# Chat in the Hat: A Portable Interpreter for Sign Language Users


Larwan Berke
Gallaudet University
Washington, DC, USA

William Thies
Microsoft Research
Bangalore, India

Danielle Bragg
Microsoft Research
Cambridge, MA, USA



## ABSTRACT

Many Deaf and Hard-of-Hearing (DHH) individuals rely on sign language interpreting to communicate with hearing peers. If on-site interpreting is not available, DHH individuals may use remote interpreting over a smartphone video-call. However, this solution requires the DHH individual to give up either 1) the use of one signing hand by holding the smartphone or 2) their ability to multitask and move around by propping the smartphone up in a fixed location. We explore this problem within the context of the workplace, and present a prototype hands-free device using augmented reality glasses with a hat-mounted fisheye camera and mic/speaker. To explore the validity of our design, we conducted 1) a video interpretability experiment, and 2) a user study with 18 participants (9 DHH, 9 hearing) in a workplace environment. Our results suggest that a hands-free device can support accurate interpretation while enhancing personal interactions.



## CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; *Mixed / augmented reality*; *Accessibility technologies*.

## KEYWORDS

Video Remote Interpreting (VRI), Deaf, DHH, Accessibility, Augmented Reality

**ACM Reference Format:**
Larwan Berke, William Thies, and Danielle Bragg. 2020. Chat in the Hat: A Portable Interpreter for Sign Language Users. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20), October 26–28, 2020, Virtual Event, Greece.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3373625.3417026


## 1 INTRODUCTION

Many Deaf and Hard-of-Hearing (DHH) individuals in the U.S. struggle to communicate with people who do not know American Sign Language (ASL) [15]. Like other sign languages, ASL is a gesture-based language, with a unique grammar and vocabulary.[1] It is the primary language of the Deaf community, but the hearing majority uses English (not ASL), resulting in communication barriers between DHH and hearing people.

To bridge the communication divide, many DHH individuals prefer sign language interpreters. ASL interpreters are professionally trained to interpret (or translate) bidirectionally between ASL and English. They can be used for one-on-one conversations, group meetings, or presentations with a mix of DHH and hearing participants. When an on-site interpreter is unavailable (due to scheduling, budget, location, etc.), and the meeting location is fixed (e.g., an office meeting), Video Remote Interpreting (VRI) may be appropriate. VRI consists of a laptop/tablet placed on a table[2] connected to a remote interpreter via video call, and is offered by several companies (e.g., Purple[3], Sorenson[4], and InDemand[5]).

However, existing remote interpreting systems restrict DHH users' ability to move around and multitask. If a DHH individual wishes to walk around during a conversation, they typically set up a remote interpreter on a smartphone, which they then hold in one hand while using the other hand to sign to the remote interpreter. At best, this setup restricts communication by limiting the signer to one-handed signing; at worst, signing becomes infeasible if their second hand is holding something (e.g., a bag, baby, or umbrella) or doing something (e.g., writing on a whiteboard). In these cases, the DHH individual must either pause the conversation to perform tasks, or pause the tasks to converse.

In particular, these difficulties create unequal access to the workplace for DHH vs. hearing workers. In-person interpreting in the workplace is not always feasible. Small companies may not have the resources to hire on-site support, and even well-resourced offices can face problems due to lack of availability or last-minute scheduling. The default fall-back solution of remote interpreting does not provide equal access, as workplaces involve mobile conversations (e.g., while walking down the hall) and conversations while multitasking (e.g., discussion at the whiteboard). Remote interpreting setups currently do not enable a DHH worker to fully engage in these conversations (as described previously). The authors have personally experienced such workplace inequality, which may also violate legal requirements for equal employment [1].

To help promote equal access to the workplace, and more generally the use of remote interpreting while mobile or multitasking, we propose our hands-free device. Our hands-free device frees the hand that otherwise would have to hold the phone, by providing a hat with a camera that captures the signing space and a speaker that provides audio. Having both hands free newly allows users to converse naturally with both hands, while having the ability to move around or multitasking during conversations. The hands-free

[1]While we focus on ASL in this work, much of our work applies to sign languages more generally.

[2](or a television with webcam on a wheeled cart)
[3]https://signlanguage.com/vri/
[4]http://www.schedulevri.com/
[5]https://www.indemandinterpreting.com/video-interpreting/



*ASSETS '20, October 26–28, 2020, Virtual Event, Greece*



ACM ISBN 978-1-4503-7103-2/20/10…$15.00
https://doi.org/10.1145/3373625.3417026

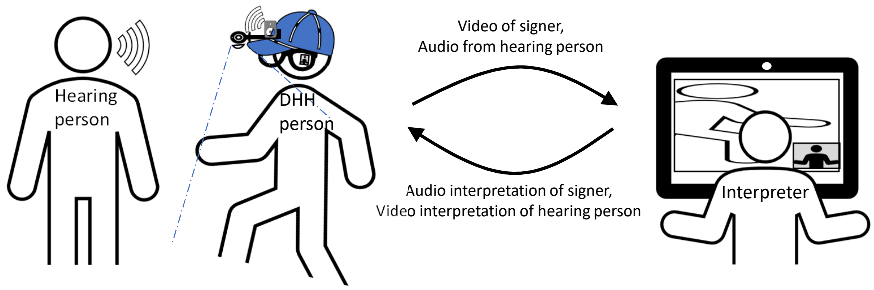



**Figure 1: Overview of the hands-free remote interpreting system (for a detailed description of the device see Figure 2). A DHH person and hearing person can communicate in person in real-time, using a remote interpreter. Both the DHH and hearing person have both hands free, and are able to move around or multitask.**

technology could also benefit DHH individuals who have lost the ability to use one of their hands (e.g. broken bones, nerve damage, etc.). Full system details are described in Section 3.

This work is guided by the following main research questions:

**RQ1**: Is it possible to design a hands-free device for remote sign language interpretation?

**RQ2**: Can a hands-free device accurately capture and convey sign language in a way that supports accurate interpretation?

**RQ3**: What are the benefits and downsides of using a hands-free device over a traditional smartphone, for DHH and/or hearing participants?

To explore these questions, we propose a hands-free prototype for remote sign language interpretation (RQ1). We also conducted two studies: one exploring the intelligibility of the audio and video captured by the system for a human interpreter (RQ2), and one on the DHH & hearing user experience of communicating over the hands-free prototype during typical workplace conversations (RQ3).

## 2 BACKGROUND AND RELATED WORK

Related work spans remote interpreting, sign language video conversations, and intelligibility of sign language over video. To this space, we contribute the first hands-free prototype for remote sign language interpreting, and an initial exploration of its effectiveness for remote sign language interpreting. In so doing, we also validate many of the factors found to be important to remote interpreting in prior work.

### 2.1 Remote Video Interpreting

In educational settings, CoUnSiL [18] developed a platform which enabled interpreters, DHH students, and hearing teachers to communicate over a single application. To minimize the student's visual dispersion, the system displayed both the presentation slides and interpreter. The CoUnSiL client application runs on a standard desktop (or laptop) computer, and the server component is hosted in a datacenter as to provide high-speed video streaming from the remote interpreter. System evaluation revealed several issues impacting the effectiveness of the interpretation: video quality and resolution; variable lighting; and interfering background noise.

