Sign Language Conversational User Interfaces Using Luminous Notification and Eye Gaze for the Deaf and Hard of Hearing

Takashi Kato*, Akihisa Shitara†, Nobuko Kato*, and Yuhki Shiraishi*
*Tsukuba University of Technology, Japan
Email: {a203101,nobuko,yuhkis}@a.tsukuba-tech.ac.jp
†University of Tsukuba, Japan
Email: theta-akihisa@digitalnature.slis.tsukuba.ac.jp

***Abstract*—We investigate the design of a user-friendly natural user interface for the deaf and hard of hearing (DHH). Voice-based conversational user interfaces (CUIs), such as Amazon Alexa and Google Assistant, are becoming increasingly popular among consumers. DHH users may not be aware of notifications from CUIs, may not be able to obtain response information, and may have difficulty waking up the CUIs. In this study, we designed a system that adds luminous notifications and sign language to the CUI and conducted Wizard of Oz experiments to investigate whether the system can provide an optimal user experience for DHH users. The results suggest that luminous notifications improve the usability and make notifications easier. After assessing the necessity of sign language/text display, we found that people with longer sign language histories tend to use sign language, and all people require the use of a text display. The percentage of DHH users who gazed at the system before entering commands into the system (93.4%) also suggests that gazing can be an effective way to wake up the system. Our findings provide guidance for future CUI designs to improve the accessibility for DHH users.**

Keywords—Deaf and hard of hearing, sign language, accessibility, user interface.

I. INTRODUCTION

The recent proliferation of voice-based interaction devices has created new accessibility barriers for many deaf and hard of hearing (DHH) users. In this paper, we propose a conversational user interface (CUI) to improve the accessibility for DHH users. In recent years, human-computer interaction (HCI) researchers have begun to evaluate sign language processing technologies from an interdisciplinary perspective [1], and as a result, this topic was addressed at a workshop on user interfaces [2]. Accessibility studies of CUIs by DHH users have reported that sign language is more suitable than gestures and text as an alternative input method to speech [3]. It has also been reported that the use of sign language is preferable to touch screens as an input method [4]. However, both of these studies substitute sign language for speech in voice user interfaces (VUIs) and do not consider the physicality of DHH users who mainly use visual information. There are few research reports on such CUIs, and HCI researchers are not working on devising or implementing design guidelines [1]. We therefore need to design CUIs accessible to DHH users, including notification methods, information transmission methods, and eye gaze.

We can classify the output of the CUIs into two patterns: the system notifying us of an "Alarm Notification" or a "Phone call," and the system describing the "Weather" or "News" that the user has requested. In the former case, DHH users do not notice the notification from the CUI, and in the latter case, they do not notice when the CUI has ended its response. In other words, it is essential to investigate the best output method of the system for DHH users, instead of the voice output provided by CUIs. By contrast, luminous notifications are familiar in deaf culture. For example, DHH users use intercoms, alarms, and fire alarms with luminous notification functions in their daily lives [5]. In addition, a luminous device that transmits the direction of the sound source of the surrounding alarms to DHH users with light has been developed [6]. In this study, we investigate whether luminous notifications can improve usability for DHH users.

Subtitling has recently become an accessibility feature for CUIs with displays [7]. Even if DHH users use this feature, there is a concern that the user experience will be lower if the system outputs subtitles. This is because an interaction will mainly be in sign language if the sign language input from the user is enabled. With devices such as Alexa entering the home, there have been reports of increased hands-free interaction with devices placed in the kitchen or living room [8]. Therefore, it is conceivable that people will be more likely to interact while doing other things. DHH users should also be able to capture responses from CUIs while doing other things. Here, it is essential to clarify whether the sign language/text output method of the system affects the user experience of DHH users, based on the difference between flowing and remaining signs for a certain period. Moreover, whereas designers must translate into the speaker's language in the case of television and the Web, CUIs are transmitted by a computer, and thus the language can be adjusted to suit the recipient. In other words, the designer should consider the output method of sign language/text, considering the user's preference in terms of attributes. Therefore, we investigate the preferences of DHH users for sign language and subtitles under the condition of parallel work when CUIs provide not only subtitles but also sign language.

