Exploring Think-aloud Method with Deaf and Hard of Hearing College Students

Si Chen
sic3@illiois.edu
Information Sciences, University of Illinois at Urbana–Champaign
Champaign, Illinois, USA

Desirée Kirst
deskirst@gmail.com
Gallaudet University
Washington, District of Columbia, USA

Qi Wang
qi.wang@gallaudet.edu
Gallaudet University
Washington, District of Columbia, USA

Yun Huang
yunhuang@illinois.edu
Information Sciences, University of Illinois at Urbana-Champaign
Champaign, Illinois, USA

ABSTRACT

The think-aloud protocol is an effective method frequently used by designers and researchers to understand how users interact with computing systems. However, there is limited research on the use of this method with deaf and hard of hearing (DHH) populations, especially in virtual settings. In this paper, we investigate the behaviors of DHH participants in virtual think-aloud sessions to better understand the challenges of conducting this type of research with this population. We conducted twelve virtual think-aloud sessions with DHH participants using Zoom, and we gathered feedback from surveys, interviews, and observations. Our results identified DHH behaviors leading to a lack of clarity in think-aloud data, such as asynchrony between signing and navigating the interfaces, as well as the use of visual descriptive signs instead of explicit terminology to ambiguously refer to interface components. Based on our findings, we provide methodological and design implications to help researchers effectively carry out virtual think-aloud studies with DHH participants (e.g., when and how to prompt for clarification).

CCS CONCEPTS

• **Human-centered computing → Accessibility design and evaluation methods**; • **Applied computing → E-learning**.

KEYWORDS

Accessibility, Reflection, Think-aloud, d/Deaf and Hard-of-Hearing, Sign Language

ACM Reference Format:
Si Chen, Desirée Kirst, Qi Wang, and Yun Huang. 2023. Exploring Think-aloud Method with Deaf and Hard of Hearing College Students. In *Designing Interactive Systems Conference (DIS '23), July 10–14, 2023, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3563657.3595980

DIS '23, July 10–14, 2023, Pittsburgh, PA, USA

ACM ISBN 978-1-4503-9893-0/23/07…$15.00
https://doi.org/10.1145/3563657.3595980

1 INTRODUCTION

The think-aloud protocol is a widely used technique in user experience studies, where participants verbalize their thoughts and actions while interacting with a system or product. While this method has proven effective in understanding the behaviors and preferences of individuals with typical hearing abilities [18, 37], there is a lack of research on its application in involving deaf and hard of hearing (DHH) participants [33, 34]. This gap is particularly notable in virtual studies [14].

A systematic review of accessibility research conducted over a span of 26 years reveals that approximately 20% of user studies have been conducted virtually [24]. During COVID-19, traditional offline design practices are impractical due to limitations on social contact. As the current situation becomes the "new normal," it is necessary for design methods to adapt and incorporate virtual user studies [23, 27]. For DHH participants, video conferencing platforms like Zoom offer new communication methods, such as automated captions and typing, which have the potential to transform the paradigm of user study sessions.

Based on the United States Census [31], an estimated 11.5 million Americans have various degrees of hearing differences, which make up 3.5% of the population. The degree of hearing loss can have an impact on an individual's speaking abilities. Individuals with severe to profound hearing loss may have more significant challenges in speech production. They may have difficulty accurately perceiving and producing certain speech sounds, resulting in speech that may be less intelligible to those who are not familiar with their communication style. Designing and testing products and systems to be fully accessible for this population with diverse communicating preferences is not only socially responsible for designers and researchers, but it is also important in the HCI field, as it advances the universal design concept and practice and fills a knowledge gap about the viability of think-aloud testing methods with DHH population. Thus, understanding the use of the think-aloud protocol with the DHH population can enable designers and researchers to gain experience and utilize this method. It is critical to consider how research methods can be made accessible for users and researchers with disabilities, as accessibility research becomes more widespread and barriers to conducting and participating in research are pervasive [24, 25].

In this study, we conducted virtual think-aloud sessions with twelve DHH college students using Zoom to examine their interactions and preferences in a post-learning dashboard that tracks and visualizes users' emotions as shown in Figure 2. In this manuscript, by observing and analyzing the behaviors of DHH participants in virtual think-aloud sessions, we aimed to better understand the challenges of conducting this type of research with this population. Our findings revealed that participants leveraged multiple communication modes concurrently and sequentially, due to their multi-lingual background, to express themselves. We also observed two behaviors in DHH's think-aloud sessions, such as interacting and explaining sequentially instead of concurrently and using visually descriptive signs instead of labeling (or spelling out the names) interface components. Our primary conclusion is methodological with recommendations in tool designs that support virtual think-aloud with DHH. We argue that hearing researchers need to understand DHH's communication in multiple modes, and the temporal and spatial (in)dependencies between multiple modes must be taken into account, as illustrated in our quote template in Figure 2.

2 RELATED WORKS

2.1 Think-aloud Protocol and Diverse Communication Preferences among Deaf and Hard of Hearing (DHH) Users

The think-aloud protocol is a user experience studying approach to understanding how users interact with a product or system. This method involves having users verbalize their thoughts and actions during their interactions, which helps researchers and designers gain valuable insights into user behavior and preferences [1, 15]. The user's interactions with the system are usually recorded along with the think-aloud for later analysis that can reveal additional user actions and behavioral rationales possibly missed during observations alone. Researchers can use these participant materials to enhance their observation details and notes as well as generate new questions that may gain further understanding of the user and system. Think-aloud has been adopted in accessibility research since it provides a robust, low-cost approach to understanding user behaviors. No special equipment is needed; researchers/designers can sit next to a user and take notes as they converse. For instance, the think-aloud protocol has been used to examine how older adults comprehend visualizations [16] and how they use mobile apps [19]. It has been used to investigate how blind and low-vision users interact with digital library interfaces [43]. Given the growing importance of accessibility research [8, 24, 25], there is a need to assess the applicability and protocols that widen design/research methods for users with disabilities. Our study aims to investigate the effectiveness think-aloud sessions with DHH users.

However, there is a lack of understanding of how to effectively apply the think-aloud protocol within DHH users, especially due to the variability of the DHH users' preferences. These users' hearing statuses vary; thus, they often employ multiple languages and modes of communication (e.g., simultaneous and consecutive use of signed, spoken, written, semiotic, and interpreted channels) to express themselves during user studies [14, 20]. The limited studies of the think-aloud protocol with DHH users have mainly considered American Sign Language (ASL). For example, [33, 34] think-aloud studies included deaf users who only communicated through ASL. Signed languages, such as ASL, are complete, natural languages, distinct from spoken languages with respect to phonology, morphology, syntax, and vocabulary [41]. However, some deaf individuals, ASL-English bilinguals, may choose to use a mix of their languages, producing words and signs or both simultaneously, which blends their communication modes.

The previous studies cited did not include hard of hearing (HoH) users who are also ASL-English bilinguals, nor did they allow the use of other communication methods. To address the deficiency studies that include various communication modes and languages for DHH participants in the think-aloud protocol, we recruited participants who self-identified their hearing status as DHH and who were comfortable communicating in ASL and English. Participants were encouraged to communicate in either/both languages supplemented with any additional communication modes they preferred. Our findings reflect the advantages and challenges such communication modes provide. Our work also expands existing exploratory virtual think-aloud work led by DHH researchers [14].

2.2 Virtual Studies In Accessibility Research

An extensive 26-year review of accessibility research reveals that virtual user studies make up around 20 %, similar to the 27 % occupied by in-lab settings [24]. Virtual user studies are an increasingly popular form of research due to their flexibility in scheduling, especially as a result of the COVID-19 pandemic [23, 27]. To ensure that participants with disabilities have access to the study materials throughout different stages, [25] identified four key dimensions (communication, materials, space, and time) for accessibility researchers to plan for and a common strategy of "anticipating with adjustments" among participants. Research has demonstrated that access must be considered and implemented throughout all stages of the research process. For example, [16] set up Zoom tutorials prior to formal study sessions to help elderly adults become familiar with the tools used in the study. In order to promote accessibility in user studies, researchers are developing guidelines to make access to labor throughout user studies visible in papers and a core consideration of study planning. Neate et al. [29] discussed the use of physical whiteboards and physical objects around disabled users to facilitate communication between researchers and participants while taking into account privacy concerns.

Focusing on the DHH users, virtual platforms like Zoom offer new communication options, such as automatic captions and typing, which can potentially enhance the accessibility of virtual user study sessions. Nonetheless, previous research has also highlighted both the challenges and adaptions that virtual platforms introduce for DHH population. For example, Seita et al. [36] proposed guidelines for conducting online co-design workshops with DHH and hearing participants to explore auto-generated captions for videoconferencing technology. They suggested providing diverse communication modalities, encouraging engagement, and considering participant characteristics. Other research has identified the challenges DHH users may face with virtual settings [24–26]. The impact of inaccessible videoconferencing systems on research facilitators such as ASL interpreters providing real-time translation service should also be taken into account in virtual studies with DHH users [35].

Meanwhile, virtual think-aloud with DHH users has not been extensively explored. In an attempt to contribute to inclusive methodologies in HCI, our research fills this gap by identifying challenges for hearing researchers encountered during the process together with DHH participants' own think-aloud quotes. In existing offline think-aloud studies with DHH participants, the hearing researcher typically reminds the participant to "keep talking" through an interpreter who provides simultaneous communication, which can lead to slight delays as the interpreter must wait for the signer's sentence to be finished before rendering an accurate equivalent and vice versa [34]. However, this study focused on usability feedback and did not explore the user experience of DHH participants, which goes beyond identifying usability problems. Usability emphasizes task efficiency and effectiveness, while user experience encompasses broader aspects of satisfaction and perception. We also listed challenges we as hearing researchers encountered when facilitating the think-aloud session.