Another team [14] evaluated a smartglass-based solution for DHH students in the classroom to display the interpreter in their field of view. Their work did not display any significant differences over the baseline (no smartglasses) for comprehension, but participants reported that they could imagine using the smartglasses in future classes. One major reported benefit of the smartglasses was reduced head movement as the students did not have to shift their attention between the interpreter, slides, teacher, and other information sources.

The ClassInFocus team [7] investigated how a multi-modal classroom could benefit students by displaying everyone on the same screen. The authors hosted a focus group and prototype session, and found that visual attention was an overlooked factor. The authors noted: "*The focused attention required to watch the interpreter may induce tunnel vision and overpower other changes in the classroom*"[7, p. 69] Based on this finding, the team then developed a notification system in order to direct the Deaf student's attention to/from the video display.

In legal settings, Napier [16] recorded several scenarios of Deaf individuals in courtrooms using interpreting services. Their work revealed that VRI was effective for courtrooms, given certain accommodations needed to facilitate high-quality interpretation. In their paper, *Scenario 2* presented a case wherein the Deaf person was in court while the interpreter was provided by VRI. Several important criteria were reported by the Deaf individual and the interpreter: proper visibility of each other, visibility of fingerspelling[6], and being able to get the attention of the viewer due to competing visual stimuli.

In Finland, a team [11] evaluated a VRI provider over several years and their team agreed that VRI made interpreters' lives easier (e.g. enabling them to work from home, or from a familiar office space). However, their interpreters mentioned several challenges to their work such as: difficulty interpreting a wide variety of topics, limited ability to prepare in advance, and the possibility of technical difficulties.

A team of researchers [10] looked into potential applications of VRI services (e.g. healthcare) and outlined VRI requirements: minimizing network latency and jitter, ability for the interpreter to see all individuals in the room (vs. only the Deaf client), and sufficient video size/quality. Importantly, the interpreters in the experiment noted that "*they were missing out on information because they were not able to easily see items being shown or demonstrated*" [10, p. 6]. To mitigate this issue, the team recommended that the

[6] Fingerspelling is an important component of sign language(s) wherein an individual produces single letters of a writing system by using the fingers in a pre-specified pose for each letter [13].

Deaf individual be made aware of the "interpreter perspective" so they know to position themselves and objects in view of the camera.

One innovative team explored projecting an interpreter onto the hearing individual's body [8]. Through a Wizard-of-Oz study, the authors identified the most important aspect of the device: the ability to fade into the background and not be noticed by the parties in the conversation. When conversation breaks occurred, the device's limitations suddenly became obvious, and the Deaf participants became frustrated with the ineffective communication and abandoned their tasks. Users were also required to "*[face] the conversation partner while speaking*", which their participants routinely violated, thus missing important visual cues such as nodding or pointing.

In a large committee meeting, the small subset of DHH members experienced accessibility challenges routinely over almost three years and published an experience report [26]. The authors present a large list (17) of recommendations for mixed local/remote meetings, and two relevant ones are highlighted here: that it is best to recruit a DHH individual fluent in sign language to judge any video-calling technologies, and that even state-of-the-art videoconferencing solutions present challenges and often require expert troubleshooting.

Researchers also investigated the potential of adding sign language interpreters to online videos in a "picture-in-picture" manner [12, 19]. A user evaluation with 19 DHH participants found that participants preferred the "tracked" mode (wherein the video of the interpreter moved vertically to align it with relevant information) to a static implementation. This preference reflects the need for Deaf individuals to control where the interpreter is located relative to their surroundings.

Our work builds upon this prior work, by proposing a novel design for remote sign language conversations. Its design is based upon many of the shortcomings and observations found in this prior work.

## 2.2 Intelligibility of Sign Language Video

Tran *et al.* [25] used a web-based platform to evaluate the intelligibility of sign language videos transmitted at low frame and bitrates. The authors explored different metrics: their own "Human Signal Intelligibility Model", as well as peak signal-to-noise ratio (PSNR) [24]. In another study, the authors [23] found that participants were able to converse intelligibly over fairly low frame rates (<25 fps) and bitrates (<100 kbps), suggesting that degraded video quality (within limits) does not necessarily translate into an unintelligible conversation.

Another team of researchers [27] investigated whether novice sign language learners would be able to converse over a mobile device. By manipulating the video resolution along with varying the sign handshape[7], the researchers were able to conduct an effective experiment which showed that even novice signers were able to reproduce the most complex signs with ease when viewed on a smartphone.

One team [6] investigated whether a dynamic video compression scheme with lossless compression around regions of interest such as the face could boost intelligibility of sign language on low-bitrate mobile devices. Inspired by prior work which used eye-tracking to validate the region of interest (face), the researchers sacrificed video quality in other regions. Study participants preferred the dynamic scheme as it enabled them to see facial expressions and assisted with comprehension.

Prior work also exists on the algorithmic intelligibility of sign language videos. State-of-the-art sign language recognition and translation systems leverage computer vision and deep learning. They achieve about 40% word (or sign) error rates on continuous recognition tasks when tested on signers not in the training set [4]. (As a result, live interpreting is the current standard accommodation.) The videos in such corpora are taken from a frontal view of the signer. Perhaps most relevant to our current work is one previous project on algorithmic accuracy of sign recognition from a top-down view [21], which achieved comparable accuracy from this view. In particular, the authors noted that the loss of facial information (from the top-down view) was not as crucial as previously believed.

To this space, we contribute an initial understanding of the human intelligibility of a top-down view of the signer, as well as the intelligibility of an interpreter displayed in augmented reality glasses for the signer.

# 3 RQ1) THE HANDS-FREE PROTOTYPE

In this section, we describe our iterative design process, outline design criteria, and present our hands-free device prototype. Based on the team's domain expertise, personal experiences, and related work, we present design criteria for a hands-free remote interpreting device. Our final design, comprising augmented reality glasses and a hat with a mounted fisheye camera and speaker/mic, was created to meet these criteria.

## 3.1 Design Process

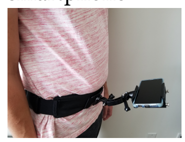
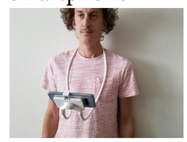
Our design team covered the primary stakeholders of remote interpreting technology: DHH signers, hearing people, and interpreters. The team also had expertise in human-centered computing, design, and systems building, as well as first-hand interpreting experience in numerous scenarios from all three perspectives (DHH, hearing, and interpreter).

The team engaged in an iterative design process, during which the team experimented with various hardware and software. We examined components individually (e.g., specific cameras) to evaluate viability. From viable components, we composed several full systems for end-to-end interpreting. We tested these systems with various subsets of our team (Deaf person, hearing person, and interpreter) in various workplace scenarios (conversations that are stationary, mobile, or interactive).

