Wearable Subtitles: Augmenting Spoken Communication with Lightweight Eyewear for All-day Captioning

Alex Olwal Kevin Balke Dmitrii Votintcev Thad Starner Paula Conn Bonnie Chinh Benoit Corda

Google Research
Mountain View, CA, USA
olwal@acm.org



Figure 1. Our Wearable Subtitles proof-of-concept shows how eyewear could benefit people who are deaf or hard of hearing. We explore hands-free access to spoken communication, situational and speaker awareness, and improved understanding while engaged in a primary task. Our lightweight (54 g) 3D-printed eyewear prototype augments the user's perception of speech and sounds in a socially acceptable form factor with an architecture that could enable up to 15 hours of continuous transcription.

ABSTRACT

Mobile solutions can help transform speech and sound into visual representations for people who are deaf or hard-of-hearing (DHH). However, where handheld phones present challenges, head-worn displays (HWDs) could further communication through privately transcribed text, hands-free use, improved mobility, and socially acceptable interactions.

Wearable Subtitles is a lightweight 3D-printed proof-of-concept HWD that explores augmenting communication through sound transcription for a full workday. Using a low-power microcontroller architecture, we enable up to 15 hours of continuous use. We describe a large survey (n=501) and three user studies with 24 deaf/hard-of-hearing participants which inform our development and help us refine our prototypes. Our studies and prior research identify critical challenges for the adoption of HWDs which we address through extended battery life, lightweight and balanced mechanical design (54 g), fitting options, and form factors that are compatible with current social norms.

Author Keywords

Assistive technology; hearing accessibility; wearables; head-worn displays; captions; low-power system; all-day

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

UIST '20, October 20–23, 2020, Virtual Event, USA
© 2020 Copyright is held by the owner/author(s).
ACM ISBN 978-1-4503-7514-6/20/10.
https://doi.org/10.1145/3379337.3415817

CSS Concepts

• **Human-centered computing~Accessibility technologies**; HCI; Mobile Computing; User Studies

INTRODUCTION: CAPTIONING FOR ACCESSIBILITY

In recent years, various technologies have been developed to help transform speech and sound into visual representations, which provide benefit to the deaf and hard-of-hearing population. While applications such as Live Transcribe [9] and Microsoft Translator [28] are available across desktop and mobile devices, they do not enable captioning within the line of sight of interlocutors [16]. Research has shown that the visual dispersion from attending handheld and desktop devices can impact mental effort and attention, causing captioned and non-captioned information to be missed or misunderstood [24, 29].

In an effort to reduce visual dispersion, various research teams have investigated the use of captions on custom [19] and commercially available head-worn displays (HWDs) [16, 29, 35, 42]. These studies corroborate surveys indicating that HWDs for captions are preferred over smartphones and other wearable technologies [3]. HWDs reduce the need to realign captions to the main point of visual attention [29], increase the perceived emotional connection to the interlocutor [16], and enable environmental awareness [16]. HWDs also yield promising results when used to assess training comprehension [43].

Profita et al. found HWDs to be socially acceptable as assistive technologies from the interlocutor and bystander perspectives [33]. This acceptance contrasts that of mobile devices during in-person conversations, which can reduce bystander social comfort [1], even when the conversation

partner is not actively engaging with the device [30]. Smartphone presence may lessen conversational quality and reduce empathic exchange [30].

Despite the benefits of HWDs for captioning, current commercially available solutions inhibit mobility and are not socially acceptable due to their poor fit or large form factor [16, 29]. Thus, a major challenge is developing a HWD with interactions and human factors that are compatible with the social contexts of everyday life, while also minimizing visual dispersion. Providing a socially acceptable design is equally important, as d/Deaf and hard-of-hearing individuals have been found to discontinue use of technologies (e.g., hearing aids [26]) if they do not align with their self-identity or aesthetic preferences.

Our HWD prototype augments the user's perception of speech and sounds in the environment (Figure 1). The 3D-printed proof-of-concept system is designed to be lightweight (54 g) and low power, with up to 15 hours of continuous transcription of speech and environmental sounds. The battery life allows for active use throughout a workday, transcribing speech into private captions.

CONTRIBUTIONS

The major contributions of this work include

- *A low-power modular architecture* to enable all-day active streaming of transcribed speech in a lightweight, socially-unobtrusive HWD.
- *Technical evaluation and characterization* of power, bandwidth and latency.
- *Usability evaluations* in a pilot and two studies with 24 deaf and hard-of-hearing participants to understand the physical and social comfort of the prototype in a range of scenarios, which align with a large-scale survey of 501 respondents.

In the following sections, we review related work suggesting initial design requirements for this project. Next, we discuss a survey regarding the situations where a captioning HWD might be most useful for people who are deaf or hard of hearing. The subsequent sections describe the iterative development of the prototype and three related user studies. We conclude with current limitations and future avenues for this work.

RELATED WORK

Technologies for the d/Deaf and Hard of Hearing

While we acknowledge that not all deaf/hard-of-hearing (DHH) individuals want to use sound or captioning technologies, prior work has demonstrated that many people would find such technologies desirable and useful in everyday activities [16]. A common real-time captioning service used today is Communication Access Real-Time Translation (CART), which is typically displayed on a computer screen and can be expensive if needed for all-day use [31]. Hubert Upton made one of the first attempts at designing a HWD for people who are hard of hearing in 1967 [41]. Upton's analog wearable computer served as an aid for lip-reading. Using highpass and lowpass filters, the system determined if a spoken phoneme was a fricative, stop, voiced fricative, voiced stop, or simply voiced. LEDs mounted on ordinary eyeglasses illuminated, indicating the phoneme type [41].