3 METHOD

We recruited twelve DHH college students from a university in the United States and conducted virtual think-aloud sessions with each participant using Zoom. Each session, which was video-recorded, lasted for around 1.5 hours, and the participants were paid $15 per hour for their participation. Prior to the think-aloud, participants were asked to watch a 15-minute video while their emotions were recognized. Next, the participants completed a think-aloud protocol while using a dashboard that tracked their learning emotions for reflections for 10-15 minutes. Afterward, an exit survey and interview were conducted to gain insight into their experience with the dashboard, the think-aloud protocol, and regarding their backgrounds. The exit survey and interview took around 45 minutes.

In our research, the DHH participants communicated through American Sign Language (ASL), written English, and spoken English with support from Zoom's auto-captioning. The participants were given the freedom to choose the mode(s) of communication that they felt the most at ease using to express themselves. The data collected from all sessions were then integrated to form the research findings, and all modes used were taken into account. Consent was obtained from each participant, and only those who gave additional consent were included in the examples in this manuscript that display names, faces, likenesses without masks, and direct quotes. This study has been approved by the university's Institutional Review Board (IRB), an independent committee responsible for ensuring the protection of human participants in research studies by reviewing and approving research proposals to ensure they meet ethical and regulatory standards.

3.1 DHH Participant Recruitment

Twelve DHH undergrad students (self-identified: four female and eight male) were recruited from a university in the United States using word-of-mouth and snowballing; students must learn both ASL and English after being admitted. Our virtual think-aloud sessions reflected this diversity and included both domestic (first language is English & or ASL) and international students (first language is neither English nor ASL) as participants. Participant background information is presented in Figure 1 and was provided through responses to multiple-choice and text-based short-answer questions during the exit survey and from any accompanying explanations during the survey process. Four self-identified as hard-of-hearing (HoH) students, and eight self-identified as deaf students. The average self-reported age was 25 years old.

3.2 Virtual Think-aloud Study Design

In this study, we adapted the 'Gestural Think-Aloud Protocol' for DHH users, known as 'self-sign'[14, 33, 34], to use a think-aloud protocol. This protocol requires participants to express their thoughts at any time or, at the very least, before they make any moves or use the mouse. Studies conducted on this protocol [14, 33, 34] show that 'self-sign' can be an effective usability evaluation method and that 'self-sign' think-aloud protocols for DHH participants are comparable to those conducted with hearing participants in terms of feedback quantity. To prepare for the think-aloud, participants were given written English instructions about the think-aloud protocol, which included a preview image of the interface and the purpose of the interface: reflection. A pre-recorded video demonstration was provided, showing the protocol being used in a different system than ours and being conducted by a hearing individual using spoken English (with captions).

To ensure the participants' understanding of the whole process, two hearing researchers, fluent in ASL and English, led the study. Both researchers were present to answer any questions asked before and during the think-aloud sessions. For each session, one of the two fluent researchers led the session. Prompting and clarifying questions are often necessary components during think-aloud studies. In prior think-aloud studies with DHH participants offline, researchers typically prompted users to keep expressing themselves in ASL via a sign language interpreter (real-time translating ASL and English) or through a tap on the participant's shoulder if they observed more than 10 seconds of silence (a few moves without any utterances or observable thought behaviors such as facial expressions or head nodding) [5, 33]. However, some of the DHH user behaviors that are noticeable to researchers offline may not be as noticeable to researchers online, especially considering DHH have varied communication modes [14]. Therefore, in our study, we slightly changed the time period to over 1 minute, and unless the participants directly asked questions, the researcher only provided visual feedback through head nods to encourage participants to "keep talking" or "keep signing" in a less interrupting way.

Non-signing hearing researcher(s) unfamiliar with ASL took observation notes and managed Zoom settings and recordings without intervening. All researchers turned on their cameras at the beginning of the session, clarified their role in the study, and then the observing researcher(s) turned off their cameras during the rest of the think-aloud session. This reduced the number of video thumbnails overlaying and obstructing the Zoom interface; further, it alleviated the complexities that arose from multiple people turn-taking in small group settings online, which was not the focus of this study.

PID	Gender	Hearing Status	Age	Age ASL Acquired	Formative Language(s)	Major	Zoom Experience
P1	Male	Deaf	30	8	ASL & English	Business	Intermediate
P2	Male	HoH	24	19	English	Business	Advanced
P3	Female	Deaf	30	0-12 months	ASL & English	Business	Advanced
P4	Female	Deaf	33	23	Saudi Sign Language	Business	Intermediate
P5	Male	Deaf	22	19	Chinese	Business	Advanced
P6	Male	Deaf	20	2	ASL	Business	Intermediate
P7	Female	Deaf	23	0-12 months	ASL & English	Business	Intermediate
P8	Female	Deaf	25	11	English	Undecided	Advanced
P9	Male	Deaf	30	6	ASL	Business	Intermediate
P10	Male	HoH	22	20	Arabic & English	Graphic Design	Advanced
P11	Male	HoH	21	13	English	Physical Education	Intermediate
P12	Male	HoH	19	18	English	Business	Advanced

Figure 1: All participants are undergraduates who provided background information through multiple-choice and text-based survey questions as well as some accompanying spoken English and signed ASL explanations. HoH = Hard of Hearing

3.2.1 Materials. The think-aloud protocol employed a post-learning dashboard that plotted the emotions recognized from the participants' facial movements while they watched an educational video on "Augmented Reality." The participants' video streams were captured while watching the stimulus video. Afterward, participants were asked to provide comments on the video, which were plotted in the dashboard to provide an additional layer of reflection. Then in the resulting dashboards in Figure 2, learners were able to reflect on their own emotions recognized from facial movements displayed - ⓒ using legend ⓕ, as well as their own comments on the video as in ⓓ. The height of the data points demonstrated the intensity of the emotions, and the color showed the recognized emotions, similar to [9]. The technical details for emotions recognized from facial movements involve feeding video streams of participants into an automatic emotion recognition model [13]. This model returns recognized emotions and their intensity for each timestamp, which is then plotted and labeled 'Your Emotion Intensity' in Figure ⓕ. By clicking on any timestamp, a still of the participant's facial expression captured will appear in the top right corner of the screen ⓔ. In addition to displaying their own emotions, the dashboard also showed the emotions and comments from their peers. ⓑ and ⓖ showed the peers' positivity and intensity, along with the legend for each. When participants hovered their cursor over either the up and down arrow icons on the vertical colored bars, they could view a selection of the comments from their peers as shown in ⓖ.

The user interaction of the tool used is similar to [9] for two specific reasons. First, the tool served as a research prototype unfamiliar to all participants. This approach ensured that participants began the study with equal levels of familiarity with the software. Second, it is a novel AI-based system that had previously undergone evaluation solely with hearing participants. By conducting user studies with DHH individuals, we gained new insights into the inclusive design and ethical considerations of the AI system when applied to diverse populations.

3.2.2 Exit Interview and Survey. During the think-aloud session, participants are given the space to think and reflect on their own, and after they have finished talking and clearly indicate 'finish,' researchers move on to the exit survey and interview, the final two parts of the user sessions. The exit survey is conducted to collect more background information, such as their English and ASL abilities and prior Zoom experience. Lastly, pre-designed questions were asked to gain further insight into the participants' behaviors and experiences with the interface design and think-aloud process.

3.2.3 Researcher Positional Statement. Our research team is composed of hearing researchers who specialize in conducting think-aloud sessions and deaf students participated in the study design, data collection, and data analysis process. Among these authors, two are interpreters of both ASL and English, and one of them has worked in interpreting for over 15 years. The other two authors are fluent in English and possess basic ASL knowledge, while the remaining three researchers are fluent in English and lack any ASL knowledge.

3.3 Analysis of Data

This research utilized thematic discourse analysis to analyze the data collected from Zoom recordings of video thumbnails, screen shares, and chatrooms during the entire study session (e.g., preparing for think-aloud, conducting the think-aloud, interview). Thematic analysis is a method for identifying, analyzing, and reporting repeated patterns of meaning across a dataset. This technique involves repetitive coding and re-coding of text and aggregation of these codes into larger themes [5]. Discourse analysis is a method of analyzing language use (written, spoken, or signed) to understand how it works in terms of meaning, structure, and context [6, 40]. To support our study, we employed the concept of 'repertoire assemblage,' which expands translanguaging theory and views languages as fluid resources that form 'semiotic assemblages.' This combination of thematic analysis along with discourse analysis has been

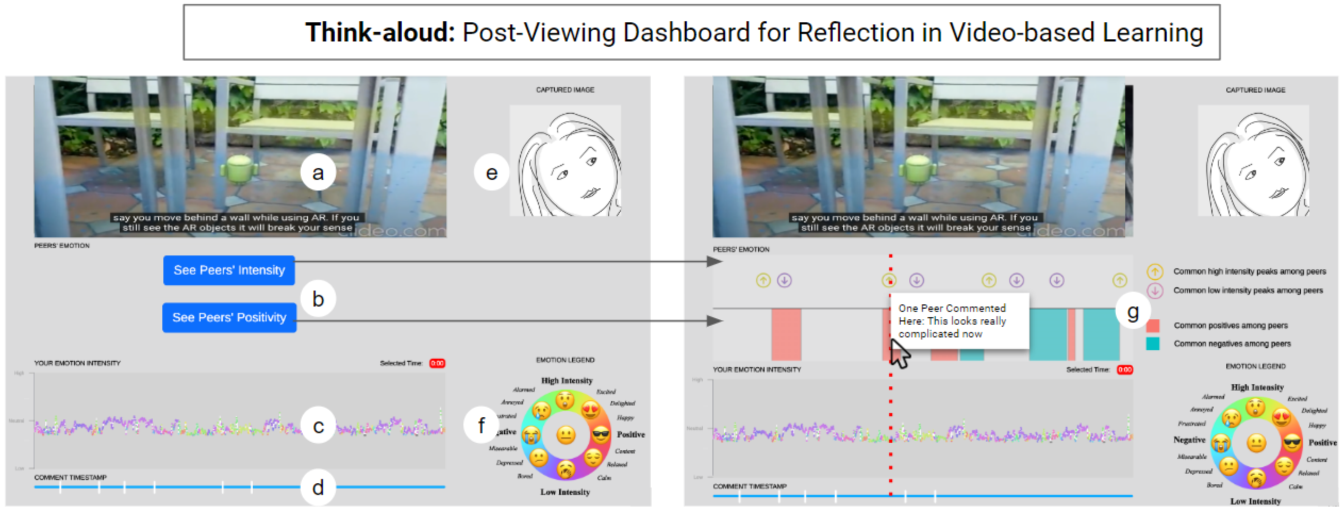

Figure 2: Material: Dashboard Used for Post-Learning Think-aloud: User reflects on their own emotions with peers' emotions on the post-learning dashboard. Post-learning dashboard has seven areas: ⓐ displaying the video recording, ⓑ displaying peers' emotions positivity and intensity, ⓒ visualizing own time-based AI-recognized emotions collected while the participant was watching the video, ⓓ displaying own comments, ⓔ providing video selfies which are collected while the participant was watching the video in at the selected timestamp (researcher illustrated to preserve privacy), ⓕ providing a legend for time-based own emotions, and ⓖ providing a legend for peers' emotions (arrows for presenting high/low intensity peaks, bars for presenting positive/negative segments). Own and peers' emotional data are temporally aligned for easy comparison. P5's emotion data is visualized in area ⓒ.

used previously to gain insights into user behavior [4, 39]. Our findings mainly focus on direct think-aloud quotes and their discourse and occasionally provide survey and interview explanations that further expand upon these quotes. Since facial expressions are an essential part of ASL grammar and vital to DHH individuals' communication, we used examples from participants who consented to have their faces unmasked. This allows for their expressions to be more accurately portrayed and thus authentically represent their experiences.