To initiate a dialog with a VUI, users need to use wake words, such as "Alexa" for Amazon Alexa, "OK, Google" for Google Assistants, and "Hey, Siri" for Siri. Although studies on the waking up of personal assistant devices have compared the methods preferred by DHH users [9], they have not examined the use of eye contact, which is essential for

 ISBN: 978-1-61208-870-9

TABLE I
LINGUISTIC MODALITIES OF HEARING AND DHH INTERACTION WITH CUIS.

User \ Type	Conversation		Call	
	CUI → User	User → CUI	CUI → User	User → CUI
Hearinng	Voice		Voice	
DHH	Sign Language / Text **		Luminnous *	Eye Gaze ***

“*”: RQ1, “**”: RQ2, “***”: RQ3.

starting a conversation in interpersonal communication with DHH signers [10]. The authors believe that eye gaze allows for a natural interaction without an explicit wake-up and increases user satisfaction. Therefore, we investigate the possibility of using an eye gaze in a CUI.

In this study, we devised the following research questions to improve the accessibility of CUIs by DHH users.

- RQ1: Does the light-based response of the CUI improve the usability for DHH users?
- RQ2: What is the best sign language/text display method for CUI for DHH users?
- RQ3: Is eye gaze an effective method of waking up to CUI for DHH users? If yes, what kind of gaze input is effective?

As shown in Table I, RQ1–RQ3 are research questions that cover the mutual input/output modalities that DHH users want to achieve when interacting with CUIs.

DHH users use sign language/text modalities when interacting with CUIs (RQ2). They use the luminous notification modality when calling from CUIs to DHH users (RQ1). By contrast, when DHH users call to (wake up) a CUI (RQ3), they use the eye gaze modality. We investigate whether using these mutual input/output modalities in CUIs can improve the user experience of DHH users.

This study contributes empirical knowledge regarding the preferences and concerns of DHH users, such as notification, sign language/text display methods, and whether eye gaze is practical for waking up a device, thereby guiding future system designers.

In Section II, we describe related studies on CUIs and eye gaze. In Section III, we provide information on the participants, the device architecture, and the experimental procedure. In Section IV, we describe the results obtained from the experiment. In Section V, we provide a discussion of the research questions. Finally, in Section VI, we give some concluding remarks and areas of future work.

II. RELATED WORK

A. Conversational User Interface

In recent years, research on natural user interfaces, which enable natural and intuitive operations in humans, has been progressing. For greater convenience and prosperity, a “conversational” interaction with users is required that enables an intuitive operation and mental support [11]. Advances in speech recognition, speech synthesis, and natural language processing have enabled humans to interact naturally with user interfaces [12].

The VUI is an interface that uses the auditory and vocal organs and utilizes their functions as a language. With the improvements in voice recognition and natural dialog technology, VUIs are attracting attention and are becoming more popular because they can be used as hands-free devices without interrupting the user’s work flow [13]. However, it is difficult for DHH people to use a VUI [14], [15].

Gestural user interfaces are those that use visual and physical functions, such as arms, fingers, and facial expressions. Example applications include motion sensing in Google Pixel 4 [16] and drones. However, gestures do not have linguistic properties, and thus their expressive power is limited.

We believe that optimal sign-language-based CUIs for DHH people should be an interface that utilizes body language, including vision, arms, fingers, facial expressions, and language. In addition, because hearing is a more passive and subconscious stimulus than sight, as long as others are nearby, speech can easily be noticed in any direction [17]. Therefore, we expect that our sign-language-based CUIs will solve the privacy problem of VUIs in an office space. Moreover, because it does not require speech recognition, it does not cause recognition problems in noisy environments. In addition, we expect sign-language-based CUIs to help overcome the problems of gestural user interfaces [18], such as reduced expressiveness and increased memory load, because they use natural language for interaction. By contrast, through the technology of sign language recognition advancement, HCI researchers have begun to consider user interfaces that can interact with sign language [19].

B. Eye Gaze in Sign Language

The preferred wake-up techniques of DHH users in descending order of preference are the use of the ASL sign-name of the device, waving in the direction of the device, clapping, using a remote control, using a phone app, and fingerspelling the English name of the device [9]. However, the comparison provided in this study does not include eye gaze. By contrast, interpersonal communication of a DHH signer requires the other person’s attention when calling out to them in comparison to those with hearing [20]. In addition, when initiating a conversation with a DHH user, it is necessary to make eye contact, which can be done by tapping the user on the shoulder or waving a hand [10]. With this background in mind, this study investigates through a Wizard of Oz experiment whether eye gaze is essential for DHH users to operate a CUI.