Most prior work on technologies for captions or the display of sound awareness cues focus on stationary contexts and the use of commercially available devices [14, 17, 18, 24, 29, 32]. While subtle prototype HWD eyeglasses have existed since 1997 [20, 37], the public availability of HWDs has increased interest in head-up captioning in the past decade. Examples include HWDs that are specifically designed for use in movie theaters, where broader adoption in everyday life is limited by monochrome displays [34] or bulky form factors [11].

Academic evaluations of HWDs suggest that commercially available solutions do not have a suitable form factor for sustained captioning. Peng et al. [32] evaluated captioning designs during seated conversations by using a large LCD screen and later a Microsoft HoloLens device. Although early co-design sessions suggested that voice transcriptions could further social participation, the final usability ratings of the captioning designs were overshadowed by the limitations of the form factor and delay in the voice recognition software. Jain et al. [16] also explored captioning with HoloLens, demonstrating the value of captions in a mobile environment as well as the increased challenges. Use was limited to 20 minutes due to device weight (579 g). Additionally, the limited field of view and tinted visor made negotiating stairs difficult, and the design decreased social comfort. Participants did not want to call attention to themselves by wearing the device, consistent with prior autoethnographic research [14].

Captioning efforts on Google Glass [6] and Epson BT-200 frames [29] also highlight the limitations in current HWDs. While Glass can be worn comfortably for a full day (though battery life is <2h with display lit), it mounts the display high in the visual field, which is designed for short, glanceable interactions as opposed to extended reading. For reasons of visual dispersion, Miller et al. [29] switched from Glass to the Epson BT-200 but discovered that its weight on the nose was too uncomfortable. In addition, these devices are still very noticeable by spectators.

Social Acceptability of Wearable Assistive Devices

Having distinctive technology that signifies a disability, such as a white cane for people who are blind, can increase bystanders' acceptance of that technology [23, 33]. However, many people do not wish to call attention to their disability for fear of exploitation or discrimination [36]. Using wearable assistive technology can lead to undesirable effects [23, 33]. Aesthetics and perception are also an issue in device adoption. For example, eyeglasses may be more desirable than hearing aids due to the perception that hearing

aids are for older adults [36]. Thus, assistive technology that resembles mainstream devices is more accepted, especially by those with "invisible disabilities" [36]. Here, we focus on a form factor that is virtually unnoticeable to bystanders. While lack of noticeability does not necessarily mean improved social acceptance [22], it does reduce the social weight [40] with naive spectators.

Lightweight, Low-power, Head-worn Display Systems

In addition to physical and social comfort, prior work [16, 25] suggests that battery life is a pertinent measure for the usability of a captioning device. Matthews et al. [25] demonstrated the interest and value of receiving transcriptions in everyday activities and conversations, such as when in a group meeting, coffee shop, or other service-oriented businesses. Jain et al. [16] also identified the need for transcription during recreational or mobile activities, such as when exercising (yoga, hiking, kayaking), in transit (bus, car), and walking.

However, existing commercially available HWD systems were not designed to support these continuous usage scenarios as they run high level operating systems to support generic applications and drivers. Current lightweight HWDs thus use power-intensive mobile processors that are intended for intermittent engagement, such as Snapdragon XR1 (Google Glass EE2), quad-core Cortex-A53 (Vuzix Blade), and quad-core Cortex-A7 (North Focals 1.0). For these devices, speech recognition could be relegated to the smartphone and cloud as in our approach, but our informal experiments suggest that these systems only provide a few hours of battery life with the display on and otherwise idle. Blade ran 2.5h with 50% brightness; Focals 1.0 ran 2h at the mid brightness setting.

Comfort, of course, is a primary concern for eyewear that might be worn all day. The weights of the lightest devices with a discrete, line-of-sight display, Vuzix Blade (97 g; 56 g nose weight) and Focals 1.0 (69 g; 49 g nose weight) illustrate the challenges of meeting the suggested 75 g weight with a maximum 40% nose weight (30 g) [38].

Thermal comfort is another engineering issue due to the small volume of eyeglasses HWDs. Focals 1.0, which has the least powerful SoC, reaches a surface temperature of 44°C after five minutes of continuous use in informal tests [12].

Our eyewear prototype weighs 54 g with 30 g on the nose. With a total power consumption of ~266 mW, our device does not exceed body temperature (37°C), and we have not observed thermal discomfort in any of our user studies.

MOBILE SCENARIOS SURVEY: 501 RESPONDENTS

To learn more about challenges in mobile contexts, we conducted a brief large-scale online survey with participants who used hearing aids, TDD/TTY (telecommunications device for the deaf/teletypewriter) [13], CART [31], and cochlear implants. We aimed to expand upon existing studies which have already identified interest in real-time captioning [3, 15] while incorporating characteristics from moving conversations [16].

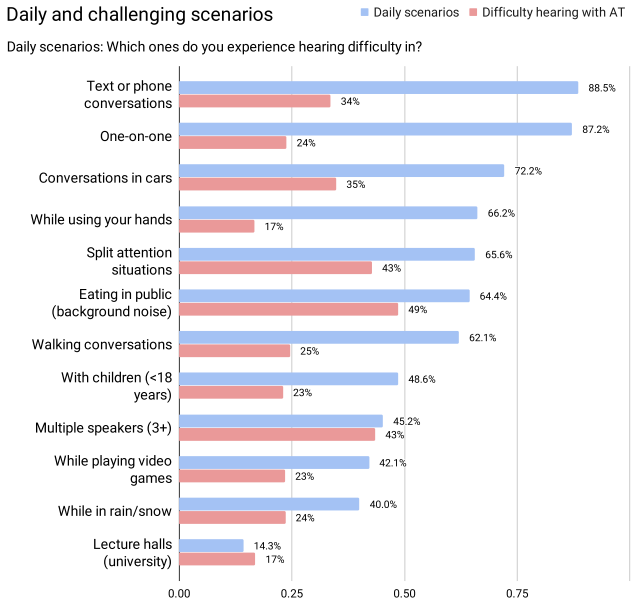

Figure 2. The 501 survey respondents who used hearing-related assistive technology (AT) indicated that activities involving eating in public and multiple speakers were likely to result in situations where it would be difficult to hear.