Three authors analyzed each session following a thematic coding process [5]. After each think-aloud session, the researchers discussed the observed themes and each rewatched the Zoom recording independently to develop an initial codebook. Through multiple sessions, they iterated on the codes until a consensus was reached. Examples of the codes included participants' language and communication modes used, switching between language and communication modes, and the researcher's response. First, all recordings were mainly analyzed by one researcher: one coder fluent in ASL and English transcribed all recordings into a text-based format as shown in Figure 3 (Left) then segmented and coded the recordings. For each session, the total word count was around 2,000, including both the researcher's and participant's utterances. Other team members with different ASL skills reviewed those codes, suggesting additions and questioning unclear codes.

One of the major themes related to all other themes was the analysis and presentation of the quotes given the multi-mode communications utilized by DHH. After considering various methods for presenting think-aloud quotes to more effectively analyze the combination of different modes, we formatted the *Multi-Mode Analysis* templates displayed in Figures 3 (Right). This template provides opportunities to reveal the (in)dependence and spatial, and temporal relationship between different modes. Each column represents one communication mode along the user study timeline. Using the new template, another researcher fluent in English with limited ASL skills coded approximately 30% of the data on her own to check for consensus with the first researcher. Upon resolution of any differences, the research team created higher-level themes presented in the findings section. The quotes presented in the rest of the manuscript follow this format of laying multiple modes along the video timeline but note that some may not present all available modes of communication in order to best illustrate the theme.

4 FINDINGS

4.1 Expressions in Multi-Mode as a Result of Multi-Lingual Backgrounds

During the think-aloud sessions, the participants expressed themselves through multiple modes of communication, including spoken English captured via Zoom audio and signed modes, which were recorded by Zoom video thumbnail. Some participants often blended multiple modes and languages simultaneously. For example, P11 primarily spoke in English while producing accompanying ASL signs. These ASL signs were conceptually accurate on their own, their grammar mirrored English patterns and therefore did not follow ASL syntax as illustrated in Figure 4.

The majority of deaf participants exclusively used ASL during the think-aloud activity. One deaf participant (P6) used ASL during

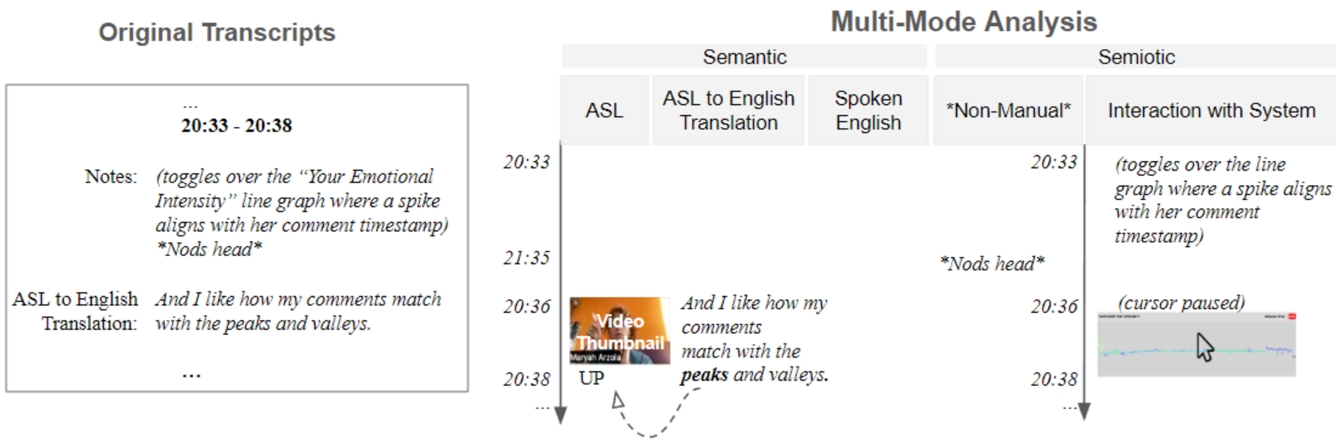

Figure 3: A Multi-Mode Format was developed to reflect the diverse communication modes used by the participants and used for analysis. The semantic columns include the language(s) used by participants while the semiotic columns include other behaviors that provided insight into the participants' actions, such as body movements, head nods, etc., as well as interactions with the interface such as moving the mouse and hovering the cursor over specific features. Capitalized English is the translation of the exact ASL sign following ASL grammar. To notice: our manuscript focuses on understanding how DHH individuals' behaviors impacted the execution of a think-aloud protocol. Therefore, the quotes included in our manuscript encompass both the think-aloud process and the preparation for it. From left to right, the sequence implies the leading language used by the participant.

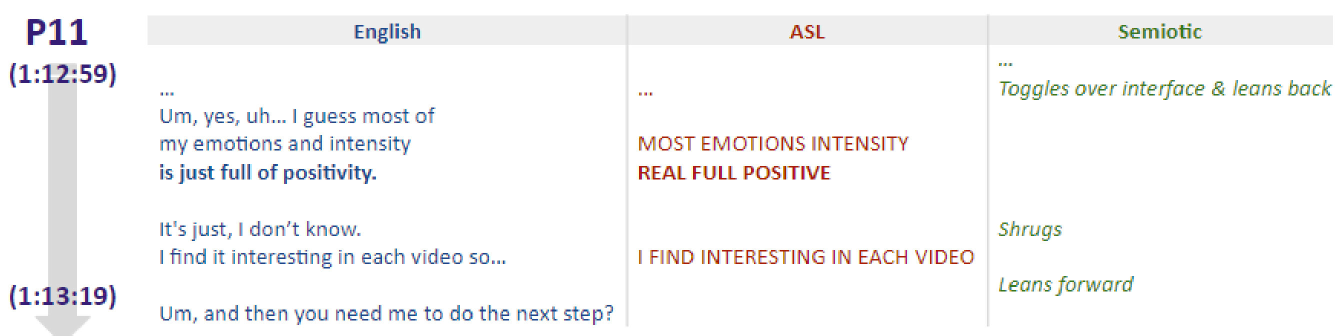

Figure 4: Participants who grew up using English in childhood also tended to use it during the think-aloud. To express their thoughts more clearly, they incorporated ASL signs while still following English grammar and syntax. Since these two languages use different modes (signed and spoken), they can be blended together. The example above shows P11's spoken sentence about the number of positive emotions he observed, which he emphasized by using the ASL sign "REAL." Meanwhile, spoken English also expresses other information such as "I guess...".

the think-aloud but a combination of both English and ASL during the preparation of the think-aloud (e.g., during instructions). In contrast, the four HoH participants made much more varied linguistic choices on how they wanted to express themselves. Three HoH participants made a single linguistic choice, exclusively using it throughout each portion of the study: P2 chose to only use spoken English during the study, while P11 used a combination of both English and ASL, and P10 only expressed himself through ASL. The other HoH participant (P12) alternated between linguistic choices, using ASL during the think-aloud but occasionally using a combination of both English and ASL during other portions of the study. The participants' linguistic decisions predominantly reflect the formative languages of their childhood.

Translation between spoken languages' vocabularies does not always correspond one-to-one, nor does it with sign language corpora. The same is true when signing ASL and speaking English simultaneously. For example, P11's ASL signs were conceptually accurate, but the syntax of his signs mirrored English patterns; some signs did not follow ASL morphology or syntax, as illustrated in the bolded portion of the quote in Figure 4. In this case, the ASL sign "REAL" means 'truly' or 'actually,' whereas the English use of the word "just" emphasizes the intensity of the spoken reflection. The

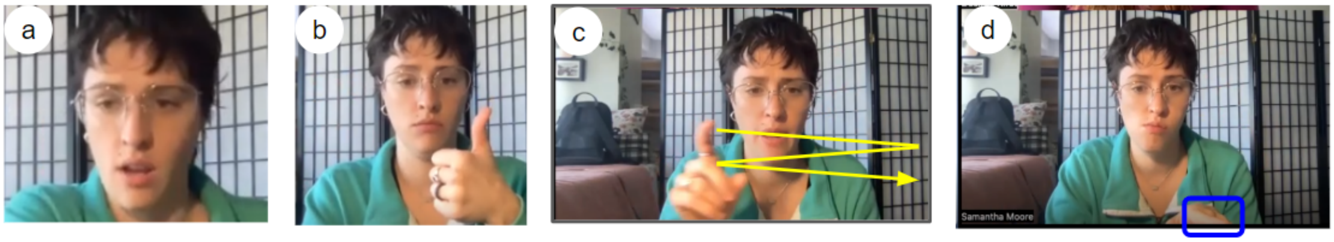

Figure 5: Examples of semiotic behaviors used by the participants to demonstrate active participation. These behaviors included using mouth shapes such as 'okay' ⓐ, thumbs up ⓑ, and moving a finger to indicate reading between the lines, yellow lines are finger-moving trajectory ⓒ. Some of these behaviors pointing at the screen during researchers' introductions to the study were outside of the camera ⓓ (blue box). (P3)

results suggest different modes contain overlapping expressions but are not exactly the same.