In thinking critically about whether the iterative designs were suitable for hands-free remote interpreting, we identified a set of design criteria, outlined subsequently, which define the components and properties. These design criteria helped us evaluate our iterative designs, as presented in Table 1. Each of the four system designs presented in the table was created in full, and tested out with a

[7] A handshape is a distinctive configuration of the hands as the signer uses them in a sign language in conjunction with other components (e.g. orientation, movement, and location) to form words. [13]

**Table 1: Analysis of design criteria for designs explored during our iterative design process. Each design was built in full, and tested with a team of a DHH person, a hearing conversation partner, and remote interpreter.**

| | | Wrist-mounted smartphone  | Belt-mounted smartphone  | Neck-mounted smartphone  | Smartglasses + hat-mounted camera, speaker, and mic (photographed in Figure 3a) |
|---|---|---|---|---|---|
| Does the device capture: | the DHH person's complete, unrestricted signing space (face, torso, hands)? | No (one hand holds/mounts the phone) | No (mount restricts movement, camera misses lower torso) | No (mount restricts movement, camera misses lower torso) | Yes (via hat-mounted camera) |
| | semantically relevant environment (e.g., objects pointed to)? | No (camera captures signer or referenced object) | No (camera captures only signer) | No (camera captures only signer) | Yes (via hat-mounted camera) |
| Is the device's display: | clear to the DHH person (i.e. understandable without strain)? | Somewhat (display is small but can move closer) | No (display is small and cannot be moved closer) | No (display is small and cannot be moved closer) | Yes (via scaled display on glasses) |
| | in the DHH person's natural line of sight (so they can always see the interpreter)? | No (need to look down at arm) | Somewhat (need to shift attention somewhat) | Somewhat (need to shift attention somewhat) | Yes (by projecting on glasses) |
| Is the overall device: | lightweight? | Yes (phone plus cloth mount) | Somewhat (phone plus large hard mount) | Somewhat (phone plus large hard mount) | Yes (hat with lightweight electronics and glasses) |
| | inconspicuous? | Somewhat (mainstream items, unusual usage) | No (large unusual mount) | No (large unusual mount) | Somewhat (mainstream items, outfitted uniquely) |

DHH person, hearing conversation partner, and remote interpreter. The analysis of each system design presented in the table is the result of that testing process.

## 3.2 Design Criteria

During the iterative design process, our team defined a set of design criteria outlining the minimal components for a successful system, as well as required full-system properties.

**Components:** A hands-free remote interpreting device must cover the four endpoints of the interaction – capturing and producing sign language, and capturing and producing spoken language. As such, the minimum required components of a hands-free device are:

- A *camera* to capture the signer's hand motions, body movements, and facial expressions, which are streamed to the remote interpreter.
- A *microphone* to capture the audio from the hearing individual, which is streamed to the remote interpreter.
- A *display* for the video stream from the remote interpreter, which would be viewed by the DHH individual.
- A *speaker* to emit audio streamed from the remote interpreter, which would be heard by the hearing individual.

**Functionality:** In testing out various solutions during the iterative design process, we gained further insight into the functional requirements for the end-to-end system. While the system must comprise the four components above, their composition must also allow for natural conversation (e.g., by capturing all the visuals that the interpreter needs, and not restricting the DHH person's signing space). Because the interpreter and hearing person's requirements can easily be solved with existing solutions (e.g., an interpreter station with a large monitor, and a speaker/mic), we focus on criteria for the DHH person's experience.

These functional requirements are presented in Table 1, separated into requirements for the capture, display, and overall system. The table also compares different designs generated and tried during the iterative design process.

## 3.3 Hands-Free Prototype Design

To meet our design criteria, we present a novel hands-free device prototype (see Figure 2). It is the result of an iterative design process (described above).

The design enables DHH individuals to use both their hands (and arms) by equipping them with a hat and smartglasses, which together handle capturing and presenting both signed and spoken language. We chose a standard baseball hat for its commonness (to help the DHH signer blend in while wearing the device), and

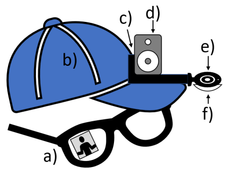



**Figure 2: Diagram of our hands-free device prototype: a) smart glasses that display the interpreter signing to the DHH person, b) hat, c) 3D-printed mount to hold components together, d) microphone/speaker for capturing the hearing person's speech and outputting the interpreter's speech to the hearing person, e) WiFi camera to capture the DHH signer, f) fisheye lens to capture the full signing space.**

for its long front bill. The length of the bill allowed us to mount a downward-facing camera far enough from the body that it could capture the full signing space, including the face. To help ensure that the full signing space is captured, we added a fisheye lens to the camera, which extends the camera's range of capture further. The hat also handled audio input and output through a mounted speaker. To ensure light weight, we chose components with minimal weight (e.g., a miniature speaker/mic, camera, and fisheye lens), and designed our mount to be lightweight as well.

The smartglasses provide the video display of the interpreter for the DHH individual. By locating the display in one eye of the glasses, it is always accessible and visible to the signer no matter where they turn their head, does not obstruct their view of the surroundings, and can be appropriately sized. We chose smartglasses that resemble ordinary glasses, to help with the unobtrusiveness of the overall design, both for passerby and for the hearing person, from whom the display of the interpreter is hidden. Unlike other AR/VR alternatives (e.g., Microsoft HoloLens or Oculus Rift), the smartglasses are also light-weight and leave most of the signer's face visible to the hearing person for eye contact and social cues.

## 3.4 Hands-Free Prototype Implementation

The physical hands-free prototype is illustrated in Figure 3, while the views from the camera (relative to a smartphone) are shown in Figure 4. The prototype was constructed out of six pieces of hardware (with letters corresponding to the parts in Figure 2):

- **a)** Vuzix Blade smartglasses[8]
- **b)** Baseball hat with sturdy, curved brim strong enough to hold the camera/lens/speaker[9]
- **c)** Custom 3D-printed mount to combine the fisheye lens with the camera and attach it to the hat (see Figure 3b)
- **d)** INSIQ portable bluetooth speaker/mic[10]
- **e)** ELValley 1080p spy-camera with Wi-Fi[11]
- **f)** NELOMO 230 degree fisheye lens[12]

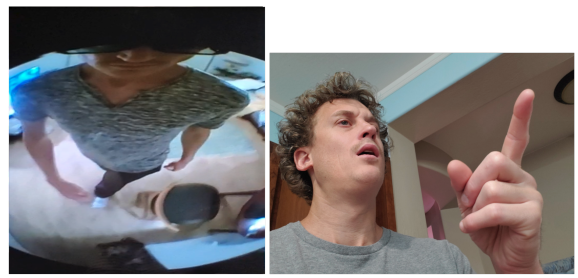
The smartglasses and the camera connected to Wi-Fi internet, to stream data to/from the interpreter station. The smartglasses

[8]https://www.vuzix.com/products/blade-smart-glasses
[9]https://www.amazon.com/dp/B07F3C1KZL
[10]https://www.amazon.com/dp/B07FCTG6SW
[11]https://www.amazon.com/dp/B07HD45MY8
[12]https://www.amazon.com/dp/B075N3KHSC

(a) Final hands-free device prototype

(b) Close-up view of 3d-printed hat mount for fisheye camera and speaker/mic

**Figure 3: Photographs of our hands-free device prototype**



(a) Top-down fisheye view from the hands-free device

(b) View from the smartphone, held in the DHH person's left hand

**Figure 4: Comparison of views of the DHH individual from the hands-free device and smartphone.**

connected to the interpreter using the Zoom video meeting app. The Zoom call was projected on the glasses' computerized display, located on one eye of the glasses. The camera used the standard IP-Camera protocols which enabled us to display the video stream on a separate monitor. The batteries in the components were as small and light as possible, and enabled the hands-free device to run for about an hour (which was sufficient for our studies).