Survey Design

We used Google Surveys [7] to deploy a short, ten-question survey. The first four questions included the informed consent and background questions (assistive technology usage, gender, and communication preferences). Due to restrictions from our institution, we did not ask participants to self-report their hearing levels or inquire about the use of signed languages for this survey, which may have resulted in an underrepresentation of Deaf participants. The remaining six questions asked about scenarios that participants experienced on a daily basis and if it was difficult to hear in any of those scenarios. Scenarios presented in the survey were based on prior work that indicated communication difficulties due to high visual dispersion, low visibility, and increased ambient noise [14, 16, 29]. The scenarios aimed to represent these difficulties at varying levels of specificity to inform potential scenarios for user evaluations.

Recruitment

Participants were recruited from Google Opinion Rewards App users [8] who are compensated up to 1 USD for each completed survey in Google Play or PayPal credit.

Participants

We received 501 responses indicating usage of assistive technologies such as a hearing aid, cochlear implant, or real-time captioning/transcription services (i.e., CART). Participant ages ranged between 18–65, with 36% identifying as female. The types of assistive technologies varied, with 30% of participants using two or more: 40% of

participants indicated the use of a hearing aid, 7% used a cochlear implant, 53% used transcription services such as CART, and 83% used TDD/TTY.

Survey Results

The most common scenarios that participants experienced on a daily basis were text or phone conversations (88.5%), one-on-one conversations (87.2%), and conversations in cars (72.2%). Nearly half of the participants indicated daily experiences with almost all scenarios presented. The scenarios where the most participants indicated difficulty with communication were eating in public (49%), conversations with three or more speakers (43%), and split attention situations (43%). Overall, these scenarios represent moments of visual dispersion (e.g., *multiple speakers, while playing video games*), low facial visibility (e.g., *phone conversations, situations in rain/snow*), and ambient noise (*eating in public with background noise, conversations with children present*). These scenarios not only introduce communication challenges, but they can occur in a mobile context or while engaging in an activity with a dynamic environment. The frequency with which participants experience these scenarios and have difficulties with them suggest that they may serve as realistic opportunities for mobile user evaluations. See Figure 2.

Participants also indicated the perceived social acceptability of their assistive technologies. The majority of participants felt positive about the social acceptability of their technologies, with 70% of participants choosing *"Perfectly acceptable"* or *"acceptable"*.

Discussion and Limitations

Our large-scale survey enabled us to identify common mobile experiences, scenarios that frequently exacerbate hearing difficulties, and attitudes toward assistive technologies. Our scenarios present both specific and open-ended situations for future user evaluations which can focus on issues of visual dispersion, facial or text visibility, and ambient noise. The range of problematic scenarios provides a motivation for **more subtle, comfortable communication devices with all-day power to mitigate difficulties in everyday scenarios.** Participants' positive social acceptability ratings (70%) of their existing technology may also indicate a favorable outlook towards assistive technologies, while there is an opportunity to also improve the solution for less satisfied participants (30%).

However, we acknowledge limitations in our survey. By asking only about assistive technology to identify DHH users, as per our institution's allowance, Deaf participants who are less likely to use assistive technologies may have been underrepresented. For this case, we reference Findlater et al. [3], who have shown that 70.5% of their participants who preferred sign language as their form of communication showed interest in captions.

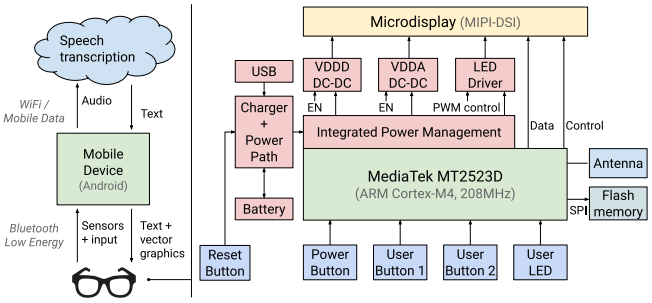

Figure 3. Left: Eyewear/phone/cloud thin-client system diagram. Right: System architecture.

SYSTEM ARCHITECTURE AND IMPLEMENTATION

To overcome the limitations of existing platforms, we developed a hybrid approach adapted to our specific application that consists of a thin-client low-power eyewear prototype coupled with a mobile phone over a wireless connection (Figure 3). This hybrid approach enables a compact form-factor by requiring only a minimal system, equipped with a wireless interface, a display interface, and sufficient CPU capacity to render graphics primitives. We implemented our prototype using a MediaTek MT2523D System-in-Package (SiP) [27], which is a single-chip Cortex-M4F with integrated Bluetooth (BT) 4.0 EDR and Bluetooth Low Energy (BLE) transceiver, power management, MIPI-DSI display controller, and memory subsystem. We optimized our electronics layout for a mostly single-sided design with a compact footprint of 14.5×60 mm (Figure 4). The major components are the MT2523D SiP, battery charger (TI BQ24230 [39]), and 1 Gbit Flash for assets, e.g., fonts and images (Winbond W25N01G [45]).