III. METHODOLOGY

A. Participants

Using a mailing list, we solicited the cooperation of 12 DHH students in their 20s to participate in the experiment.

We also investigated the characteristics of the participants to analyze the effect of their attributes on the results of the

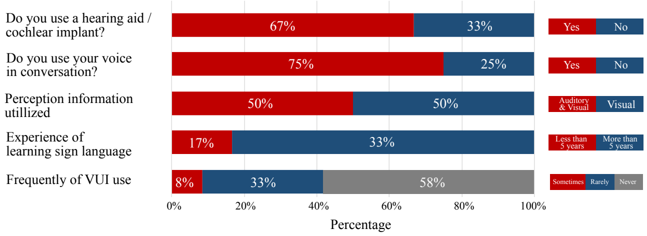

Figure 1. Graph of prior experience with voice user interfaces for 12 participants.

experiment. Specifically, we conducted a preliminary questionnaire survey on the age, gender, and cochlear implant/hearing aid use of the participants to determine whether they use their voices when communicating, whether they mainly use auditory or visual communication, and whether they use both, as well as their sign language history and their experience using VUIs. Figure 1 shows the results. The age of our participants, 8 male and 4 female, ranged between 20 and 24. We asked the participants to rate their experience of using VUIs on a 4-point Likert scale (1 = usually, 2 = sometimes, 3 = rarely, and 4 = never). The results showed that the response of one participant was 2 = sometimes, four participants responded 3 = rarely, and seven participants responded 4 = never. The majority of the participants commented that “they could not have their voices recognized” or that “they lived a life where they did not speak using voice.” Research mentioning that a minimal number of DHH users use personal-assistant devices [21] indicates a similar trend to that of the participants in this experiment.

This study was approved by the Research Ethics Review of Tsukuba University of Technology, where the experiment was conducted. The duration of the experiment was 90 min, and the honorarium paid to the participants was 1,305 yen (approximately $12).

B. Device Architecture

The basic configuration of the system built in this experiment was an iPad, a Meross Smart Wi-Fi LED Bulb (LED Bulb), and a GoPro HERO9 camera. Figure 2 shows the appearance and operation of the system. We set four tasks that the system can perform: “Phone call,” “Alarm settings and notifications,” “Checking the Weather,” and “Checking the News.” An Apple iPad simulated Alexa, and the display was created using Microsoft PowerPoint 2019 and combined with the signer’s video. To switch screens remotely, the remote function of Keynote was used. LED bulbs can be set to any color (16 million RGB colors) and a blink cycle, and it can be controlled remotely from a smartphone app. An LED bulb flashes yellow when the system notifies the user of a “Phone call” or “Alarm notification” and light green when the system provides “Weather” or “News” to the user. In addition, the GoPro camera views the sign language input from the user.

C. Procedure

The recognition rate in a real-life continuous sign language recognition system developed in 2019 was 39.6% [22]. Therefore, it is impossible to conduct experiments incorporating sign language recognition technology to interact with a user interface using sign language. The Wizard of Oz method [23], [24] is a solution to this problem. With this method, a human called a wizard pretends to be the system and interacts with the user. In the Wizard of Oz method, even if the entire system is not yet complete, the wizard can complement the undeveloped parts of the system and make it work. To verify RQ1–RQ3, we conducted an experiment based on the Wizard of Oz method. Figure 3 shows the experimental environment. In this experimental environment, we assumed that the participants interacted with the system while working on their PCs. Therefore, we placed the system on the left side of the desk in front of the participants at 45 degrees and the work PC in front of them. We tried to make the participants aware of the system response while the system was operating so that they did not have to constantly look at the system. We also aligned the system at the eye level of the participants. The instruction device prompted the participant to issue commands to the system at certain times. We incorporated a program in PsychoPy (v2021.1) [25] to display numbers and/or English letters at random positions on the screen. In addition, the frame rate of the installed camera was 50 fps.

The critical points for the participants and the experimenter (Wizard) in this environment are as follows.