# 4 RQ2) INTERPRETABILITY TEST

While the hands-free prototype enabled DHH individuals to sign naturally with both hands, it also introduced other changes that could affect the intelligibility of the conversation. It required the interpreter to view the conversation from above, and for the DHH individual to view the interpreter using smartglasses. To help quantify the impact of these changes, we performed an exploratory test of how well words could be conveyed and understood using the hands-free device, relative to a baseline in which the DHH individual performed one-handed signing with a smartphone.

## 4.1 Procedure

We tested interpretability by connecting a DHH individual to a remote interpreter under a variety of settings. We varied four aspects of the interaction:

- Device: Hands-free or smartphone.
- Activity: Standing or walking.
- Direction: DHH signing to interpreter or vice versa.
- Person interpreting: One of two professional interpreters.

For each combination of these factors, 100 individual words were signed in one direction (either by the DHH individual or the remote interpreter) for a total of 1600 signs in the experiment. The list of words to be communicated was constructed in advance (details below) and printed on a piece of paper for the sender to read and translate into ASL.

To test whether each sign was understood as intended, the receiver was asked to repeat each sign that they observed. Videos were recorded from a natural vantage point of both the sender and receiver. Subsequently, a different professional interpreter inspected the video recordings and determined whether the same sign was used by both the sender and the receiver. (We did not check the "correctness" of the sender's sign relative to the word list; so long as the same sign was expressed by the sender and receiver, we considered the sign to be communicated and understood correctly.)

The words used for the experiment were drawn from a popular ASL textbook [20]. First, we created a master word list consisting of the first 400 words from the textbook. Then, we created two copies of this list, each in a random order. One copy was used by the DHH individual when they were the sender, while the other copy was used by the interpreters when they were the sender. Each consecutive block of 100 words was used for a different combination of device and activity.

## 4.2 Physical Setup

The study was run in-lab. The study tasks took place in a single office room (standing condition) and the adjacent hallway (walking condition). During the study, the interpreter was situated in a room adjacent to the user study room to minimize connectivity issues such as bandwidth, latency, and dropped calls. The setup mimicked a VRS/VRI call center, with multiple monitors and a comfortable chair.

## 4.3 Participants

The DHH individual was the same in all cases, and was a research team member. The interpreters were the same as those in our broader user study (described later in RQ3). Because the interpretability test was performed near the end of the user study, both the DHH individual and the interpreters already had a few hours of experience communicating via the device. We counter-balanced the order of the devices between the interpreters, but kept the other orderings constant for simplicity (DHH signs to interpreter first, DHH stands first).

## 4.4 Results

The results of the test appear in Table 2. Overall, the hands-free device shows good accuracy. For signs viewed in the smart glasses (the upper left quadrant of the results table), approximately 96% are understood correctly. The most challenging case is signs conveyed through the hat camera (upper right quadrant of results), in which 83% of signs were understood on average. The results show that signing into a smartphone is also challenging: in part due to the requirement of signing with only one hand, only 88% of such signs are understood by the interpreter. The most reliable case is when signs are viewed on a smartphone. Out of 400 trials of this condition, only one mistake was logged.

**Table 2: Number of words understood correctly, out of a max of 100, for each condition in the interpretability test. The scores shown represent the average of two interpreters, while the range (difference between them) is shown in parentheses.**

| | | Interpreter signing | DHH Signing |
|---|---|---|---|
| Hands-free | standing | 85.5 (7) | 97 (0) |
| | walking | 80.5 (1) | 94.5 (3) |
| Smartphone | standing | 90 (4) | 100 (0) |
| | walking | 86 (6) | 99.5 (1) |

**Table 3: Logistic regression predicting interpretability of remote signing. Abbreviations: Est. estimate, S.E. standard error. Significance codes: *** < .001, ** < .01, * < .05**

| Coefficient | Est. | S.E. | $z$ | $p$ | |
|---|---|---|---|---|---|
| interpreter | -0.034 | 0.186 | -0.186 | 0.852 | |
| hands-free device | -0.633 | 0.191 | -3.302 | <.001 | *** |
| DHH sender | -2.014 | 0.259 | -7.763 | <.001 | *** |
| walking | -0.419 | 0.188 | -2.222 | 0.026 | * |

To assess which variables in our 2x2x2x2 design impacted the interpretability of remote signs, we conducted a logistic regression. The resulting model is shown in Table 3. The results indicate that there are significantly more mistakes with the hands-free device than with smartphone. There are also significantly more mistakes when the DHH individual is signing to the interpreter (via the overhead camera) than when the DHH individual is viewing signs from the interpreter (via the smartglasses). There is a slightly smaller but significant effect of walking, with more errors exhibited when the DHH individual was walking as opposed to standing. The two interpreters did not have significantly different results.

There were some common patterns regarding which words were misunderstood. Of the 68 misunderstandings stemming from via the hat camera, 29% were words misunderstood by both interpreters. Some of these words used hands positions that were challenging to see via the hat. For example, the word "father" (which stems from the forehead) was misunderstood, likely because the hand placement and motion was not completely in view. Similarly, the word "banana", misunderstood by both interpreters, includes an upright finger and up/down peeling motion that was difficult to see from above.

Of the 48 misunderstandings arising from one-handed signing into a smartphone, 46% were common across interpreters. The most common cause of misunderstanding was that truncating a sign from one hand to two can make it ambiguous. For example, the sign for "clock" was interpreted as "need" by both interpreters; these signs are the same with one hand, but different with two hands.

Of the 17 misunderstandings arising from display of signs in the smartglasses, none of the misunderstood words were common across interpreters. However, several of these words were displayed when the DHH individual was turning around (they walked back and forth in a short hallway, requiring frequent turns). The rapidly changing background behind the glasses, coupled with the cognitive overhead of shuffling positions, may have contributed to these misunderstandings.

## 4.5 Discussion

While the hands-free device resulted in more misunderstandings than the smartphone, it nonetheless led to relatively high accuracy rates. About 96% of words viewed in the smartglasses were understood, while 83% of words signed into the hat camera were understood. It is important to note that recognizing individual words, without conversational context, is inherently difficult and rarely necessary. Just as individual spoken words are often difficult to understand over a telephone, but easily understood in context, we believe the same is true of words signed via our hands-free device. This exploration was intended as a stress test to probe the relative strengths and weaknesses of each approach. In everyday conversations, we believe the observable differences in interpretability will be far less than we have outlined here. We turn our attention to such real-world scenarios in the next section.