Embedded System and Communication protocol

The embedded system and phone communicate using BLE. The application-layer protocol is implemented using Protocol Buffers [10] with message types for drawing graphics primitives, configuring sensors and peripherals, and changing device modes. Another set of message types transmit sensor data, compressed audio captured from the onboard microphones, and remote procedure call (RPC) return values back to the phone.

For graphics effects that require animation, such as our smoothly-scrolling transcript text, we implement interpolation primitives which are executed on the device to drive position and scale parameters of other primitives. This

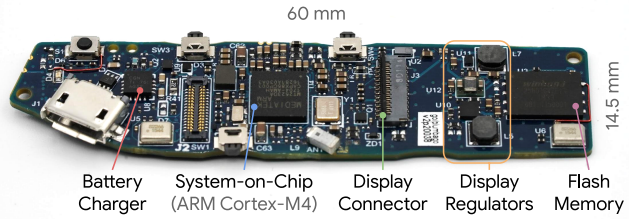

Figure 4. The 8-layer FR4 PCB assembly of our eyewear electronics with wireless communication and display.

approach reduces the bandwidth required to drive these effects, since only the key frames (initial position, animation target, and interpolation strategy) are specified.

APIs: Android, JavaScript and Python for Prototyping

A core service application manages the BLE connection and exposes an API for controlling the display and capturing sensor data through an Android Interface Definition Language (AIDL) interface. External applications such as Live Transcribe [9] communicate with the service to control the display.

Display System

Our prototype system uses a 1-lane MIPI-DSI compatible microdisplay engine with 30 fps graphics rendering. For our proof-of-concept, we use an optical combiner similar to the 15° diagonal example in Cakmakci et al. [2]. It is a monocular right-eyed display and is horizontally offset from the central vision for all users, though the position in the visual field varies depending on the mechanical fit for the individual user. Since the optical see-through display can only add light, we render the text in white to maximize contrast and visibility.

MECHANICAL DESIGN FOR RAPID PROTOTYPING

Our plastic parts are 3D printed with a biocompatible material and use a multi-step finishing process. To maximize material strength and robustness in our rapid prototyping for all-day use, we use SLS (selective laser sintering).

Our first prototype frame included the board in the right temple while the battery was embedded in the left temple (Figure 5 left). After studies 1 and 2 below, we redesigned the prototypes to include two 180 mAh 1-cell LiPo batteries symmetrically placed at the back of left and right temples, which creates a more balanced design and reduced the nose weight from 78% (42g) to 56% (30g). See Figure 5, right. Our prototypes weigh 54 g, which is below the suggested 75 g maximum weight for all-day use HWDs [38].

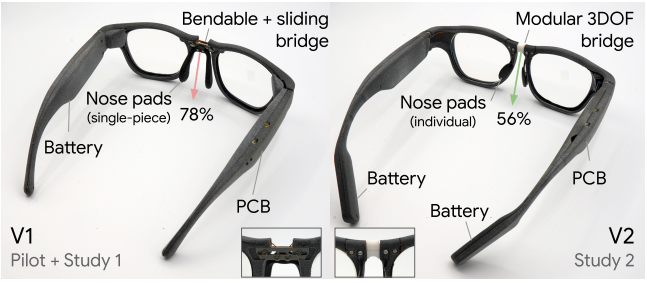

Figure 5. 3D-printed prototype. Left: V1, used in pilot and Study 1. Uses a sliding nose bridge. 54 g weight with 78% on the nose (42 g). Right: V2, used in Study 2. Uses fixed nose bridge modules. Same weight, but two batteries behind the ears create a more balanced design (56% nose weight, 30 g).

USABILITY EVALUATIONS: THREE STUDIES WITH DEAF/HARD-OF-HEARING PARTICIPANTS

We are particularly interested in our proof-of-concept's potential to augment communication and perception for DHH individuals. To improve the prototype iteratively, we conducted a brief pilot study and rapidly iterated with two usability studies:

- We launched a formative in-lab pilot study using our prototypes with external participants (outside of our institution) with a wide range of hearing loss.
- We recruited a group of DHH participants from our institution for a usability study to use the prototype over three days and provide feedback on the contexts of use and the challenges encountered.
- Based on these studies, we incorporated many changes including software improvements, re-engineering the nose bridge, and placing the batteries behind the ears. The second usability study evaluated the improved prototype in more challenging scenarios, such as conversation while walking, interactions in a cafeteria, and a social setting that incorporated group dynamics, multi-party conversation and turn-taking.

Prototype Apparatus and Fitting

All studies used our eyewear prototype with transcribed speech sent wirelessly from an Android Pixel 3 phone. In cases where the eyewear experience was compared to a mobile device, we used an updated version of Live Transcribe [9] running on the same phone. All participants were fitted in individual in-person sessions with our HWD, which included mechanical adjustments (nose pads, nose bridge) to optimize fit. Participants were instructed on how to use the system and how it connected to the phone.

PILOT STUDIES: FORMATIVE IN-LAB RESEARCH

We ran formative in-lab pilot studies with 14 participants in the U.S. (*New York, NY*=6, *Mountain View, CA*=8), whose self-reported hearing loss ranged from moderate (41-70 dB) to profound (>95 dB). All participants used hearing aids. The majority of participants also used captions (n=13), followed by transcription services (n=2), cochlear implants (n=1), and CART services (n=1).