Semantic information was used to explain the interaction with the system and aid the think-aloud process. For instance, in Figure 4, the participant said "my emotions and intensity is just full of positive" to express his thoughts on his emotion line graph. After he was finished, he asked "and then you need me to do the next step" to signal to the researcher that he was done. Further examples of using semantic information to ask for clarification and indicate completion to the researcher can be seen in Figure 6, Figure 7, and Figure 8.

In addition to language preferences, the Zoom recordings also captured various semiotic information via video thumbnail. Semiotic information includes additional details that can be gleaned from gestures and facial expressions, which mainly occurred before and after the think-aloud self-sign to indicate that the researcher is ready/ done to facilitate the study. Examples in Figure 5 show P3 demonstrating mouthing 'okay' ⓐ, displaying a thumbs-up gesture ⓑ, and pointing with an extended index finger ⓒ and ⓓ. The most commonly observed behavior among the participants was nodding of the head, which could indicate a range of meanings, such as engagement in the conversation, attentiveness, agreement, or cognitive processing. An additional example of a semiotic behavior is illustrated by P1 in Figure 6.

Besides the video feeds of the participants, their mouse movements were also collected via screen sharing, providing further insight into their interactions with the system. Examples of this can be seen in Figure 7 and Figure 9. While some of the gestures used in face-to-face communication may not be as clear during virtual sessions, the participant in this example pointed to the screen at 57:55 and then moved the cursor to the button she was referring to and pointed to the screen again (58:56, 59:47) to avoid confusion. Such observation suggests, different modes of semiotic information were utilized together to improve expression clarity.

4.2 Asynchronicity Between Signing and Navigating the Interface

The use of multiple modes of communication by multi-lingual individuals in user study sessions has resulted in miscommunication and created obstacles to successful user studies. For example, DHH participants needed to pause interacting with the system (semiotic information) in order to use their hand(s) to sign their thoughts and observations (semantic information), causing asynchronicity issues and making it difficult for them to fully and accurately explain their experience with the interface

Therefore, the participants who did not use spoken English (11 out of 12 participants) as the main language in the think-aloud session experienced a delay between their interaction with the system and their explanations of the interactions. Some participants naturally pointed at the screen instead of using the mouse to navigate the interface, as shown in Figure 5 (e). Hearing participants can perform both actions simultaneously, but the signing participants must switch between the actions resulting in a delay between the action and the thought process. The delay is caused by the switch between actions, as both require the use of the hand(s). Occasionally, these pointing behaviors, a common form of pronominalization (e.g. "this") in sign languages, became unclear for observing researchers because the virtual settings made it difficult for the researcher(s) to recognize what was being pointed at. The delay continued until the participants expounded on what they were referencing with their point. For example, in Figure 7, P7 communicated only in ASL. In this instance, P7 looked at the interface and pointed at the screen (timestamp 57:55), holding the pointing gesture until the researcher responded (57:56). The cursor remained over the 'See Peers' Intensity' button, but P7 did sign a sentence. Later, P7 hovered over this button again, pointed at the screen, signed "CLICK," and pointed again (58:55-58:56). The researcher responded affirmatively, and P7 signed "CLICK" again and pointed a third time before clicking the button (timestamp 58:55). This similar pattern was repeated with the "See Peers' Positivity' button (59:47-59:48). This switching between semiotic behaviors (using the cursor and gesturing) and signing sentences ("CLICK") delayed responses. Furthermore, some information can be cut off by the video camera leading to misunderstood behaviors. For example, after reading a peer's comment (1:00:01), P7's hand moved below the camera frame. The researchers had to pay particular attention to the semiotic environment in order to understand what P7 was referred to in her signed sentence.

Due to the asynchronicity between interactions and explanations, many participants clicked on and paused their cursor movements

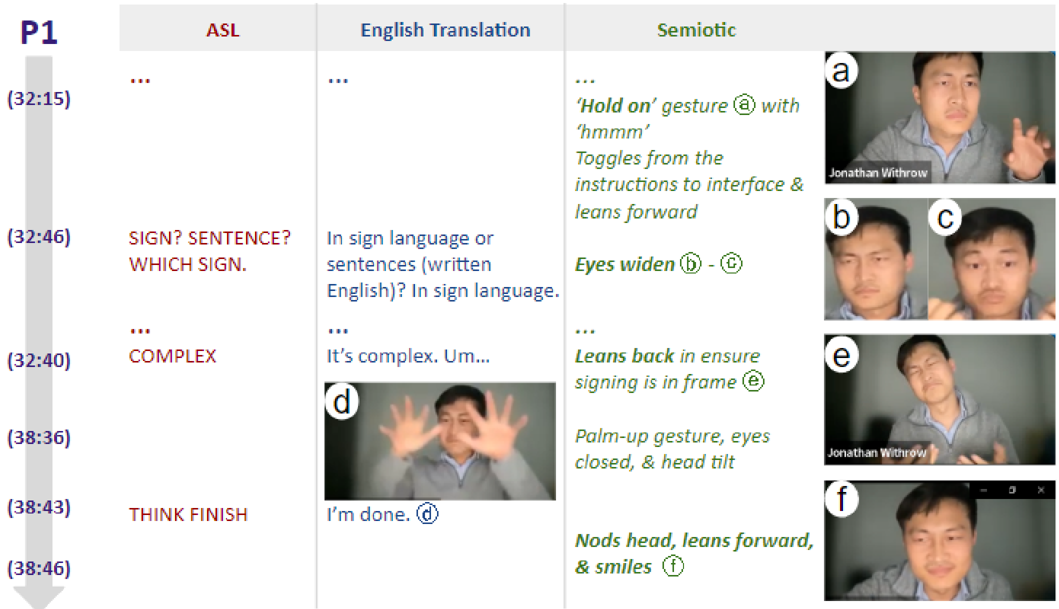

Figure 6: During the think-aloud sessions, the majority of the deaf participants used ASL as their only language. They supplemented their communication with additional semiotic behaviors to both express their thoughts further and to gain clarification from researchers. Examples of semiotic behaviors used by the participants to capture researchers' attention, included using gestures for 'hold on' ⓐ, enlarging their eyes and asking if the ASL was allowed ⓑ and ⓒ, leaning backward to ensure signs get captured by camera ⓓ, leaning forward and smiling to capture the researcher's attention and indicating he was finished with the think-aloud ⓔ. (P1). All faces and quotes shown in the manuscript are given consent by our participants.

over multiple places on the interface. During some pauses, participants would recall their previous actions and then explain those behaviors. However, often they would interact with more interface components than they would give explanations for. This may be due to the inconvenience of not being able to sign and interact with the interface simultaneously. The think-aloud process had to be completed sequentially for signing participants. Thus, participants' interactions tended to include more actions than their utterances conveyed, excluding some of their previous behaviors and thought processes. For example, in Figure 8, P4 only interacted with the tool for four minutes (22:07-26:46). After three minutes without any explanation of her thoughts or behaviors, the researcher gave an elicitation prompt(24:57-24:58), "While you are playing, what are you thinking?" The subsequent sentence by P4 summarized her general experience but did not address her thoughts about the specific features she had been interacting with. Within the next minute, P4 signed that she had finished exploring the interface. Clarity on P4's interactions was not discussed until the interview portion of the study because P4 talked in big chunks, and the researchers avoided interrupting and cutting off thinking processes or leading the participant's remarks. This created a challenge when deciding when to ask for clarification.

Participants in our study used 'silences' as a source of chronemic to engage the researcher and receive more responses from the researcher without explicitly asking a question - the use of time in communication. According to the study design adopted from prior research, researchers actively nodded their heads more frequently. However, this was not sufficient for some participants, as P4 and P7 resumed their interactions after 45 and 17 seconds of silence, respectively, as well as multiple pauses afterward in an effort to engage the researcher and receive more responses. This indicates that the current way for researchers to intervene may not be sufficient and natural enough for participants, and suggests that other methods of engagement should be explored. Another confirming quote from P1. When the researcher asked the participant to begin the think-aloud process, the participant paused and expressed an utterance that indicated they were treating it as a conversation rather than a form of self-talk." Oh. I want to think through my feelings before answering and make sure I don't jump into it without figuring out my thoughts and analyzing my feelings first. Once I feel confident in that, I'll go ahead and answer. I don't want to jump ahead and just give an answer without thinking it through. I want to slow down and make sure it makes sense before going forward. I wanna think about it."