# 5 RQ3) USER STUDY

To explore the trade-offs of using our hands-free device for DHH and hearing users, we ran an in-lab user study with 18 participants (9 DHH, 9 hearing). Continuing our focus on equal access to the workplace, the study focused on three typical workplace interactions between a DHH and hearing person.

## 5.1 Procedure

Our experiment had a within-subjects design, and took about one hour. Each participant experienced both conditions in random order: baseline (smartphone in one hand), and our hands-free device. The main part of the experiment consisted of the participant (either DHH or hearing) in one-on-one conversation with an actor (hearing for DHH participants, and DHH for hearing participants), with the DHH individual using either the smartphone or hands-free device for interpreting.

**Consent and Demographics (15 min):** The study began with obtaining consent, followed by a 15-minute survey covering basic demographics. DHH participants were also asked about their past experience with VRI and VRS, while hearing participants were asked about prior conversations with DHH individuals and interpreting services more generally.

**Main Tasks (30 min):** Next, each participant participated in our three main tasks, in the following order. They completed the activities twice: once with the hands-free device and once with the smartphone baseline, in random order.

(1) Standing/sitting and chatting (3 minutes)
(2) Walking around the office and chatting (3 minutes)
(3) Playing a game of Pictionary on a whiteboard (5 minutes)

Our study consisted of three main conversational tasks, to enable us to evaluate our hands-free device in a diversity of use cases. The tasks are designed to probe the following fundamental components of face-to-face conversation: **eye contact** - a fundamental way people connect during conversation; **mobility** - the ability to walk or move around while conversing; and **attentivity** - the ability for either person to get the other's attention. We selected three common workplace tasks which cover these components (summarized in Table 4): **1) stationary chat**, **2) walking chat**, and **3) whiteboard activity**.

For the whiteboard activity, we picked the game "Pictionary", as it involves many conversational turns. The game consists of two players taking turns drawing something, while the other player tries to guess what they are drawing. Performing this activity on a whiteboard closely mimics a typical team meeting (i.e., someone drawing a diagram or flow-chart, and others giving feedback or corrections on the fly). It also forces the participant who is at the whiteboard to simultaneously complete their drawing task and be aware of guesses from the other individual (seated behind them).

To help ensure consistency of experience across participants, we employed a single hearing actor and single DHH actor, both of whom were members of our research team. We hired two certified, professional sign language interpreters to take turns as the remote interpreter during these tasks.

**Exit Survey (15 min):** The study concluded with an exit survey, consisting of questions about their experience (using standard 5-point Likert scale questions) and which device they preferred along various aspects of interaction. We also gave them the opportunity to express open-ended opinions and feedback on the hands-free device compared to the smartphone. To minimize communication barriers and misunderstandings, we provided longer questions in ASL video, in addition to English (as in [3, 9]). We also offered the option to provide feedback in interview format in ASL to our DHH researcher.

## 5.2 Physical Setup

The physical setup mirrored that of the interpretability test. The available room contained chairs for sitting during the stationary chat, if desired. On the room table, we placed a container of paper strips with random words[13] for Pictionary.

## 5.3 Participants

We recruited 18 participants (9 DHH, 9 hearing). DHH participants were recruited through social media and local DHH organizations. Recruitment criteria included having the ability to walk around and use both hands, and not needing to wear prescription glasses. Hearing participants were drawn primarily from within our organization and recruited through an internal email list. For participating in the study, each hearing participant received a $25 gift card, and each DHH participant received a $50 gift card (to compensate for the time and expense of travel in addition to the study itself).

Basic self-reported demographics for our participants were:

- **DHH (9):** 6 Deaf, 2 deaf, 1 Hard-of-Hearing; mean age 41.5 (SD=11.3); 5 male, 4 female; all had used VRS and VRI extensively.
- **Hearing (9):** mean age 33.45 (SD=9.7); 5 female, 4 male; 4 never conversed with a DHH individual before this study.

[13] Varying difficulty levels, taken from a Pictionary website: https://hobbylark.com/party-games/pictionary-words.

**Table 4: Relation between our three experiment tasks and fundamental components of face-to-face conversation.**

| | Eye Contact | Mobility | Attentiveness |
|---|---|---|---|
| **1. Stationary chat** (face-to-face, stationary) | YES | NO | NO |
| **2. Walking chat** (side-by-side conversation) | SOME | YES | SOME |
| **3. Whiteboard activity** (conversing from behind) | NO | SOME | YES |

## 6 RESULTS

Our user study results suggest that our hands-free device provided benefits for both DHH and hearing participants in terms of interpersonal interactions and freeing the signer to use both hands and move around, at the expense of familiarity and technical robustness.

### 6.1 Overall Preferences

First, we present high-level preference data from both DHH and hearing participants. Across DHH and hearing groups, participants preferred the hands-free device to the smartphone for various aspects of interaction: attention-getting of the other person, eye contact, mobility, and fun (see Figure 5). However, both groups preferred the smartphone to the hands-free device for ease of conversation, and slightly preferred the smartphone overall.

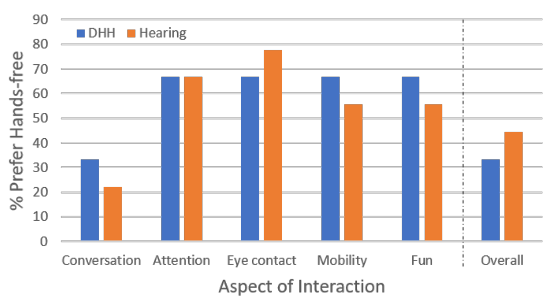



**Figure 5: Percent of participants (both DHH and hearing) who preferred the hands-free device to the smartphone, both overall and for various types of interaction (ease of conversation, attention-getting, eye contact, mobility, and fun).**

### 6.2 Usability Questions

Next, we present results from questions asked only of DHH participants, to help us better understand their experience with the device. Specifically, we asked the DHH participants to evaluate the usability of the hands-free device using several standard questions. Of these questions, two focused on the hands-free device (Net Promoter Score and Adjective System Usability Scale questions). We also asked three questions about each of the hands-free device and smartphone, with results given on a Likert scale (SUS-1, SUS-8, and a last question on discomfort).

In response to the Net Promoter Score (NPS) question [17]: "*How likely is it that you would recommend this system to a friend or colleague?*", our DHH participants scored[14] [$\mu$=7.11, SD=1.45], which indicates that the participants are "promoters" and were not detracted by the prototype. In response to the Adjective System Usability Scale [2]: "*Overall, I would rate the user-friendliness of this product as:*", the DHH participants replied "Good" on average [$\mu$=3.89, SD=0.78][15], which indicates that the hands-free device was perceived as user-friendly to our participants.