Participants were asked to wear the prototype throughout various pre-planned activities, which included single and multi-party conversations (locally and over video call), communication while working on manual tasks (card sorting, assembly tasks on a computer), watching a movie, and mingling with other participants in a happy hour. In the single-party conversations, the researcher asked the participant questions about themselves. In the multi-party conversations, participants engaged in a meeting with a mock agenda for discussion. For the happy hour, participants were given a card which directed them to find out information about other participants and start a conversation (e.g., 'what is your secret talent?'). P1 was unable to attend the happy hour and movie. P5 did not participate in the multi-party conversations and manual tasks due to fit issues with the device. The tasks selected for this study aimed to explore challenges such as visual dispersion, low visibility, and split attention in scenarios similar to those in the survey and prior work [14, 16, 29].

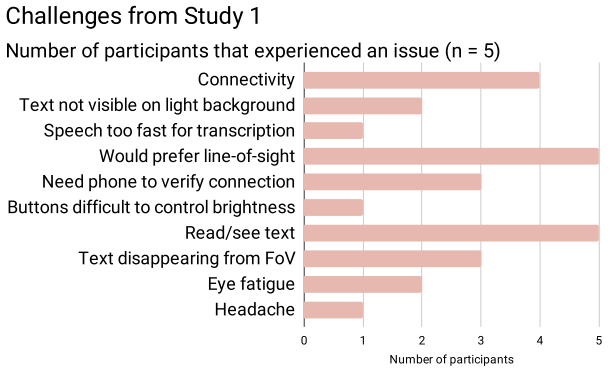

Figure 6. Participants reported a range of challenges with the first version of the system. The most common reported challenges were line of sight, legibility of text, and connectivity.

Participants suggested that **testing in group conversations and more challenging auditory environments** could be helpful to fully understand the benefits of the HWD. For example, it was noted that the quiet lab setting used for the activities was too ideal, compared to settings with more environmental noise. The feedback from this study aligns with the focus on more challenging tasks and procedures for the in-the-wild scenarios described in the next section.

STUDY 1: IN-THE-WILD USE (THREE DAYS)

To gather feedback from in-the-wild scenarios, we conducted a study where participants could use Version 1 of our prototype over three days.

Participants

A total of five DHH participants were recruited from our institution. Four participants used a cochlear implant and all were active users of the Live Transcribe mobile app. We did not collect hearing loss data due to restrictions about data collection on participants from our institution. All participants volunteered to provide feedback and did not receive compensation for using the eyewear prototype.

Task and Procedure

Throughout the three days, participants were encouraged to wear the eyewear prototype during several activities: ordering and consuming beverages/food, watching TV or movies, walking around campus, and participating in meetings, presentations, and conversations. Participants were instructed to try different types of conversations including one-on-one and group conversations, conversing while a car passenger, and in situations with mixed signing and vocal communication. At the end of each day, participants completed a questionnaire detailing their usage time, contexts where they used the prototype, and feedback on how it could be improved.

Results

Participants used Version 1 of the prototype in a variety of settings and scenarios. Continuous usage time ranged from 15 mins to 3 hours, depending on the activity. Two out of the five participants indicated using the prototype in sessions longer than two hours.

Scenarios

Participants limited themselves mainly to scenarios in the office work environment (at desk, meetings, video calls) and in conversation. One participant watched TV at home. Participants who wore the device in three hour sessions mostly spent time at meetings or at their desk.

Display of Transcriptions

All participants experienced issues with Version 1 of the prototype. These issues included the ability to read/see the text and the display disappearing from the field-of-view. Two participants reported eye fatigue and one participant mentioned that they had gotten a headache. Participants had been instructed to stop using the device if issues occurred.

Comfort

We received feedback that the eyewear temple overlapped with cochlear implants and all four cochlear implant users experienced discomfort. The participant without a cochlear implant had issues with the frames being too large.

User Experience

Participants gave feedback on problems that affected their user experience, noting challenges in five main areas: **text visibility, speed of transcription, connectivity, physical button ergonomics, and text placement.** For text visibility, two participants found that bright backgrounds diminished visibility, leading them to take the device off: *"I took it off [when] working on my computer. Too bright to read it even though I want[ed] to see what other [people] around my desk [were saying]." (P2)*. Regarding the speed of transcription, one participant found that the transcription *"fell so far behind to not be useful anymore." (P4)*. Three participants had connectivity challenges when pairing the phone to the eyewear, and they had to repeatedly check their phone to ensure connectivity.

For physical button ergonomics, one participant felt that it was difficult to use the button to change the display settings. Lastly, all users requested a UI closer to their central vision, in order to maintain better focus on the text while talking to peers, although some suggested that it may be a matter of adaptation: *"Getting more used to placement of screen." (P4)*. The comments from participants signaled that the work was premature, but were encouraging and showed interest in the device: *"Good idea but execution needs some improvement. Please keep trying and happy to try again later!" (P5)*.

DISCUSSION AND DESIGN IMPROVEMENTS

Study 1 identified several important issues that we focused on addressing before Study 2.

Connectivity and Pairing

To provide visual feedback to the wearer about when the prototype was connected to the phone, we implemented a real-time audio level meter in the lower right corner of the display (Figure 1). The visualization showed that the system was listening and ready for transcription, also helping to communicate any latency in transcription.

Study 1 also helped us discover and address antenna issues. We improved the RSSI from -70 to -40 dBm, which eliminated unwanted disconnects with the phone anywhere in an approximately 5 m radius (free space).