4.3 Using Visually Descriptive Signs Over Explicit Terminology

Our study's participants generally did not use the labels and terminology explicitly written on the dashboard interface. Instead, they used a variety of signs and visual descriptions. However, because of the virtual nature of the think-aloud sessions, the referents of their descriptions were difficult to ascertain, making it hard for

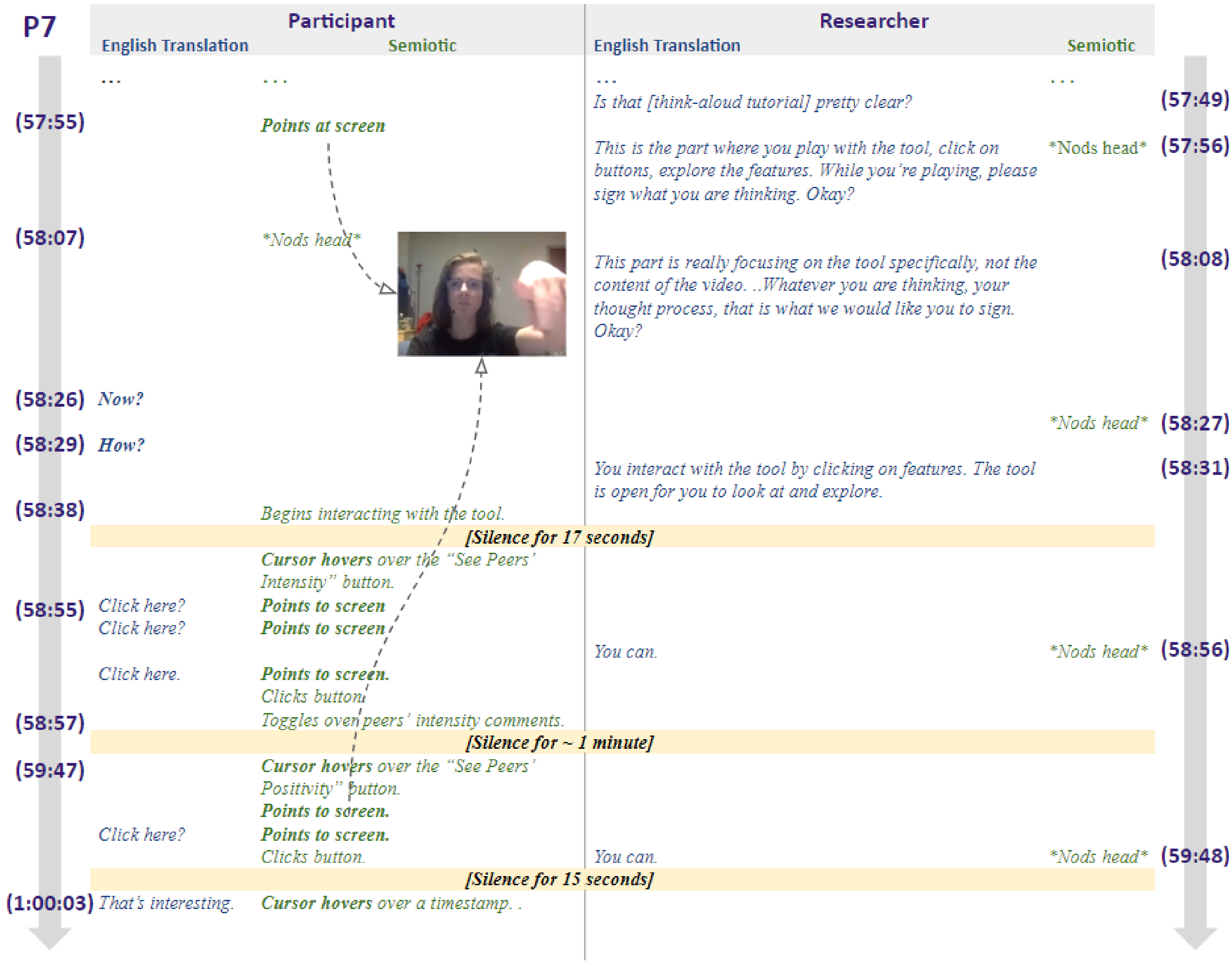

Figure 7: Some of the gestures used in face-to-face communication are not as clear during virtual sessions, as seen when the participant in this example pointed to the screen at 57:55. To help avoid confusion, the participant moved the cursor to the button she was referring to and pointed to the screen again (58:59). Additionally, instead of watching the tutorial, the participant asked for step-by-step guidance from the researcher, such as when her silence at 56:50 was not responded to and she actively asked "now?" and "how" for further guidance. 'Silence' refers to no semantic or semiotic information being observed and the Zoom looks like it 'freezes.'

the researchers to understand what part of the interface was exactly being referred to. For example, the synthesized information from previous participants was used to generate two charts: 'Peers' Intensity' and 'Peers' Positivity' along with their corresponding legends (Figure 2 - ⓑ and ⓖ). The DHH participants referred to these charts and the information therein in a variety of ways.

For example, P8 signed the transcribed sentence, "I like how my comments match the time," and then immediately she pointed up and down. Figure 9 [a1] shows P8 signing UP. The pointing is vague as it could refer to several components on the interface. The only qualifier is the sign "TIME," which also could refer to several components (e.g., the timeline on the video learning material, the Comment Timestamp bar, etc.) At that moment, it was notable that the participant's cursor was hovering over a peak in her emotional intensity line graph Figure 9 [a0]. Understanding the implied meaning but wishing to clarify, the researcher asked, "Do you mean how they [the peak on the line graph and tick mark on the timestamp bar] align?" P8 confirmed with a head nod. In this case, P8 used the signs "UP" and "DOWN" to reference the intensity shown in her graph. See the signs UP and DOWN simultaneously demonstrated by a researcher in Figure 9 [a2].

These same signs, "UP" and "DOWN," were used by other participants in other contexts, such as referring to the Peers' Intensity chart. Once the button is clicked, the chart appears, showing the

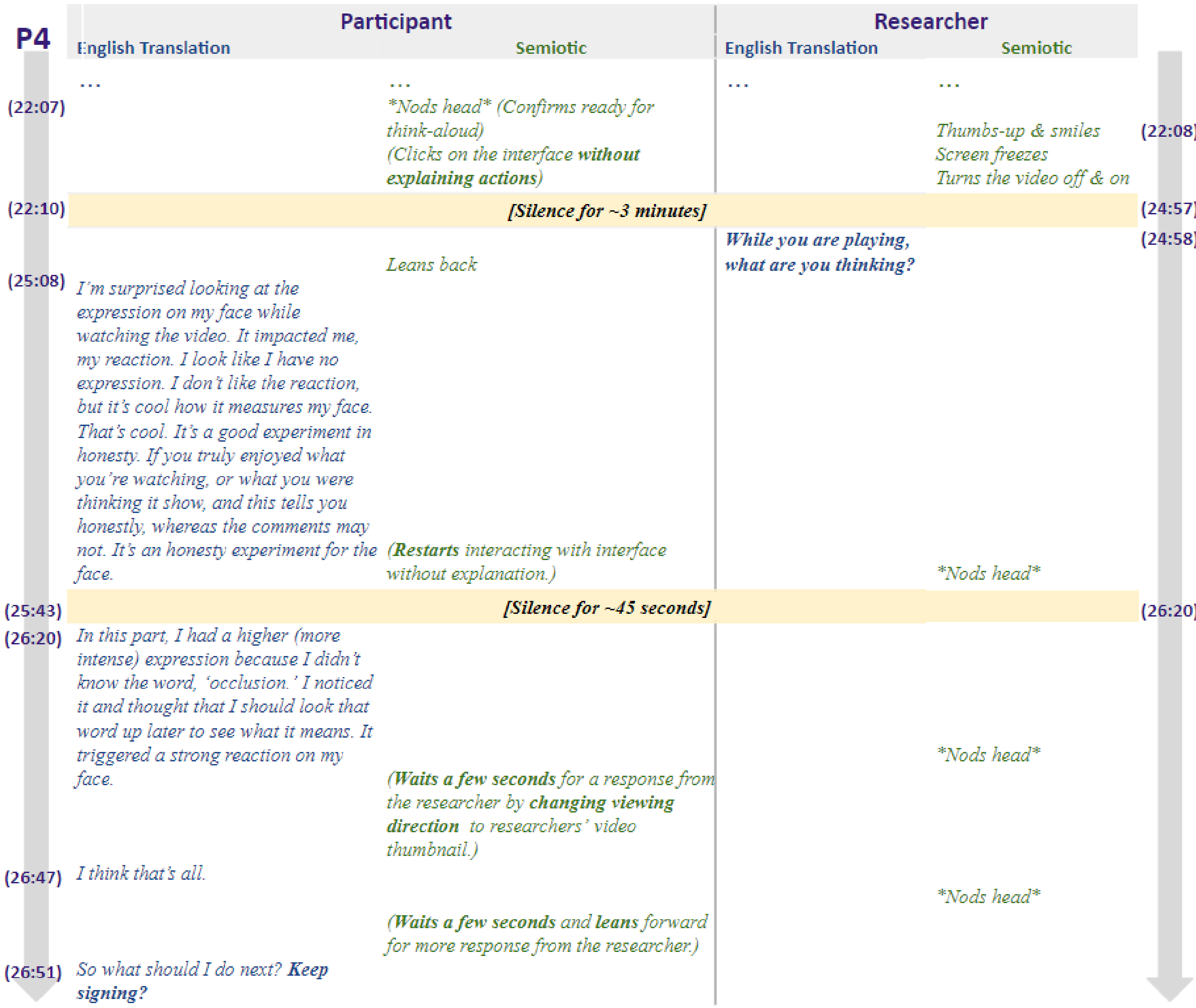

Figure 8: Synchronicity conflicts between signing and navigating the interface were also observed, exemplified in the sequence instead of concurrent signing and interacting with the interface in the 25:08 and 26:20. Participants showed an eagerness to engage in conversation with the researchers, employing strategies such as pausing (26:20, 26:47), changing their viewing direction (26:36), and leaning forward (26:49). The researcher encountered some problems with the internet connection (around 23:58), so she had to turn the camera on and off. This process took her over one minute to complete and prompt P4 to continue. P4 requested anonymity, so none of P4's images are shown.

peers' high intensity with upward pointing arrows and low intensity with downward pointing arrows in Figure 2 -ⓖ. Many participants signed "UP" and "DOWN," mirroring how the arrow-shaped icons represented peer intensity. These are the same signs P8 demonstrated, but they were used to refer to very different parts of the interface. Researchers diligently observe the participants' semiotic behaviors to fully understand what was being referenced during the think-aloud sessions.