Participant responses to our two System Usability Scale [5][16] questions showed no significant difference between devices (by the Kruskal-Wallis rank sum test). For SUS-1) "*I think that I would like to use this system frequently*", there was no significant difference [$\chi^2$=1.6771, $p$=0.1953], and the Hands-free device scored [$\mu$=3.44, SD=0.881] while the Smartphone scored [$\mu$=3.89, SD=1.269]. For SUS-8) "*I found the system very cumbersome to use*", there was no significant difference [$\chi^2$=0.13387, $p$=0.7144], and the Hands-free device scored [$\mu$=3, SD=1.118] while the Smartphone scored [$\mu$=3.11, SD=1.167], which indicated that our participants did not find either system cumbersome to use.

Finally, we followed up with a 5-point Likert question on device comfortability: "*I experienced discomfort while using the device.*"[17] There was no significant difference between the Hands-free device and the Smartphone [$\chi^2$=0.002382, $p$=0.9611], both devices had the exact same score [$\mu$=3.22, SD=1.202], indicating that our participants did not feel more discomfort with one device than the other.

### 6.3 Comparison of the Devices

To analyze the free-response survey questions, three members of the research team (including the DHH and hearing actors from the experiment) independently reviewed all of the responses, assigned codes to key sentiments and grouped them into themes [22]. We then met as a group to compare our interpretations, reconcile differences, and reach a set of consensus themes. We analyzed a total of 113 comments (2,899 words) from both DHH and hearing participants. The resulting themes are summarized in Table 5, and explicated in the text, grouped by device.

### 6.4 The Smartphone

The main benefits participants found in using the smartphone were familiarity and good sound quality. The main downsides had to do with the signer having to hold the phone, and the obvious presence of a third party in the conversation.

*6.4.1 (+) Familiarity.* Several DHH participants reported that they were familiar with the smartphone, using it daily in their lives

[14]The NPS is a 11-point scale, ranging from 0 to 10.

[15]The SUS-Adj is a 7-point scale, ranging from "Worst Imaginable" to "Best Imaginable" with "OK" as the middle point.

[16]The SUS-1 and SUS-8 questions are answered on a 5-point scale, ranging from "Strongly disagree" to "Strongly agree".

[17]This last question was also answered on a 5-point scale from "Strongly disagree" to "Strongly agree".

**Table 5: A high-level overview of the comparison between the hands-free device and the smartphone with their pros and cons.**

| | Hands-free | | Smartphone | |
|---|---|---|---|---|
| | **Pro** | **Con** | **Pro** | **Con** |
| **DHH** | eye contact; ability to multitask; both hands free | poor brightness/contrast; eye health concerns; heavy; incompatible with glasses | familiarity | one hand for signing; having to hold phone; limited view for terp |
| **Hearing** | eye contact; mobility; personal connection due to "invisible" terp | lack of feedback; sound issues; obscured face of other person | familiarity | awkward third party involvement |

so they preferred it over the hands-free device: "*Experienced with Smartphone. Hands-free device is so new and possibly needs more time to experience with it.*" (P2 - DHH)

*6.4.2 (+) Good Sound Quality.* The hearing participants seemed to appreciate the higher fidelity audio coming from the smartphone (possibly because the DHH actor held the phone in front of them, thus having better audio transmission/reception compared to the mic/speaker on the hat): "*Sound was much clearer, so the conversation was never interrupted and flowed well.*" (P3 - Hearing)

*6.4.3 (–) Occupied Hand - 1-Hand Signing, Fatigue.* As expected, many of our DHH participants felt constrained by signing with only one hand and after some time experienced fatigue: "*Sometimes hard to sign with one hand. Heavy. Have to look up and down when walking.*" (P4 - DHH)

*6.4.4 (–) Limited Interpreter View.* Some DHH participants kept forgetting that they were mobile (as opposed to propping the smartphone on a fixed location) and it presented challenges for the interpreter to comprehend what they were signing when the DHH participant was out of view: "*Keeping myself in focus for an interpreter to see was a challenge while signing with one hand.*" (D2 - DHH)

*6.4.5 (–) Third Party Presence.* The mere fact that the DHH actor was looking at the smartphone made the hearing participants feel like there was a third party to the conversation that they were not privy to, and caused some social tension: "*It was not difficult but I didn't like the fact that there was a third person that he was looking at who I couldn't see.*" (P7 - Hearing)

## 6.5 The Hands-Free Device

The main benefits participants reported from using the hands-free device had to do with the DHH person having both hands free to sign or interact, and increased personal connection. The main drawbacks were technical or physical in nature.

*6.5.1 (+) Two Hands Free - Multitasking, Mobility.* Almost all of our DHH participants immediately appreciated the ability to sign with both hands and/or multitask (such as drawing on the whiteboard during the Pictionary phase of the study) with the hands-free device: "*I can using my ASL both my hands rather than holding my smartphone.*" (P9 - DHH)

*6.5.2 (+) Eye Contact.* Another benefit of the hands-free device was the awareness brought on by increased eye contact during conversations: "*I could see the interpreter and the hearing person at the same time. I always had the habit of signing directly to the interpreter, instead of the hearing person before using the hands-free device. Now, I could use my eye contact with the hearing person.*" (P7 - DHH)

*6.5.3 (+) Heightened Personal Connection.* Both DHH and hearing participants had a sense of personal connection when the hands-free device was being worn which was missing when the smartphone was being used: "*The hands-free device made the interaction smoother as it was easier to interact with each other via non-verbal interactions such as hand gestures or eye contact while conversing.*" (P4 - Hearing)

*6.5.4 (–) Brightness/Contrast & Health Concerns.* A few participants mentioned that it was hard to see the interpreter on the display in the smartglasses during certain scenarios: "*It was kind of hard to see the interpreter especially on a white board or through glass. It would be good if the screen is 100% visible, not blurry, just like a HDTV.* " (P7 - DHH) Also, some of them were worried about the smartglasses' technology being harmful to their eyes: "*Just have to be careful with eyes. I know prolong use of VR / hand-free device can affect eyes. I believe hand-free device is good idea and can use for short time if needed.*" (P6 - DHH)

*6.5.5 (–) Physicality - Weight, Glasses Compatibility, Hat Shadow.* Some participants struggled to have the smartglasses fit their face snugly or did not like the weight: "*I would like different style of glasses, that one you give me is wrong shape fit my face because my nose cannot hold it.*" (P1 - DHH)

*6.5.6 (–) Low Feedback to Hearing Person.* A few of the hearing participants felt "left out" when the hands-free device was used because they didn't know what was happening when there was delays with the interpreter's audio stream (because of technical or cognitive issues): "*I wasn't always sure when it was ok for me to talk or when I would be interrupting.*" (P2 - Hearing)

## 6.6 The Interpreter's Experience

To explore the experience of interpreting from a top-down fisheye view of the signer, we conducted a brief exit survey with the two interpreters who interpreted for our user study. We asked about the benefits and difficulties they experienced interpreting with each device, and within what contexts they would prefer to interpret with the hands-free device.

Both interpreters reported that it was easier to interpret the smartphone view. On a scale of 1 (very easy) to 5 (very hard), on average our interpreters rated the smartphone a 1.5 (1 and 2), and the hands-free device a 3 (2 and 4). It is possible that this difference in difficulty was due to the novelty of the hands-free view, compared to the familiar smartphone view.