Participant

1) Owing to the nature of the Wizard of Oz method, the participants assume that there is a sign language recognition system and do not know that a person (wizard) is operating the system.
2) During the experiment, the participant has to continuously work on the task of “entering numbers/English letters displayed at random positions on the work PC screen with the keyboard as they appear.”
3) The participant commands the system using sign language commands for “Setting the alarm,” “Checking the weather,” and “Checking the news.” The participant presses the button as soon as the end of the description

Figure 2. System prototype.

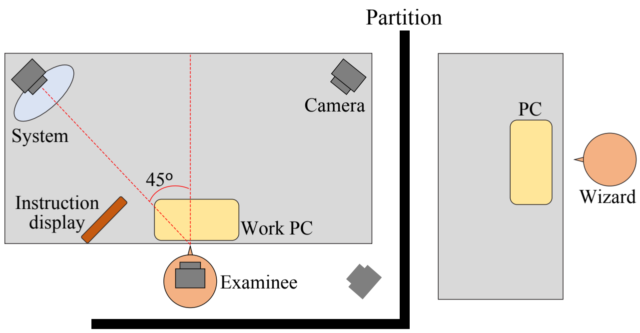

Figure 3. Experiment setup.

of "Weather" or "News" from the system is noted.
4) During the work, the participant uses sign language commands to stop the "Alarm notification" and "Phone call" sent by the system.
5) During the experiment, the user wears a GoPro attached to a head strap mount.

Experimenter (Wizard)

1) Owing to the nature of the Wizard of Oz method, the experimenter must not let the participants know that the experimenter is operating the system when performing the sign language recognition system.
2) During the experiment, the experimenter operates the system as well as the LED bulb.
3) When we asked the participants to conduct a specific task at an arbitrary time, we showed them the content of the task and an example of the command to be performed, and we immediately turned off the screen after confirming that the participants understood the task.

Before the experiment, we explained how to use the system and how the system behaves for each of the four tasks. In addition, to familiarize the participants with command execution using sign language, we gave them a practice session to perform a task equivalent to the real one before the actual experiment was conducted. The participants conducted each of the four tasks once and repeated them twice. To eliminate the order effects, the order of the tasks for each participant and the two conditions, "Luminous/Conventional," were counterbalanced.

D. Analysis Method

For the time analysis using video, we applied the ELAN [26] tool.

For RQ1, we used the system usability scale (SUS) [27], a widely applied evaluation index for a quantitative evaluation of usability, to examine the usability of "Luminous" and "Conventional." In addition, we believe that improved usability is also related to awareness. To evaluate the awareness of the notifications from the system, we measured by video the time between the notifications and when the participant noticed and reacted to them. We defined the reaction time for a "Call" as the time between the change in the display screen as the reaction start point and the users turning their eyes to the screen as the reaction endpoint. However, if the light turns on before the screen changes, the reaction start point is when the light turns on. We defined the reaction time for a "Response" as the time between the change in the display screen as the reaction start point and the user pressing the button as the reaction endpoint. However, if the light turns off before the screen changes, the reaction start point is defined as the time when the light turns off.

For RQ2, we examined the participants' need for sign language/text. After the experiment, we administered a questionnaire to determine the need for sign language/text using a five-point Likert scale (1 = agree, 2 = agree a little, 3 = neutral, 4 = disagree a little, and 5 = disagree).

For RQ3, we examined whether the participants gazed at the system before giving a command in sign language. For this purpose, we measured the percentage of the total number of times the participants gazed at the system at least once in the 5 s before the sign language command and the time between the start of the gazing and sign language using video. For the data to be analyzed, there was a scene during the experiment in which the system responded to an "Alarm notification" or "Phone call," and the user was given a command to stop the system. The user looks at the response screen before making a sign language command and does not provide analysis data to investigate whether the user gazed at the screen before the sign language command. The data for analysis are the three tasks for which the user actively gives a sign language command, i.e., "Setting the alarm," "Checking the weather," and "Checking the news."

IV. RESULTS

A. System Usability Scales

Figure 4 shows the results of the SUS investigated after the experiment. The mean SUS value of "Luminous" was 80.67, with a standard deviation of 7.62, and that of "Conventional" was 68.96, with a standard deviation of 14.6. As a result of the Wilcoxon signed-rank test, "Luminous" was found to be significantly higher ($p < .05$).