The last example comparing the use of the term intensity revisits the quote from P11 using simultaneous spoken English and signed ASL in Figure 4. When he recognized the peaks on his emotional intensity line graph, P11 spoke the word "intensity" and simultaneously fingerspelled the word "INTENSITY." Fingerspelling involves configuring the fingers into handshapes that represent individual letters; each handshape can be signed to spell out a word, like writing each letter one out at a time in English. P11's decision differs from the "UP" and "DOWN" signs other participants used and the "high/HIGH" and "low/LOW" descriptions P2 used. In his reflective think-aloud utterances, P11 used the terminology explicitly provided by the interface's text. As mentioned previously, the choice

Figure 9: P8 used a sign pointing upward [a1] to refer to the intensity spike in her line graph [a0] with her cursor hovered on the graph, instead of fingerspelling it. Fingerspelling is a method of labeling words by spelling them out letter by letter using the manual alphabet, instead of visually explaining them. This same sign was also observed being used by other participants to refer to patterns in line graphs in Figure 2 -ⓒ, as well as arrows representing their peers' intensity in Figure 2 -ⓖ (demonstrated in [a2] by a researcher). Therefore, the researcher was confused as to what P8 was referring to. Signing provides a convenient way for DHH user to visually express themselves; however, signs that are not explicitly labeled may lack clarity for researchers to understand which interface component the sign is referring to. P8 has given permission to share her face and likeness.

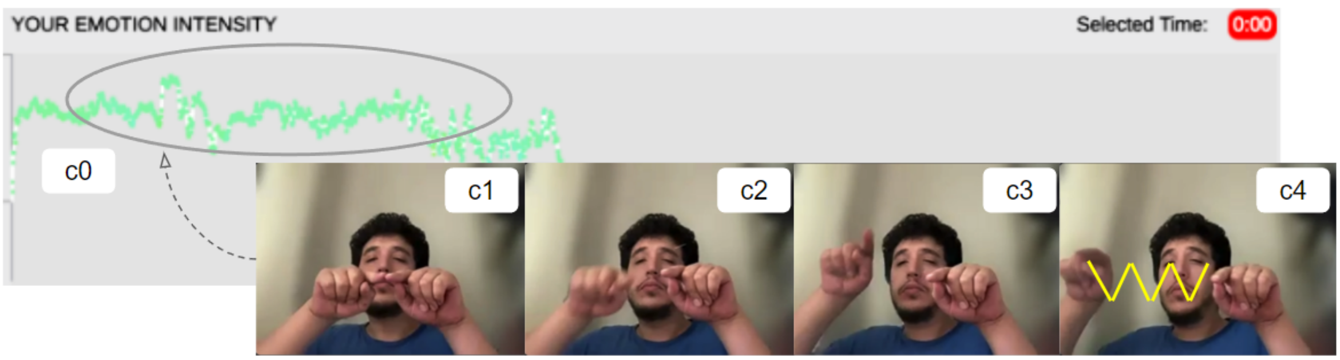

Figure 10: Participants used different signs to refer to the same interface component due to differences in visual features. Our participants traced the shape of the line graph in the air with their extended index fingers as if they are drawing the peaks and valleys on an invisible whiteboard in the space in front of them as shown by P9 [c1-4] (Yellow lines are finger-moving trajectories between [c4] that matches the visual features of own emotion intensity line [c0]. In comparison, P8 refers to its own emotion intensity graph in Figure 9 [a0] using [a1].

to spell out a word in ASL typically requires more time because each letter of the word must be signed individually. In the case of P11, that would require nine handshapes for one spoken word. He could have chosen to only say the word "intensity" without any ASL as he did for other words in this same sentence. Or he could have used an ASL sign. However, P11 chose to speak and fingerspell. However, due to the inherent time difference between the two, P11 prevented only signing three of the nine letters: IEY. This self-mediation allowed the spoken and signed modes to align smoothly without a pause or elongated speech.

Sign languages often capitalize on the use of multiple articulators (e.g., hands, body, face, etc.), whereas spoken languages are primarily limited to the use of the vocal tract, e.g., the tongue. Sign language allowed our participants to employ 3D space, layering phonetic elements to convey a wealth of visual information that spoken languages can only express sequentially and linearly. While signs are more descriptive due to their inherent visual elements, fingerspelling is more explicit, providing a specific label for the word being produced. To illustrate, consider the term 'line graph.' In ASL, the sign for this concept is highly iconic, and it looks very much like the provided machine-generated image (Figure 10 [c0]). During the interview portion of the study, participants not only produced this sign, but they also manipulated it. To illustrate, compare how P7 signed "LINE GRAPH" four different ways within twenty seconds in Figure 11.

Participants utilized ASL's multiple articulators along with signs and gestures during their think-aloud activity. This showcased how resourcefully visual information can be conveyed. Sometimes this required researchers to ask clarifying questions or view the participants' cursor actions to identify what features were being discussed. While the participants could have spelled out terms such as "line graph," sign language already maximizes the visual information often inherent to ASL signs, which our participants then modified to express unique visual information that the English terminology alone lacked. It is important to recognize the value of understanding how DHH participants describe their user experience, which reflects their community culture as well as an array of language choices that could go beyond these current two findings.

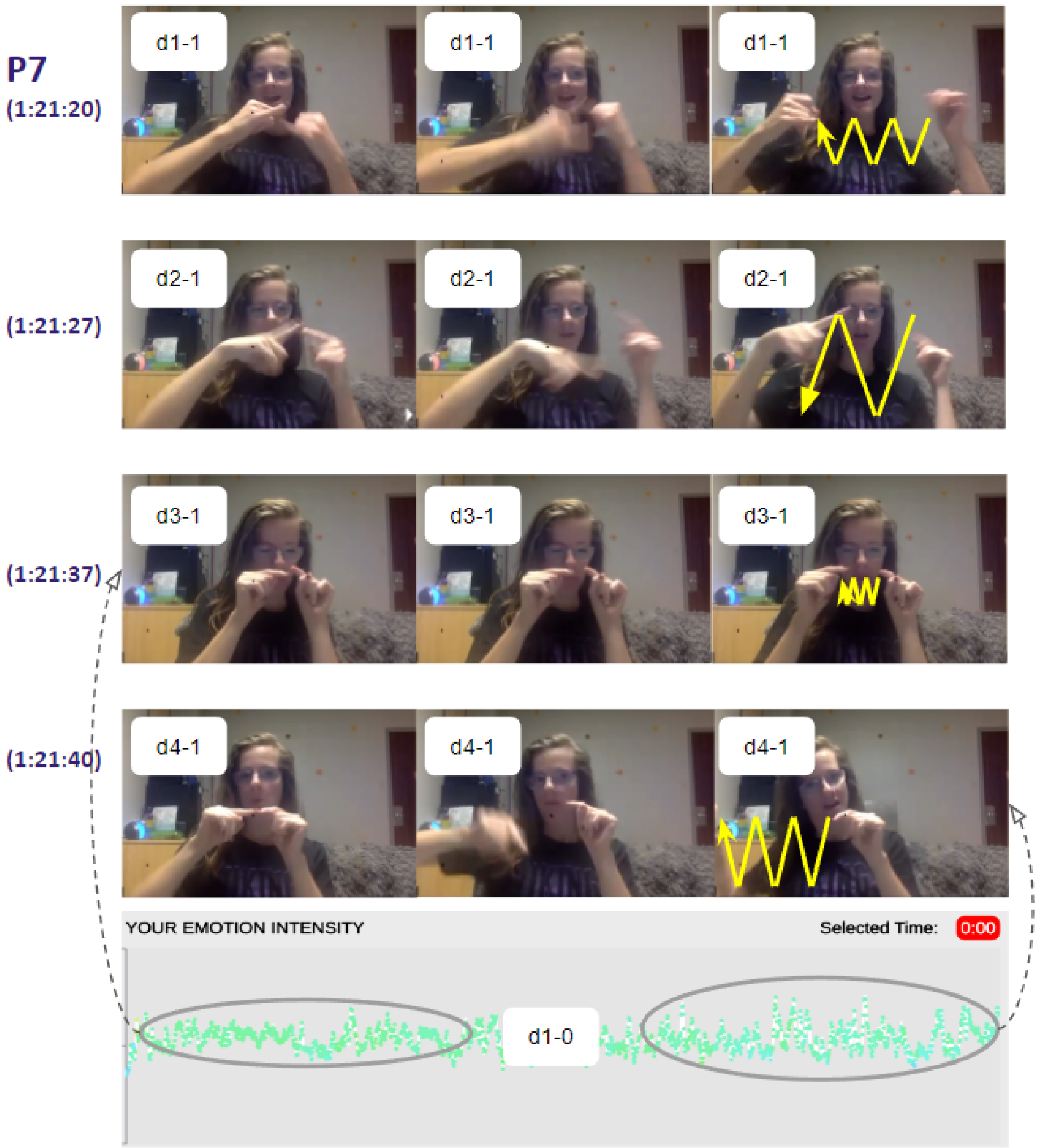

Figure 11: When referring to the same interface component, participants use various signs to visually describe it. Zoom video thumbnail captured four ways ([d1-1] to [d1-3], [d2-1] to [d2-3], [d3-1] to [d3-3], [d4-3] to [d4-3]) P7 referred to the emotion intensity graph [d0]). The first two examples, [d1-1] and [d2-1], are both plain nouns that are not manipulated. The difference in these productions is how [d1-1] was produced with pinky fingers while [d2-1] was produced with index fingers, similarly to P9 in Figure 10. The second two examples illustrate how P7 modified the movement of the sign "LINE GRAPH" to codify additional information. In [d3-1], P7 made a small zig-zag movement up and down to represent the changes at the beginning when her graph. Moments later, she enlarged the movement to show the stronger fluctuation at the end of her graph in [d4-1]. Yellow lines are finger-moving trajectories.

5 DISCUSSIONS

5.1 Multi-Mode Think-aloud Protocol: Data Collection and Analysis

Our findings in section 4.1 revealed that DHH participants used multiple modes to communicate that relate to their diverse linguistic backgrounds (e.g., formative language and hearing status). Our findings are takeaways from participants, who are DHH learners with diverse linguistic and language abilities. These findings expand the focus of previous DHH think-aloud studies that only analyzed DHH's semantic information signed via ASL [5, 14, 33].

