and interaction with a personal assistant device that understands sign-language input, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, 2022, pp. 1–12.

[7] V. Mande, A. Glasser, B. Dingman, M. Huenerfauth, Deaf users' preferences among wake-up approaches during sign-language interaction with personal assistant devices, in: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, 2021, pp. 1–6.

[8] J. R. Cauchard, Mobile multi-display environments, in: Proceedings of the 24th annual ACM symposium adjunct on User interface software and technology, 2011, pp. 39–42.

[9] Lightform announces project lfx | steerable projected interfaces, 2021. URL: https://youtu.be/3XIGxNP3-Mk.

[10] R. Mathew, G. W. Tigwell, R. L. Peiris, Deaf and hard of hearing people's perspectives on augmented reality interfaces for improving the accessibility of smart speakers, in: International Conference on Human-Computer Interaction, Springer, 2024, pp. 334–357.

[11] R. E. Mitchell, T. A. Young, B. Bachelda, M. A. Karchmer, How many people use asl in the united states? why estimates need updating, Sign Language Studies 6 (2006) 306–335.

[12] G. Laput, K. Ahuja, M. Goel, C. Harrison, Ubicoustics: Plug-and-play acoustic activity recognition, in: Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, 2018, pp. 213–224.

[13] J. Wu, C. Harrison, J. P. Bigham, G. Laput, Automated class discovery and one-shot interactions for acoustic activity recognition, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–14.

[14] W. K. Edwards, R. E. Grinter, At home with ubiquitous computing: Seven challenges, in: Ubicomp 2001: Ubiquitous Computing: International Conference Atlanta Georgia, USA, September 30–October 2, 2001 Proceedings 3, Springer, 2001, pp. 256–272.

[15] F. Lamberti, F. Manuri, A. Sanna, G. Paravati, P. Pezzolla, P. Montuschi, Challenges, opportunities, and future trends of emerging techniques for augmented reality-based maintenance, IEEE Transactions on Emerging Topics in Computing 2 (2014) 411–421.

[16] Y. J. Park, H. Ro, N. K. Lee, T.-D. Han, Deep-care: Projection-based home care augmented reality system with deep learning for elderly, Applied Sciences 9 (2019) 3897.

[17] M. Speicher, B. D. Hall, A. Yu, B. Zhang, H. Zhang, J. Nebeling, M. Nebeling, Xd-ar: Challenges and opportunities in cross-device augmented reality application development, Proceedings of the ACM on Human-Computer Interaction 2 (2018) 1–24.

[18] N. Tran, P. S. DeVries, M. Seita, R. Kushalnagar, A. Glasser, C. Vogler, Assessment of sign language-based versus touch-based input for deaf users interacting with intelligent personal assistants, in: Proceedings of the CHI Conference on Human Factors in Computing Systems, 2024, pp. 1–15.

[19] A. Pradhan, K. Mehta, L. Findlater, "accessibility came by accident" use of voice-controlled intelligent personal assistants by people with disabilities, in: Proceedings of the 2018 CHI Conference on human factors in computing systems, 2018, pp. 1–13.

[20] J. Blair, S. Abdullah, Understanding the needs and challenges of using conversational agents for deaf older adults, in: Companion Publication of the 2019 Conference on Computer Supported Cooperative Work and Social Computing, 2019, pp. 161–165.

[21] D. Jain, A. Lin, R. Guttman, M. Amalachandran, A. Zeng, L. Findlater, J. Froehlich, Exploring sound awareness in the home for people who are deaf or hard of hearing, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–13.