B. Reaction Time

Figure 5 shows the results of the reaction time. The mean reaction time to "Alarm notification" and "Phone call" of

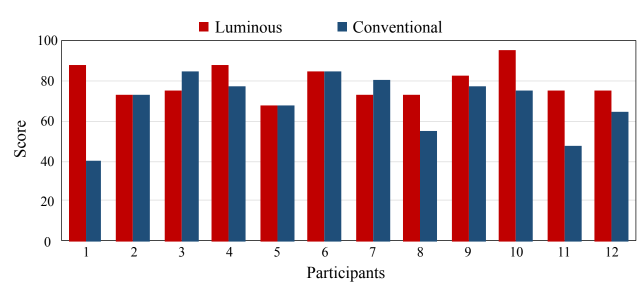

Figure 4. SUS score for each participant (N = 12).

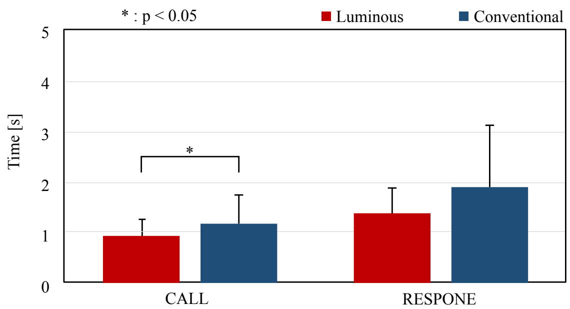

Figure 5. Reaction time for each feedback from the system.

"Luminous" was 0.91 s with a standard deviation of 0.35 s, and the mean value of "Conventional" was 1.19 s with a standard deviation of 0.57 s. The Wilcoxon signed-rank test showed that the reaction time was significantly shorter for "Luminous" ($p < .01$). The mean reaction time to the end of "Weather" and "News" for "Luminous" was 1.37 s with a standard deviation of 0.50 s, and the mean value of "Conventional" was 1.91 s with a standard deviation of 1.22 s. The Wilcoxon signed-rank test showed no significant differences between "Luminous" and "Conventional" ($p = .36$).

C. Necessity of Sign Language/Text

Figure 6 shows the results of a 5-point Likert scale to assess the need for sign language and text for the 12 participants, respectively. In terms of sign language, we found the following: 1: "Agree" was reported by four participants. 2: "Agree slightly" was reported by two participants. 3: "Neutral" was reported by three participants. 4: "Disagree slightly" was reported by two participants. 5: "Disagree" was reported by one participant. There were three participants who did not need sign language (4,5), and their sign language experience was, in order of shortest to longest, three years (1st), five years (2nd), and fifteen years (5th). By contrast, for text, "1 = Agree" was reported by nine of the participants, and "2 = Agree slightly" was reported by three of the participants.

D. Eye Gaze

During the experiment, a pattern occurred in which the experimenter turned off the screen of the instructional device late, indicating the task to be performed, and the participant gave a sign language command while reading. We removed these data from our analysis because they were unsuitable for

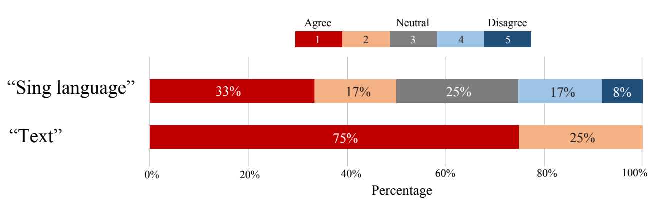

Figure 6. Necessity of "Sign language" / "Text".

examining whether the participants were gazing at the system. The participants (N = 12) input sign language commands into the system 69 times: 23 times for "Setting the alarm," 24 times for "Checking the weather," and 22 times for "Checking the news." Table II lists the percentage of the total number of times the participants gazed at the system at least once during the 5 s before the sign language command and the average time from the start of gazing to the start of the sign language, as well as the standard deviation and minimum and maximum values.