To accurately analyze and present the message expressed by DHH users, we proposed a data scheme in Figure 3 to analyze think-aloud data. This format consists of four key dimensions: semantic (e.g., English, ASL), semiotic, temporal relationship, and spatial relationship. It is important to note that English and ASL possess different grammar and may not necessarily map one to one sequentially, and based on DHH's student's linguistic background, the leading semantic mode differs. Furthermore, this scheme allows researchers to examine how different modes explain or do not explain each other. More specifically, our section 4.2 demonstrates using our scheme to understand the temporal relationship between different modes (e.g., sequentially interacting with interfaces and explaining interaction) and our section 4.3 demonstrates using our scheme to understand the spacial relationship between different modes (e.g., the mapping between visual signs, cursor, and interface components). Expanding on [5], which suggested that the data collection process for DHH users is the main challenging point, we argue that the analysis and presentation of the data should also be taken into account to make it readable for researchers/designers not familiar with DHH research.

Our research team developed the proposed format in Figure 3.3 to address the issues with text-based transcripts generated from video recordings used in various user studies with disabilities (e.g., [29]). This approach disregards many communication modes used by DHH participants and the relationships between them. Consequently, researchers had to switch between different recordings to gain a holistic understanding of user behavior. Moreover, our research team - with members possessing a range of language abilities in ASL and English - had difficulty synthesizing the different modes due to the nature of how different modes explain or do not explain each other. Therefore, our proposed schema is intended to help hearing researchers and designers with minimal understanding of deaf culture gain insight into linguistic approaches toward DHH user studies, and ultimately contribute to strategies for anticipating inclusive and access needs at all stages of the research process [25].

Moreover, it is important to acknowledge that the effectiveness of coding multi-mode responses may be somewhat limited for users with less diverse backgrounds or for tools that prioritize lower levels of self-expressiveness, such as workspace tools. Furthermore, it is worth considering that the degree of hearing loss can affect an individual's speaking abilities. In this regard, coding multi-mode responses could potentially offer greater inclusivity for individuals with more severe hearing impairments, who often experience lower speaking abilities. However, further research is necessary to delve deeper into this topic and gain a more comprehensive understanding.

5.2 Inclusive Virtual Think-aloud Protocols in User Studies with DHH

Recent papers suggest critical reflections on the method of conducting accessibility research [22]. This is in response to the third wave of HCI research, which emphasizes the importance of recognizing how perspectives and biases can shape user experience and research outcomes [3]. Additionally, it is important to consider the potential impact of research on different users and communities by striving for greater inclusion and diversity practices in research [32].

5.2.1 Lack of Clarity in Signed Discourse during Self-Talk. Our findings revealed two unique aspects of signed discourse that lead to a lack of clarity in conducting virtual think-aloud sessions in section 4.2 and section 4.3 - preferences for staying in a signing mode and preferences for using visual representations to refer to objects rather than labeling them (in our case, interface components e.g. line chart). These two aspects are rooted in DHH participants relying heavily on multiple modes to express preferences in signed discourse compared to spoken discourse [7, 21, 42]. This requires researchers to be aware of all modes in real time and to ask for clarifications immediately. For example, in section 4.3 Figure 9, researchers need to quickly map the relationship between participants' self-identified and likely ambiguous and diverse visual representations to interface components and understand what sign/term(s) such as "this/these" refer to. Then researchers need to ask corresponding clarification questions after processing this multi-mode information (e.g., Zoom video thumbnail, shared screen, and audio). In order to ensure clear communication in our think-aloud sessions, our lead researchers are proficient in both ASL and English. However, if an interpreter is required for hearing researchers unskilled in sign language(s), as explored in [33, 34], this can complicate the real-time interactions.

Platforms selected for conducting virtual think-aloud studies can impede researchers from fully capturing high-quality multi-mode data. Researchers must be aware of the limitations of the platform they are using. For instance, in our case, using Zoom for virtual think-aloud sessions may limit the quality of data collected, even when participants attempt to adjust their bodies to ensure their expressions are fully captured (e.g., P1 leaned backward in Figure 6 ⓓ). Additionally, some expressions may still be signed out of the camera view (e.g., TIME and DOWN were partially covered by the P8's name tag in Figure 9, P3's hands were not captured by the camera in Figure 5). Additionally, during our initial data analysis, we encountered challenges using Zoom that prevented the inclusion of some data in this paper. For instance, the screen-thumbnail ratio of some participants was too low, making it difficult to interpret their ASL expressions (an extreme case of 9 [a1]). Additionally, Zoom's auto-pinning feature, which auto-selected the video thumbnail of the loudest Zoom user, prevented some signed expressions from being recorded, as this user was not necessarily the person who was signing. The inclusion of DHH researchers in collecting data, leading think-aloud sessions, and analyzing think-aloud data is also necessary [14].

5.2.2 When and How to Prompt DHH Users for Clarification? Prompting and clarifying questions are often necessary components of think-aloud studies. However, inexperienced facilitators can easily bias user behavior with their interruptions [30]. Traditionally, when conducting think-aloud with DHH offline, users may have lapses in explanations (approximately 10 seconds or more). Researchers encouraged users to continue sharing explanations of their behaviors and thoughts by verbalizing orally which is in real-time interpreted to ASL or by tapping on the participant's shoulder [5, 33]. However, some user behaviors are not as noticeable offline. Therefore, in our study, we slightly changed the guidelines for prompting and clarifying questions by extending the time period to more than 1 minute and nodding the head more often to show "keep talking" in a less interruptive manner. By doing so, we gained a new understanding of DHH think aloud that there is an asynchronicity conflict between signing and navigating the interface in section 4.2, which leads participants to sequentially switch between explaining and interacting with the interface. This expands findings in [14], which observed that some DHH only signed their thoughts prior to interacting with the interface, and explains that this is possible because of the synchronicity conflict.

Our findings in Section 4.3 suggest that these prompts may not be sufficient in virtual think-aloud sessions, as DHH participants require close attention and feedback from researchers in order to continue. Moreover, the lack of clarity in sign discourse necessitates more clarification questions and interruptions from the researcher, increasing the potential for researchers to bias users. To ensure a fair and unbiased study, further research needs to develop guidelines for proper prompting and clarifying of questions when conducting think-aloud studies with DHH. Furthermore, researchers could explore innovative ways to facilitate think-aloud sessions without biasing and interrupting user behavior. Drawing inspiration from studies of aphasia charities adapting communication to videoconferencing [29], researchers can use clear and big physical visual aids such as a 'thumbs-up' to encourage expression without interruption or a "explain once again" to request clarification instead of using only facial expressions and ASL.

5.3 Design Implications for Tools to Support Think-aloud Protocol

5.3.1 Human-AI Collaboration to Prompt for Clarifications. To improve the lack of clarity in the think-aloud protocol outlined in section 4.2, tools can be developed to better capture user experience and prompt explanations. This can help maintain the "in-the-moment" benefit of the think-aloud protocol.

AI assisting the researcher to understand when and where to prompt for clarification. The use of think-aloud analysis tools to analyze the user experience of hearing participants [17, 18, 38] has provided insight into how AI could be used to help the researcher to detect user experience leveraging multi-mode communication in DHH users. Specifically, this includes plug-ins detecting clarity and reminding researchers to ask clarification questions when executing virtual think-aloud. Semiotic features, such as leaning forward to indicate "finish" (when to ask), nodding the head to indicate active engagement in interactions (when to ask - not nodding head), and pointing to the screen to indicate interface components (where) were identified as potentially useful.

AI prompting DHH users for clarification with consideration of visual attention. Tools such as plugins that are designed to improve the clarity of think-aloud sessions for DHH users can include prompts for explanations after detecting head nods or unclear pointing fingers, as mentioned in the reminder for researchers. Zoom could also introduce an optional setting to visualize participants' gestural expressiveness (calculating algorithm [28]) to help them become more aware of their own expressions and pauses. Additionally, the system might consider incorporating different levels of visually demanding prompts to reduce the split in visual attention between the tested interface and the researcher's video thumbnail [20, 37]. Examples of such prompts could include more visually demanding pop-up windows for clarification, or less visually demanding interventions such as a small green light or avatar next to the tested interface to encourage users to keep talking. However, caution must be taken when designing real-time prompts, as they may bias further users' explorations and create other concerns such as privacy [10, 29].

5.3.2 Supporting Expressiveness with Personalized Visual Signs Glossary in HCI Research. The current think-aloud protocol may be ambiguous in the mapping between visually descriptive signs and interface components as shown in section 4.3. To address this issue, tools should be designed to support personalized expressions that clearly match the interface components without standardizing and oppressing their expressions. For example, after a user finishes a think-aloud, the interface can prompt participants to create their own glossary, establishing ASL signs and visual descriptions that refer to the interface components (e.g., intensity and line graph). This glossary can be reused for different user studies and be expanded as needed (e.g., the line graph can be used for various interfaces with visualizations). redundant language, as previously suggested as a reason for not using think-aloud with DHH [20]. Our proposed data scheme also provides initial insights into designing a visual sign language glossary, allowing users to attach video clips and images of signs to other communication modes over the course of a user study.

Moreover, this approach can be applied to general science, technology, engineering, and mathematics (STEM) online learning to address the limited amount of conventionalized signs for DHH user within these fields [11, 12] and generate crowd-sourced signs for terminologies in those fields for further inclusive online learning [2]. Furthermore, researchers need to recognize the potential of understanding how DHH participants embody and describe their user experience, as pointed out in section 4.3. The repeated use of language and modes does not necessarily equate to redundant language as mentioned as a reason for not using think-aloud with DHH [20].

6 CONCLUSIONS

This study sought to investigate the use of the think-aloud method in virtual user experience studies with DHH populations, an area with limited current research. By conducting twelve virtual think-aloud sessions on Zoom with DHH participants, the study found that multiple communication modes were often used, including concurrent and sequential semantic (e.g., ASL, English) and semiotic (e.g., gestural and facial) expressions. These modes sometimes resulted in asynchrony between signing and navigating interfaces, as well as the use of visual descriptive signs instead of explicit labels. Such behaviors could present challenges for obtaining clear think-aloud data. The study provides implications for researchers to conduct virtual think-aloud studies with DHH participants, including when and how to prompt for clarification. It is recommended that researchers understand the multiple communication modes used by DHH individuals and consider the temporal and spatial dependencies among them. It is important to acknowledge that the participants in our study are college students who do not have a background in STEM. This could potentially influence the way they articulate interface components, and it is necessary to further explore this in future research by including DHH users from diverse fields.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 2119589. The authors would like to thank the anonymous reviewers for their efforts and valuable comments. We extend our sincere appreciation to Saumya Malhotra for her valuable contributions in data annotation, figure editing, and feedback provided for this manuscript. Additionally, we would like to express our gratitude to all the participants for their thoughtful and reflective feedback.