The primary benefit interpreters reported from using the smartphone was the ability to see the DHH person's **face** as reported by both interpreters. (Conversely, limited facial view was the primary difficulty with the hands-free device.) As facial expressions can be semantically meaningful, they are important for interpreters to see. The smartphone provided a clearer view of the face when pointed at the signer's face, compared to the fisheye camera view, which captured the face, but did not provide a front-on view (see Figure 4).

The primary benefit interpreters reported from using the hands-free device was the ability to view the signer's surroundings (also reported by both interpreters). (Conversely, limited view of surroundings was the primary difficulty with the smartphone.) Context is essential to semantic meaning (e.g., a signer can point to a nearby object to refer to it), so being able to view the surroundings was beneficial. The hands-free device offered more environmental information than the smartphone, since the top-mounted camera provided a wider field of vision and moved with the signer.

Both interpreters reported that a preference for the hands-free device in various situations, based on the increased ability to view surroundings. Examples included public events and arenas, business and board room meetings, and health care. The interpreters illustrated why the hands-free device would be particularly beneficial in these situations: the signer would "*be able to point to the body if in pain*" (I1) and have the interpreter communicate the symptom to the doctor, and generally the interpreter "*being able to see who is in the room and what is happening would improve the Deaf person's experience.*" (I2)

## 7 DISCUSSION

It is possible that device inconspicuousness may actually introduce communication difficulties. One of the hands-free device's strengths is that it enabled more personal connections (e.g., increased eye contact and reduced interpreter presence). As the device approached "invisibility", it may have become easy to forget that an interpreter was still in the loop, which introduces delays. At the same time, as expectations of a natural interaction increased, sensitivity to delays may also have increased, as glitches in the hidden technology became more unexpected. In contrast, the smartphone reminded hearing participants of the interpreter's presence, and calibrated expectations for the time needed to translate each utterance. One participant described this mismatch: "The delay between signing and speech makes things a little tricky–this was true for both [devices], but for some reason the overlap was more noticeable for me and created more problems during the hands-free interaction."

Displaying the interpreter in the DHH person's line of sight may have heightened these problems by making it harder to discern whether the DHH person was waiting for the interpreter to finish signing or for the hearing person to speak. One hearing participant summarized: "I wasn't always sure when it was OK for me to talk or when I would be interrupting." In contrast, the smartphone makes it obvious where the DHH person's attention is directed.

It is also possible that DHH participants who tried our hands-free device had raised expectations about performance that backfired in their evaluation of the system. Because we presented them with a new device involving sophisticated-looking components (i.e., augmented reality glasses, and a 3D printed hat piece), their expectations might have been elevated for its performance. At the same time, the device they experienced was a prototype, and as such had occasional technical difficulties (e.g., WiFi connectivity issues), that detracted from the experience. It is possible that this mismatch in expectations accounts for most (>50%) of DHH and hearing participants preferring the hands-free device for 4 of 5 interaction aspects (Figure 5), yet stating an overall preference for the smartphone.

While reducing the burden for the DHH individual was a major motivation for this work, we note that our solution still puts the burden disproportionately on the DHH person (as opposed to the hearing person). The burden of accommodating a disability is often placed on the person with the disability, and it would be exciting to design solutions that distribute this burden more evenly. In the context of this work, we envision interpreting devices that distribute the burden equally across participants, or even put the burden solely on the hearing person. For example, it might be possible to design a comparable hat and glasses that are worn by the hearing person, rather than the DHH person. For this type of solution to work, the camera would have to track the DHH person as they move, and dynamically adjust the interpreter display to be in the signer's line of sight. The social implications of shifting the burden away from the DHH person would also be exciting to study.

There remain several opportunities to further improve our system. The interpreter's view of the DHH person is from an unusual perspective (overhead) and with an unusual lens (fisheye), as indicated by their feedback. In our interpretability test, these factors also contributed to increased misunderstanding of words seen through the hat camera compared to other configurations. In the future, it could be possible to improve the solution with additional image processing, e.g., to correct for the fisheye distortion. Furthermore, entirely alternative solutions to using a fisheye may be beneficial, for example angling or widening the camera lens, as the fisheye lens itself may be obtrusive to wear. It is also possible that setting up a training tool to help interpreters acclimate to the new viewpoint would be beneficial. More generally, studying the interpreter learning curve would also be interesting.

It would also be interesting to explore providing feedback to the hearing person on what the interpreter is doing, and when to speaking. Our user study revealed that many hearing participants thought the hands-free device was not working, when in fact the interpreter was still interpreting. They expected a working system to have no delays, not accounting for interpretation delays. These problems were likely exacerbated by many hearing participants' (33%) lack of prior experience conversing with DHH individuals. Providing feedback to the hearing person could help address these problems. Such feedback could be visual (e.g., a projection on the outward facing side of the smartglasses), or audio (e.g., playing a sound equivalent of a loading icon while the interpreter signs). Exploring and evaluating possible solutions makes exciting future work.

## 8 CONCLUSION

Remote sign language interpretation is a promising approach to bolster social inclusion of DHH individuals, who often lack reliable access to in-person interpreters in a given location. However, existing interfaces for remote interpreting are awkward at best, requiring DHH individuals to devote one of their signing hands to conducting a video call on a smart phone.

In this work, we bring a fresh HCI lens to the problem of remote video interpretation, with a special focus on enabling independence, mobility, and natural expression on the part of DHH individuals. After exploring various points in the design space, we prototyped a novel hands-free solution by piecing together several off-the-shelf components. Our exploratory evaluation suggests that a hands-free device may offer benefits, including enabling DHH signers to multitask while engendering more personal and direct connection with hearing conversation partners. Our evaluation of interpretability shows that DHH people and interpreters can understand each other, even if there is still room for improvement.

Variants of our hands-free device could be used to facilitate more natural conversation (or multitasking) in other use cases. For example, a sign language user could converse with any remote sign language user, using ASL exclusively instead of translating to spoken language. In the future, it would be beneficial to the research community to investigate whether alternative uses of the hands-free device would also facilitate novel conversation modalities.