TABLE II
PERCENTAGE OF EYE GAZE, MEAN AND STANDARD DEVIATION OF TIME OF EYE GAZE

Task	*Percentage* (%)	*Mean*±*SD* (*s*)	*Min* (*s*)	*Max* (*s*)
Alarm	100	0.76±0.61	0.20	3.18
Weather	100	0.43±0.23	0.10	1.08
Alarm	86.4	0.59±0.44	0.20	2.08
Total	93.4	0.59±0.47	0.10	3.18

A high percentage of the total number of users gazed at the system before using the sign language commands.

Three participants, P3, P8, and P9, waived before applying the sign language. These three participants had experience using VUIs and knew that they should use a waking command. The interviews also revealed that they thought it was necessary to take explicit action before talking to the system during this experiment.

V. DISCUSSION

A. RQ1: Efficacy of Luminous Notification

The results described in Section IV-A suggest that luminous notification improves the usability of DHH users in noticing notifications from the CUI. In addition, the reaction times to "Alarm notification" and "Phone call" were significantly shorter when using a luminous notification, suggesting that it is easier to notice such notifications from the system.

Participants commented, "I am familiar with luminous notification methods, such as the intercom system in my house, which notifies me by light, so it would be more impressive to add light to the system as well. I can notice the light notification even when I am concentrating on my work." However, there were also comments such as "I feel uncomfortable with the luminous notification because I live my life relying on sound. Therefore, the system may not be suitable for people who use their daily hearing functions.

From Figure 4, we can see that the usability of P3 and P7 decreases with a luminous notification. They commented that they did not feel the need to use a luminous notification because they only noticed the change in the system screen. This may be because there were cases in which DHH users could respond to conventional methods [28]. In this experimental environment, the system was placed on the left side of the desk in front of the user at a 45 degree angle and within the peripheral vision. During this experiment, we placed the system within the peripheral vision of the front of the user, and thus some of the subjects noticed changes in the screen without looking at the system.

By contrast, there were no significant differences in the reaction time to the end of the "Weather" and "News" responses when using the luminous notification, as described in Section IV-B. However, some of the participants commented positively that "it was convenient to know when the response ended without having to look at the system." By contrast, others commented negatively that "luminous notification was not necessary for information (weather and news) that I wanted," and "the light was too bright." As a result, the usability of the system can be improved by reducing the light exposure and improving the luminous notification method, although the noticeability remains the same.

A participant commented that it is preferable to increase the brightness of the display, as in ON AIR, instead of directly informing us with LED bulbs. For a luminous notification, we used LED bulbs, which were initially used as lighting fixtures. Therefore, we need to consider a way to change or blink the brightness of the display directly instead of using external LEDs.

In this study, we incorporated a luminous notification as a means of responding to DHH users. However, some participants commented, "I think it would be easier to notice if there was a notification method using vibration as well as light." In the future, we believe that it will be necessary to conduct a verification experiment that includes a vibration notification. In addition, because we placed the system at the front of the participants in this experiment, we need to find a way to make them aware of the notifications from the rear.

B. RQ2:How to Suitably Display Sign Language/Text

From Section IV-C, we can see that all of the participants need to display text regardless of the user attributes. By contrast, the necessity for the sign language display varies from participant to participant. In addition, we can see that those who have not signed for a long time tend not to believe that sign language is necessary.

The participants who did not need sign language commented that they did not understand sign language and had trouble processing information when both sign language and text were output simultaneously. By contrast, the participants who needed to use sign language commented that the sign language display made it easier for them to remember the system responses. Some of the participants commented, "When I look at the task screen while working, the remaining text is better than the flow of the sign language."

For hearing users, interaction with VUIs has the advantage of being eyes-free [13]. Therefore, users frequently interact with CUIs while conducting other tasks. However, in the case of DHH users, the advantage of eyes-free interaction is lost because they cannot acquire audio information and instead gaze at the screen. To complement this, we anticipate that DHH users will need text information that they can recognize, even if they look away for a moment. One possible solution to this problem is to stop the sign language when the user looks away and start again when the user looks back.

C. RQ3:Efficacy of Eye Gaze

Section IV-D shows that the participants tend to gaze at the screen before speaking in sign language.

During this experiment, we did not provide instructions on how to wake up the device. Nevertheless, the participants naturally gazed at the system with a high probability.