REFERENCES

[1] Peter Afflerbach. 2001. Verbal reports and protocol analysis. In *Methods of literacy research*. Routledge, 97–114.
[2] Bhavya Bhavya, Si Chen, Zhilin Zhang, Wenting Li, Chengxiang Zhai, Lawrence Angrave, and Yun Huang. 2022. Exploring collaborative caption editing to augment video-based learning. *Educational technology research and development* 70, 5 (2022), 1755–1779.
[3] Susanne Bødker. 2006. When Second Wave HCI Meets Third Wave Challenges. In *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles* (Oslo, Norway) *(NordiCHI '06)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/1182475.1182476
[4] Riley Botelle and Chris Willott. 2020. Birth, attitudes and placentophagy: a thematic discourse analysis of discussions on UK parenting forums. *BMC Pregnancy and Childbirth* 20, 1 (2020), 1–10.
[5] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
[6] Erica Burman and Ian Parker. 2016. *Discourse analytic research: Repertoires and readings of texts in action*. Routledge.
[7] Olga Capirci and Chiara Bonsignori. 2022. Beyond Orality: The Case of Sign Languages. *Rethinking Orality I* (2022), 69.
[8] Sambhavi Chandrashekar, Tony Stockman, Deborah Fels, and Rachel Benedyk. 2006. Using think aloud protocol with blind users: a case for inclusive usability evaluation methods. In *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*. 251–252.
[9] Si Chen, Yixin Liu, Risheng Lu, Yuqian Zhou, Yi-Chieh Lee, and Yun Huang. 2022. "Mirror, Mirror, on the Wall" - Promoting Self-Regulated Learning Using Affective States Recognition via Facial Movements. In *Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 1300–1314. https://doi.org/10.1145/3532106.3533500
[10] Xinyue Chen, Si Chen, Xu Wang, and Yun Huang. 2021. "I Was Afraid, but Now I Enjoy Being a Streamer!": Understanding the Challenges and Prospects of Using Live Streaming for Online Education. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3, Article 237 (jan 2021), 32 pages. https://doi.org/10.1145/3432936
[11] Mel Chua, Ian Smith, Miriam Nathan Lerner, Sarah Jacobs, Rita Straubhaar, Ruth Anna Spooner, and Perseus McDaniel. 2019. Does" Affordance" Mean" Thing-inform"?: Case Studies in Seeing Engineering Meaning Differently Through the Process of Technical ASL Vocabulary Creation. In *2019 ASEE Annual Conference & Exposition*.
[12] Kaitlyn Clark, Asma Sheikh, Jennifer Swartzenberg, Ashley Gleason, Cody Cummings, Jonathan Dominguez, Michelle Mailhot, and Christina Goudreau Collison. 2021. Sign Language Incorporation in Chemistry Education (SLICE): Building a Lexicon to Support the Understanding of Organic Chemistry. *Journal of Chemical Education* 99, 1 (2021), 122–128.
[13] Didan Deng, Zhaokang Chen, and Bertram E Shi. 2020. Multitask emotion recognition with incomplete labels. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*. IEEE, 592–599.
[14] Becca Dingman, Garreth W Tigwell, and Kristen Shinohara. 2021. Interview and Think Aloud Accessibility for Deaf and Hard of Hearing Participants in Design Research. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–3.
[15] K Anders Ericsson and Herbert A Simon. 1998. How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity* 5, 3 (1998), 178–186.
[16] Mingming Fan, Yiwen Wang, Yuxi Xie, Franklin Mingzhe Li, and Chunyang Chen. 2022. Understanding How Older Adults Comprehend COVID-19 Interactive Visualizations via Think-Aloud Protocol. *International Journal of Human–Computer Interaction* (2022), 1–17.
[17] Mingming Fan, Ke Wu, Jian Zhao, Yue Li, Winter Wei, and Khai N. Truong. 2020. VisTA: Integrating Machine Intelligence with Visualization to Support the Investigation of Think-Aloud Sessions. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (2020), 343–352. https://doi.org/10.1109/TVCG.2019.2934797
[18] Mingming Fan, Qiwen Zhao, and Vinita Tibdewal. 2021. Older adults' think-aloud verbalizations and speech features for identifying user experience problems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.
[19] Mingming Fan, Qiwen Zhao, and Vinita Tibdewal. 2021. Older Adults' Think-Aloud Verbalizations and Speech Features for Identifying User Experience Problems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 358, 13 pages. https://doi.org/10.1145/3411764.3445680
[20] Ryo Iijima, Akihisa Shitara, and Yoichi Ochiai. 2022. Designing Gestures for Digital Musical Instruments: Gesture Elicitation Study with Deaf and Hard of Hearing People. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–8.
[21] Adam Kendon. 1988. Parallels and divergences between Warlpiri sign language and spoken Warlpiri: Analyses of signed and spoken discourses. *Oceania* 58, 4 (1988), 239–254.
[22] Amanda Lazar, Ben Jelen, Alisha Pradhan, and Katie A Siek. 2021. Adopting diffractive reading to advance HCI research: a case study on technology for aging. *ACM Transactions on Computer-Human Interaction (TOCHI)* 28, 5 (2021), 1–29.
[23] Duri Long. 2023. Conducting Remote Design Research on Embodied, Collaborative Museum Exhibits. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–8.
[24] Kelly Mack, Emma McDonnell, Dhruv Jain, Lucy Lu Wang, Jon E. Froehlich, and Leah Findlater. 2021. What do we mean by "accessibility research"? A literature survey of accessibility papers in CHI and ASSETS from 1994 to 2019. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–18.
[25] Kelly Mack, Emma McDonnell, Venkatesh Potluri, Maggie Xu, Jailyn Zabala, Jeffrey Bigham, Jennifer Mankoff, and Cynthia Bennett. 2022. Anticipate and adjust: Cultivating access in human-centered methods. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.
[26] Emma J McDonnell, Ping Liu, Steven M Goodman, Raja Kushalnagar, Jon E Froehlich, and Leah Findlater. 2021. Social, environmental, and technical: Factors at play in the current use and future design of small-group captioning. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–25.
[27] Laura Meskanen-Kundu et al. 2021. Remote user study: a reflection on remote user research methods during the COVID-19 pandemic. (2021).
[28] Prasanth Murali, Javier Hernandez, Daniel McDuff, Kael Rowan, Jina Suh, and Mary Czerwinski. 2021. Affectivespotlight: Facilitating the communication of affective responses from audience members during online presentations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.
[29] Timothy Neate, Vasiliki Kladouchou, Stephanie Wilson, and Shehzmani Shams. 2022. "Just Not Together": The Experience of Videoconferencing for People with Aphasia during the Covid-19 Pandemic. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16.
[30] Jakob Nielsen. 2012. *Thinking Aloud: The #1 Usability Tool*. https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/
[31] Colorado Department of Human Services. 2022. *Deaf, Hard of Hearing, and DeafBlind Demographics Guide*. https://ccdhhdb.com/wp-content/uploads/2022/09/DHHDB-Demographics.pdf
[32] Sirpa Riihiaho et al. 2015. Experiences with usability testing: Effects of thinking aloud and moderator presence. (2015).
[33] Vera Roberts and Deborah Fels. 2002. Methods for inclusion: employing think aloud protocol with individuals who are deaf. In *Computers Helping People with Special Needs: 8th International Conference, ICCHP 2002 Linz, Austria, July 15–20, 2002 Proceedings 8*. Springer, 284–291.
[34] Vera Louise Roberts and Deborah I Fels. 2006. Methods for inclusion: Employing think aloud protocols in software usability studies with individuals who are deaf. *International Journal of Human-Computer Studies* 64, 6 (2006), 489–501.
[35] Jazz Rui Xia Ang, Ping Liu, Emma McDonnell, and Sarah Coppola. 2022. "In This Online Environment, We're Limited": Exploring Inclusive Video Conferencing Design for Signers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 609, 16 pages. https://doi.org/10.1145/3491102.3517488
[36] Matthew Seita, Sooyeon Lee, Sarah Andrew, Kristen Shinohara, and Matt Huenerfauth. 2022. Remotely Co-Designing Features for Communication Applications using Automatic Captioning with Deaf and Hearing Pairs. In *CHI Conference on Human Factors in Computing Systems*. 1–13.
[37] Marcos Serrano, Jolee Finch, Pourang Irani, Andres Lucero, and Anne Roudaut. 2022. Mold-It: Understanding how Physical Shapes affect Interaction with Handheld Freeform Devices. In *Proceedings of the 2022 CHI Conference on Human*

Factors in Computing Systems. 1–14.
[38] Ehsan Jahangirzadeh Soure, Emily Kuang, Mingming Fan, and Jian Zhao. 2021. CoUX: collaborative visual analysis of think-aloud usability test videos for digital interfaces. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2021), 643–653.
[39] Gary W Taylor and Jane M Ussher. 2001. Making sense of S&M: A discourse analytic account. *Sexualities* 4, 3 (2001), 293–314.
[40] Paul Ten Have. 2007. Doing conversation analysis. *Doing Conversation Analysis* (2007), 1–264.
[41] Clayton Valli and Ceil Lucas. 2000. *Linguistics of American sign language: An introduction.* Gallaudet University Press.
[42] Sébastien Vandenitte. 2022. Making referents seen and heard across signed and spoken languages: Documenting and interpreting cross-modal differences in the use of enactment. *Frontiers in psychology* 13 (2022), 784339.
[43] Iris Xie, Rakesh Babu, Tae Hee Lee, Melissa Davey Castillo, Sukjin You, and Ann M Hanlon. 2020. Enhancing usability of digital libraries: Designing help features to support blind and visually impaired users. *Information Processing & Management* 57, 3 (2020), 102110.