Display of Transcriptions

Interacting with our participants also highlighted the importance of smooth scrolling in order to mitigate eye fatigue. In Version 1 of the prototype, the verbatim output of the automatic speech recognition (ASR) engine was sent to the wearable, re-rendering the whole display whenever a new transcription result arrived. The ASR engine is balanced for latency and accuracy in such a way that low-confidence predictions arrive first, and as confidence improves with additional context audio, the transcript is updated. As a result, transcript updates would cause visual "jumping" or discontinuity as words changed length or moved around within sentences. In Version 2 of the eyewear prototype, we minimized this abrupt motion by preserving line breaks for the high-confidence portion of the ASR result, such that only the text subject to modification as new audio is collected would reflow. We also implemented scrolling animations that would smoothly roll the text upwards as new lines are added, helping to guide the user's eye along as the transcript moves.

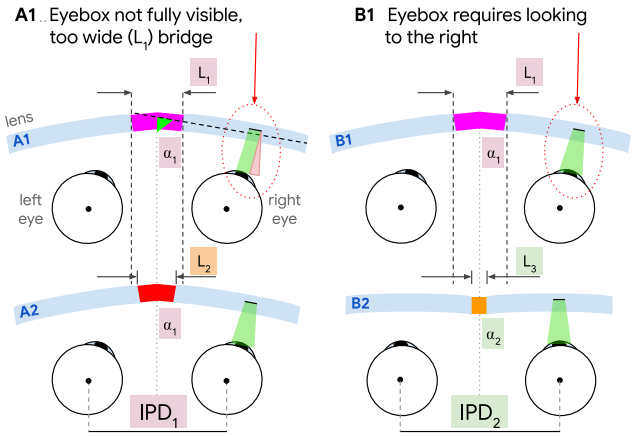

Figure 7. Our modular nose bridge provides independent IPD and wrap adjustments. A1/A2: Reducing nose bridge width addresses display visibility. B1/B2: Adjusting nose bridge with and without wrap rotates the eyebox into a central position.

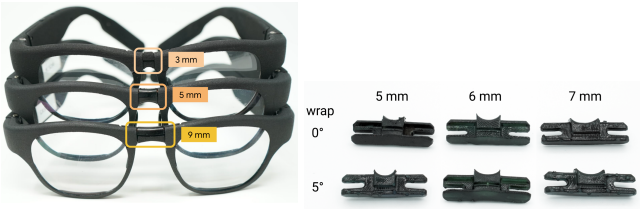

Figure 8. 3D-printed modular nose bridges facilitate personalization for display visibility for small eyebox displays, and maximize prototype reusability.

Comfort: Nose Weight and Balance

To offset the weight of the optics and the display engine in the front, which can lead to slipping and nose pressure, we redesigned the frame to place the batteries at the back of the temple (Figure 5, right), similar to other head-worn systems such as Vuzix Blade and Google Glass. This improvement reduced the nose weight from 42 g (78%) to 30 g (56%).

By exchanging the location of the battery from the temple (Version 1) to the temple tip (Version 2), most of the left temple earpiece remained empty. This provides an opportunity to include more batteries in the future, if the electronics can be rebalanced to avoid adding nose weight.

Eyebox Adjustments: Personalized 3DOF Nose Bridge

To accommodate different and/or asymmetric interpupillary distances (IPDs), we prototyped a more adaptable nose bridge and nose piece. We also investigated shifting the UI towards participants' line-of-sight and display contrast improvements.

Minimal in-lens displays with small eyeboxes, such as North Focals 1.0 and our prototype, require alignment of the optics to the user's face and eye geometry. To see the image, the user's pupil needs to be inside the eyebox volume, which is typically a pyramidal frustum that originates from the display. Thus, with different face geometries and small eyeboxes, mechanical adjustments are needed to align the pupil inside the eyebox. To accommodate mechanical design, North fits Focals 1.0 users with a 3D face-scanning process that identifies appropriate frame sizes [12]. The final fit can be adjusted manually by shaping the nose pads.

To enable personalization in a single frame design without the need for 3D scanning, we developed an interchangeable nose bridge, which provides independent adjustments of IPD, wrap, and cyclo-rotation.

Manual adjustment of the nose pads also supports final adjustments of the eyewear height, nose bridge fit, and monocular display visibility. Wrap angle and distance adjustments control the orientation and position of the eyebox relative to the user's eye, as shown in Figure 7. Based on the feedback from Study 1, we adjusted the wrap and width of the nose bridges to provide a more central display location (See Figure 7, B1/B2).

We generated 12 nose bridges, which we 3D-printed to cover interlens distances from 3–9 mm and 0°, 5°, and 15° wrap. We did not vary cyclo-rotation in these studies. Our fitting process takes about 10 minutes and captures IPD and head width, followed by selection of an optimal nose bridge. The result maintains the subtle eyewear form factor (Figure 8).

STUDY 2: MOBILE AND GROUP CONVERSATIONS

With the revisions to the Version 1 prototype discussed in the previous section, we conducted an additional study to understand the physical and social comfort of Version 2 and associated software improvements. Related work [16] shows that head-worn captions are particularly useful in mobile

scenarios or when hands are occupied. Since our Study 1 participants mostly spent time at work, we decided to emphasize more varied scenarios in our continued experiments, including controlled environments with structured tasks. We wanted to collect feedback on the more central position of the transcriptions and the updated mechanical design in Version 2 of our prototype to see if it would increase perceived physical and social comfort, better accommodating different head sizes, IPDs, and face geometries.

Participants

Five newly-recruited participants completed all parts of the study. An additional sixth participant was recruited for the study, but was excluded from analysis as they completed the tasks but not any of the feedback questionnaires. Of the five participants who completed the study, only one participant used a sign language interpreter. Three participants had moderate (41-70 dB), one had severe (71-95 dB) and one had profound (>95 dB) hearing loss. All participants used hearing aids. One had a cochlear implant, two used real-time captions (e.g., CART), and four used closed captions or subtitles on a daily or near-daily basis. Participants were compensated with up to $150 USD, based on their completion of the activities.