By contrast, 3 of the 12 participants did not gaze at the system but made hand gestures instead. When DHH users use waving as a wake-up method, there is a concern that signs made while talking to another person may be recognized as waving at unexpected times, such as during a "Phone call" or "Alarm notification." In addition, we believe that gazing is a more natural way of interacting than waving every time a command is used. These results suggest that gazing is a compelling wake-up method.

When Alexa waits for a response from the user, there is a time limit of 8.0 s [29]. From Table II, the maximum time between gazing at the system and the start of sign language was 3.18 s. In other words, when DHH users use gazing as the wake-up method, they can use commands within the system's waiting time.

D. Limitations

In this study, the age range of the participants was low, and the sample size was small, with all participants being university students. In addition, 92% of the participants had little or no experience using VUIs. During this experiment, we did not present a line of sight from the system. In other words, the user cannot judge whether the system and the user's gaze coincide.

VI. CONCLUSION AND FUTURE WORK

In this study, a system was designed that adds a luminous notification and sign language to CUIs to provide an optimal user experience for DHH users. We surveyed DHH users (N = 12) to determine the optimal input and output modalities of the CUIs.

We tested the effectiveness of a luminescent modality applied to call from the CUIs to DHH users using the SUS and found that the usability of the system was improved using a luminous notification. The reaction time of the luminous notification to the call from the system (0.91 ± 0.35 s) was significantly shorter than that of the conventional system (1.19 ± 0.57 s). After assessing the necessity of sign language/text display, 50% of the participants answered that they needed to use sign language for the display method of CUIs, and 100% of the participants answered that they needed to use text. We also found that people with longer sign language histories tended to use sign language. As a result of the experiment on whether the gazing modality is effective for DHH users to wake up the CUIs, the percentage of times they gazed at least once in the system during the five seconds before the sign language commands was as high as 93.4%, suggesting that gazing is a compelling wake-up method.

Future studies will include a more diverse group of people such as children and the elderly. We also plan to conduct an evaluation experiment after using a system entirely accustomed to the participants. We are also planning to conduct an evaluation experiment to determine whether gaze can be used as a wake-up command.

We expect that our findings will become one of the design guidelines used for CUIs suitable for DHH users.

ACKNOWLEDGMENT

This work was partially supported by JSPS KAKENHI Grant Number #19K11411 and the Promotional Projects for Advanced Education and Research in NTUT. We would like to thank Editage (www.editage.com) for English language editing.

REFERENCES

[1] D. Bragg et al., “Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective,” Proceedings of ASSETS ’19: The 21st International ACM SIGACCESS Conference on Computers and Accessibility, pp. 16–31, NY, USA, Oct. 2019.

[2] D. Bragg et al., “Sign Language Interfaces: Discussing the Field’s Biggest Challenges,” Proceedings of CHI EA ’20: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–5, NY, USA, April 2020.

[3] G. Evan, K. Raja S., R. Jason, V. Christian, and W. Brittany, “Accessibility of Voice-Activated Agents for People Who Are Deaf or Hard of Hearing,” Proceedings of CSUN ’19: 34th Annual Assistive Technology Conference Scientific/Research, vol. 7, pp. 144–156, San Diego, 2019.

[4] W. Gilmore et al., “Alexa, Can You See Me?” Making Individual Personal Assistants for the Home Accessible to Deaf Consumers,” Proceedings of ASSETS ’19: 35th Annual Assistive Technology Conference Scientific/Research, vol. 8, pp. 16–31, San Diego, 2020.

[5] J. Berke, [retrieved: June 30, 2021], “Assistive Listening Devices for the Deaf and HOH, https://www.verywellhealth.com/assistive-listening-devices-1046105

[6] A. Matsuda, M. Sugaya, and H. Nakamura, “Luminous Device for the Deaf and Hard of Hearing People,” in Proceedings of the Second International Conference on Human-Agent Interaction (HAI ’14), pp. 201–204, Oct. 2014.

[7] Amazon, [retrieved: June 30, 2021], “Hearing Communicate and stay connected with Alexa.” Available: https://www.amazon.com/b/ref=ods_afe_hop_hp?node=21213721011

[8] F. Bentley et al., “Understanding the long-term use of smart speaker assistants,” Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 2, no. 91, pp 1–24, Sept. 2018.