## REFERENCES

[1] 1990. Americans with Disabilities Act of 1990. Pub. L. No. 101-336, 104 Stat. 328.
[2] Aaron Bangor, Philip T. Kortum, and James T. Miller. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies* 4, 3 (2009), 114–123.
[3] Larwan Berke, Matt Huenerfauth, and Kasmira Patel. 2019. Design and Psychometric Evaluation of American Sign Language Translations of Usability Questionnaires. *ACM Transactions on Accessible Computing* 12, 2, Article 6 (June 2019), 43 pages. https://doi.org/10.1145/3314205
[4] Danielle Bragg, Oscar Koller, Mary Bellard, Larwan Berke, Patrick Boudrealt, Annelies Braffort, Naomi Caselli, Matt Huenerfauth, Hernisa Kacorri, Tessa Verhoef, et al. 2019. Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective. *arXiv preprint arXiv:1908.08597* (2019).
[5] John Brooke. 1996. SUS-A quick and dirty usability scale. *Usability Evaluation in Industry* 189, 194 (1996), 4–7.
[6] Anna Cavender, Richard E. Ladner, and Eve A. Riskin. 2006. MobileASL:: Intelligibility of Sign Language Video As Constrained by Mobile Phone Technology. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility* (Portland, Oregon, USA) *(Assets '06)*. ACM, New York, NY, USA, 71–78. https://doi.org/10.1145/1168987.1169001
[7] Anna C. Cavender, Jeffrey P. Bigham, and Richard E. Ladner. 2009. ClassInFocus: Enabling Improved Visual Attention Strategies for Deaf and Hard of Hearing Students. In *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, Pennsylvania, USA) *(Assets '09)*. ACM, New York, NY, USA, 67–74. https://doi.org/10.1145/1639642.1639656
[8] Jan Gugenheimer, Katrin Plaumann, Florian Schaub, Patrizia Di Campli San Vito, Saskia Duck, Melanie Rabus, and Enrico Rukzio. 2017. The Impact of Assistive Technology on Communication Quality Between Deaf and Hearing Individuals. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. ACM, New York, NY, USA, 669–682. https://doi.org/10.1145/2998181.2998203
[9] Matt Huenerfauth, Kasmira Patel, and Larwan Berke. 2017. Design and Psychometric Evaluation of an American Sign Language Translation of the System Usability Scale. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore, Maryland, USA) *(ASSETS '17)*. ACM, New York, NY, USA, 175–184. https://doi.org/10.1145/3132525.3132540
[10] G. Hughes, B. Hudgins, and J. MacDougall. 2004. Remote sign language interpretation using the Internet. In *Proceedings of the Second Annual Conference on Communication Networks and Services Research*. 345–350. https://doi.org/10.1109/DNSR.2004.1344750
[11] Jouko Kokko, Erkki Kemppainen, and Aulikki Rautavaara. 2007. Technology and Regional Social Structures: Evaluation of Remote Sign Language Interpretation in Finland. In *Universal Acess in Human Computer Interaction. Coping with Diversity*, Constantine Stephanidis (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 716–721.
[12] Raja Kushalnagar, Matthew Seita, and Abraham Glasser. 2017. Closed ASL Interpreting for Online Videos. In *Proceedings of the 14th Web for All Conference on The Future of Accessible Work* (Perth, Western Australia, Australia) *(W4A '17)*. ACM, New York, NY, USA, Article 32, 4 pages. https://doi.org/10.1145/3058555.3058578
[13] Scott K. Liddell. 2002. *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge University Press, Cambridge, UK.
[14] Ashley Miller, Joan Malasig, Brenda Castro, Vicki L. Hanson, Hugo Nicolau, and Alessandra Brandão. 2017. The Use of Smart Glasses for Lecture Comprehension by Deaf and Hard of Hearing Students. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI EA '17)*. ACM, New York, NY, USA, 1909–1915. https://doi.org/10.1145/3027063.3053117
[15] Ross E. Mitchell, Travas A. Young, Bellamie Bachleda, and Michael A. Karchmer. 2006. How many people use ASL in the United States? Why estimates need updating. *Sign Language Studies* 6, 3 (2006), 306–335.
[16] Jemina Napier, Laurence Meurant, Aurelie Sinte, Mieke Van Herreweghe, and Myriam Vermeerbergen. 2013. *"You get that vibe": A pragmatic analysis of clarification and communicative accommodation in legal video remote interpreting* (1 ed.). De Gruyter, 85–110. http://www.jstor.org/stable/j.ctvbkk4dr.7
[17] Frederick F. Reichheld. 2003. The One Number You Need to Grow. *Harvard business review* 81, 12 (2003), 46–55.
[18] V. Rusňák, P. Troubil, D. Daxnerová, P. Kajaba, M. Minárik, S. Ondra, T. Sklenák, and E. Hladká. 2016. CoUnSiL: A video conferencing environment for interpretation of sign language in higher education. In *2016 15th International Conference on Information Technology Based Higher Education and Training (ITHET)*. 1–8. https://doi.org/10.1109/ITHET.2016.7760711
[19] Matthew Seita. 2016. Closed ASL Interpreting for Online Videos. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (Reno, Nevada, USA) *(ASSETS '16)*. ACM, New York, NY, USA, 337–338. https://doi.org/10.1145/2982142.2982147
[20] Cheri Smith, Ella Mae Lentz, and Ken Mikos. 2008. *Signing Naturally: Student Workbook, Units 1-6*. DawnSignPress, San Diego, CA, USA.
[21] Thad Starner, Joshua Weaver, and Alex Pentland. 1998. Real-time american sign language recognition using desk and wearable computer based video. *IEEE Transactions on pattern analysis and machine intelligence* 20, 12 (Dec 1998), 1371–1375. https://doi.org/10.1109/34.735811
[22] Anselm Strauss and Juliet M. Corbin. 1998. *Basics of qualitative research: Grounded theory procedures and techniques* (2 ed.). SAGE Publications, Inc., Thousand Oaks, CA, USA.
[23] Jessica J. Tran, Ben Flowers, Eve A. Risken, Richard J. Ladner, and Jacob O. Wobbrock. 2014. Analyzing the Intelligibility of Real-time Mobile Sign Language Video Transmitted Below Recommended Standards. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility* (Rochester, New York, USA) *(ASSETS '14)*. ACM, New York, NY, USA, 177–184. https://doi.org/10.1145/2661334.2661358
[24] Jessica J. Tran, Joy Kim, Jaehong Chon, Eve A. Riskin, Richard E. Ladner, and Jacob O. Wobbrock. 2011. Evaluating Quality and Comprehension of Real-time Sign Language Video on Mobile Phones. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility* (Dundee, Scotland, UK) *(ASSETS '11)*. ACM, New York, NY, USA, 115–122. https://doi.org/10.1145/2049536.2049558
[25] Jessica J. Tran, Rafael Rodriguez, Eve A. Riskin, and Jacob O. Wobbrock. 2013. A Web-based Intelligibility Evaluation of Sign Language Video Transmitted at Low Frame Rates and Bitrates. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (Bellevue, Washington) *(ASSETS '13)*. ACM, New York, NY, USA, Article 2, 8 pages. https://doi.org/10.1145/2513383.2513432
[26] Christian Vogler, Paula Tucker, and Norman Williams. 2013. Mixed Local and Remote Participation in Teleconferences from a Deaf and Hard of Hearing Perspective. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (Bellevue, Washington) *(ASSETS '13)*. ACM, New York, NY, USA, Article 30, 5 pages. https://doi.org/10.1145/2513383.2517035
[27] Kimberly A. Weaver, Thad Starner, and Harley Hamilton. 2010. An Evaluation of Video Intelligibility for Novice American Sign Language Learners on a Mobile Device. In *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility* (Orlando, Florida, USA) *(ASSETS '10)*. ACM, New York, NY, USA, 107–114. https://doi.org/10.1145/1878803.1878824