Task and Procedure

Participants provided feedback after completing the following activities:

- Walking to and from a lab to a coffee shop in a connected building, where they were encouraged to order a complimentary beverage. Their path consisted of hallways, elevators, and open collaborative spaces. They used the eyewear prototype when traveling in one direction, and phone transcription for the other direction (~30 min total).
- A game of charades with two experimenters and the participant's companion. Three people viewed a prompt card and described the contents to the guesser. After each round, roles rotated. Participants used the eyewear prototype for one round and phone transcription for the other round (15 min per round).

After each task, participants completed a feedback questionnaire regarding device comfort, ranking a series of statements on 5-point ordinal scales, ranging from 1 = "Not at all..." to 5 = "Extremely…". The order of eyewear and mobile conditions was counterbalanced across participants.

Results

The feedback questionnaires revealed the benefits of the eyewear prototype in comparison to the experience on the mobile phone. The following subsections describe the ratings for social dynamics, comprehension, and user satisfaction.

Social Dynamics: Discreetness and Awareness

Participants found that in mobile contexts ("on-the-go"), the eyewear transcriptions were more discreet ($\tilde{x}$=4, *IQR*=1) compared to a handheld mobile phone ($\tilde{x}$=2, *IQR*=1). (Figure 9, left). For group conversation, the level of perceived discreetness was similar. However, participants did express concern that the apparent use of transcriptions on a phone may be misinterpreted as ignoring other bystanders, which may happen if they needed to look away: *"With [this eyewear], I could just be a person getting support versus 'I'm ignoring you and you don't know if I am reading my Facebook feed [on a mobile phone]'"* (P1).

Participants' ratings also suggest that the eyewear prototypes helped participants become more aware of their surroundings ($\tilde{x}_{eyewear}$=4 vs. $\tilde{x}_{phone}$=3), and who was currently speaking ($\tilde{x}_{eyewear}$=4 vs. $\tilde{x}_{phone}$=2) while using the eyewear, whereas awareness of body language was similar for both conditions (Figure 9, center). Participants explained that the eyewear kept them engaged in conversations: "*The capacity to have other people talking and have me actually know what's going on, instead of me going off [of] other people's body movements, that feels freeing… I miss a lot of cues, a lot of laughs, but I miss a lot… I'm already discreet and faking [that I can hear], this gives me another tool in my pocket so that I can fake less"* (P3).

Comprehension of Presented Contents

Participants rated the text rate on the phone while on-the-go slightly more positive ($\tilde{x}$=4, *IQR*=1) compared to the neutral ratings for the eyewear ($\tilde{x}$=3, *IQR*=1). Participants were, however, neutral about the text rate for the group activity in both conditions, and overall rated their ability to understand the spoken communication as *"Extremely well"* in both conditions ($\tilde{x}$=5; IQR_{phone}=0, $IQR_{eyewear}$=1).

Prototype Eyewear Experience and Desired Use

Participants also provided specific feedback on the eyewear at the end of the study. Participants ranked both fit and the

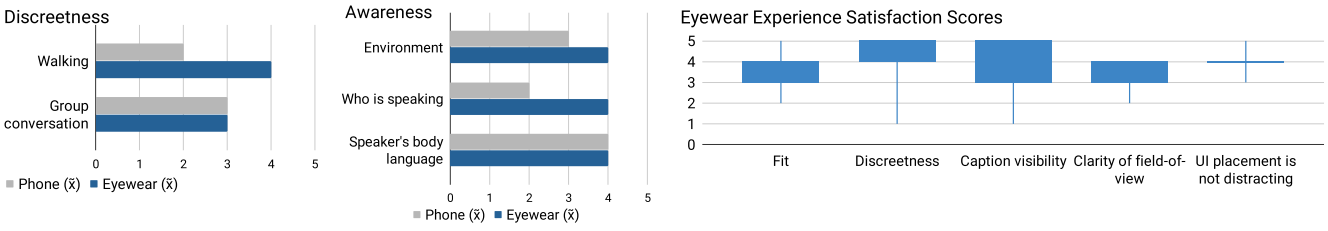

Figure 9. Left: The prototype eyewear was rated more discreet for on-the-go use than phone-based transcription. Center: The prototype eyewear was rated higher for environmental and speaker awareness. Right: The updated v2 design was rated favorable along all dimensions, suggesting that important challenges may have been addressed.

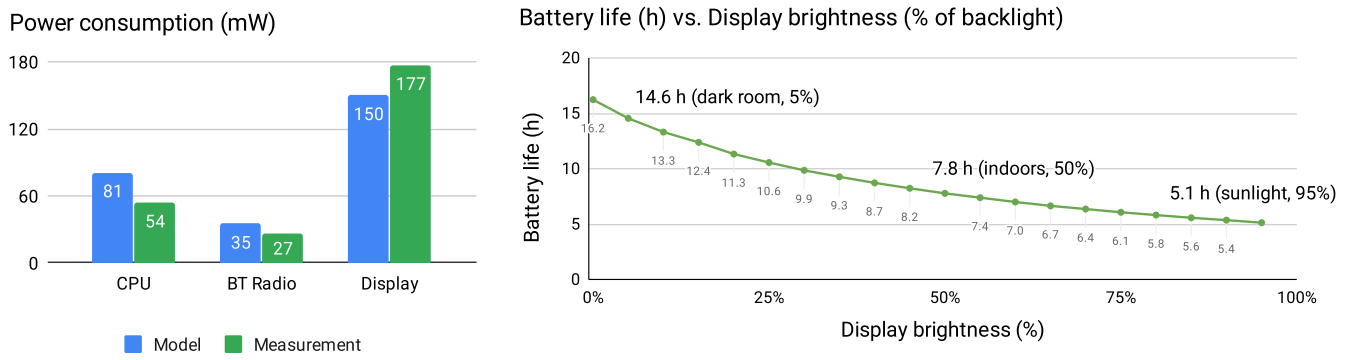

Figure 10. Left: Measured power numbers vs. predicted worst case values from the datasheets. Right: Battery life vs display brightness for display contrast under continuous usage.