[9] V. Mande, A. Glasser, B. Dingman, and M. Huenerfauth, “Deaf Users’ Preferences Among Wake-Up Approaches during Sign-Language Interaction with Personal Assistant Devices,” Proceedings of CHI EA ’21: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, no. 370, pp 1–6, May 2021.

[10] SignGenius, [retrieved: June 30, 2021], “Do’s & Don’ts - Getting Attention in the Deaf Community.” Available: https://www.signgenius.com/info-do’s&don’ts.shtml

[11] R. Byron and N. Clifford, “The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places,” Bibliovault OAI Repository, University of Chicago Press, pp. 18–36, 1996.

[12] H. Candello et al., “CUI@CHI: Mapping Grand Challenges for the Conversational User Interface Community,” in Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA ’20), pp. 1–8, NY, USA, April 2020.

[13] C. Pearl, “Designing Voice User Interfaces: Principles of Conversational Experiences,” O’Reilly Media, 2017.

[14] J. P. Bigham, R Kushalnagar, T. K. Huang, J. P. Flores, and S. Savage, “On How Deaf People Might Use Speech to Control Devices,” in Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS ’17), NY, USA, pp. 383–384, Oct. 2017.

[15] A. Glasser, “Automatic Speech Recognition Services: Deaf and Hard-of-Hearing Usability,” in Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA ’19), no. SRC06, New York, USA, pp. 1–6, May 2019.

[16] verizon 2020 [retrieved: June 30, 2021], “Google Pixel 4 - Motions and Gestures.” Available: https://www.verizon.com/support/knowledge-base-228703/

[17] National Institutes of Health (US); Biological Sciences Curriculum Study, “NIH Curriculum Supplement Series,” [Internet], Bethesda (MD): Information about Hearing, Communication, and Understanding, 2007.

[18] M. Henschke, T. Gedeon, and R. Jones, “Touchless Gestural Interaction with Wizard-of-Oz: Analysing User Behaviour,” in Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction (OzCHI ’15), New York, USA, pp. 207–211, Dec. 2015.

[19] A. Glasser, V. Mande, and M. Huenerfauth, “Accessibility for Deaf and Hard of Hearing Users: Sign Language Conversational User Interfaces,” in Proceedings of the 2nd Conference on Conversational User Interfaces (CUI ’20), New York, USA, no. 55, pp. 1–3, July 2020.

[20] A. M. Lieberman, “Attention-getting skills of deaf children using American Sign Language in a preschool classroom,” Applied Psycholinguistics, vol. 36, no. 4, pp. 855–873, July 2016.

[21] A. Pradhan, K. Mehta, and 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 (CHI ’18), New York, USA, no. 459, pp. 1–13.

[22] R. Cui, H. Liu, and C. Zhang, “A deep neural framework for continuous sign language recognition by iterative training,” IEEE Trans. Multimedia, vol. 21, no. 7, pp. 1880–1891, July 2019.

[23] N. Crook, [retrieved: June 30, 2021], “Wizard of Oz testing – a method of testing a system that does not yet exist.” Available: https://www.simpleusability.com/inspiration/2018/08/wizard-of-oz-testing-a-method-of-testing-a-system-that-does-not-yet-exist/

[24] N. Fraser and N. Gilbert, “Simulating speech systems. Computer Speech & Language,” vol. 5, pp. 81–99, Jan. 1991.

[25] J. Peirce and J. R. Gray, “Simpson, S. et al. PsychoPy2: Experiments in behavior made easy,” Behavior Research Methods, vol. 51, pp. 195–203, Feb. 2019.

[26] The Language Archive, [retrieved: June 30, 2021], “ELAN.” Available: https://archive.mpi.nl/tla/elan

[27] J. Brooke, “SUS: A quick and dirty usability scale,” Usability Eval. Ind., vol. 189, pp. 1–7, Nov. 1995.

[28] D. Bavelier et al., “Visual attention to the periphery is enhanced in congenitally deaf individuals,” The Journal of Neuroscience: The official journal of the Society for Neuroscience, vol. 20, no. 17, pp. 1–8, Sept. 2000.

[29] Developer documentationamazon alexa, [retrieved: June 30, 2021], “Alexa Design Guide.” Available: https://developer.amazon.com/en-GB/docs/alexa/alexa-design/available.html