overall discreteness as an accessibility technology positive ($\tilde{x}$=4, *IQR*=1). Participants rated the UI (placed more towards the center) not distracting ($\tilde{x}$=4, *IQR*=0). They felt that their field-of-view was clear while using the eyewear ($\tilde{x}$=3, *IQR*=1) but had a neutral rating on the visibility of the captions ($\tilde{x}$=3, *IQR*=2). See Figure 9, right.

When asked about how long they would like to use the system, all participants expressed a desire to use it for multiple hours on a daily basis. Two participants expressed that they would want to use it for *"10 hours or more"* (the highest option), whereas the other three participants chose 3, 4, and 6 hours, respectively. For instance, P2 mentioned the desire for ongoing transcription throughout their work day, "*I don't know the [range of the eyewear transcription] but like, if you're sitting over here and your boss is over there having a conversation, and you really wish you could be a part of that [conversation] but you can't hear it, [this tool could let you] know if you should get out of your seat and go be a part of it."*

Discussion

Study 2 incorporated both a mobile phone and our prototype eyewear in walking and multi-speaker interactions to bring further insights into the potential for a HWD for captions. Our results align with the work from Jain et al. [16] who showed that head-worn captions can provide unique benefits in mobile contexts, where hands-free access, social acceptance, and environmental awareness are of particular importance. Specifically, participants found our eyewear prototype to be more discreet than captions on a phone in a mobile context, and also felt that it allowed them to be more aware of their surroundings. Our previous studies identified group conversations as an area of priority, and Study 2 results suggest that head-worn captions also bring benefits through better speaker awareness, which enabled better participation in the conversations.

These preliminary results also suggest that Version 2 of the prototype addresses some of the main technical challenges identified in the previous studies. This study shows positive ratings for fit, as well as display ergonomics, such as UI placement, clarity of their field-of-view, and overall comprehension. There is still work to do to also improve the currently neutral ratings on the visibility of the captions.

With these improvements and positive feedback, we are excited about opportunities to further validate the potential through quantitative methods for attention, as well as through more extended usage in continued diary studies.

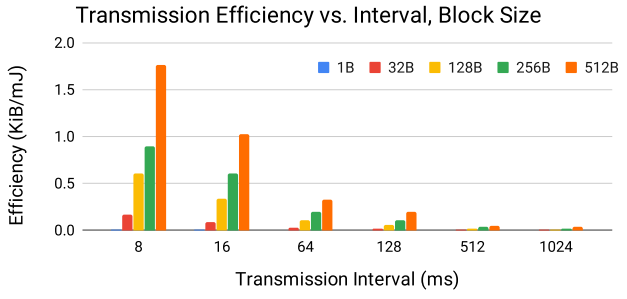

Figure 11. Energy efficiency of data transmission in various Bluetooth usage modes.

TECHNICAL EVALUATION

As discussed in the Related Work, prior research indicated a need for reliable transcription and sufficient battery throughout daily tasks [16, 25]. In this section, we evaluate the power/performance relationship of our proof-of-concept system across power, latency, and bandwidth metrics.

System Power Consumption

Our power consumption modeling from device specifications maps well to measured values. At full display brightness, we estimated 266 mW and measured 258 mW. (Figure 10, left). With a 360 mAh LiPo battery (3.7V) and 90% power efficiency, we estimate 4.5 hours in sunlight (100% display), 8 hours indoors (50% display), and approximately 15 hours in a dark room (5%) (Figure 10, right). Note that our LCOS display backlight accounts for >70% of the power consumption at maximum brightness. More efficient displays would greatly extend our battery life. The thin client hardware architecture is inherently low power as the firmware is running on a single Cortex-M4 CPU at 208 MHz. We use high-efficiency (>90%) switched-mode power

supplies for the most power-intensive components, i.e., for the back-light driver, radio, and CPU core.

Power Consumption for Communications Scenarios

We evaluated BLE power consumption under various usage scenarios. Figure 11 shows the energy efficiency of communication (KiB/mJ) for different transmission intervals and transmission sizes (block size). The rate of individual transmission events correlates with power consumption, so more efficiency is achieved by buffering multiple data packets and transmitting them together. For latency-critical components (such as audio streaming), we do not perform this buffering. We evaluated peak transmit bandwidth using a dummy firmware image to keep the transmit queue full. We achieve an efficiency of 2.2 KiB/mJ at a 645 kbps peak data rate (65% of the 1 Mbps PHY signaling rate). We also evaluated system latency by measuring the round-trip time from a test script to the firmware and back over BLE. This latency is largely controlled by the connection interval and slave latency of the BLE connection, which we configured to their minimum possible values (7.5 ms and 0 ms, respectively). We observed a mean latency of 21.0 ms with a standard deviation of 7.7 ms (N=200). One-way latency can be approximated by half of the round-trip latency. Latency is critical for conversation comprehension.

LIMITATIONS AND FUTURE WORK

Our user studies included 24 DHH participants, 19 from outside our institution, and five from within. We recruited DHH individuals who used a breadth of communication methods and assistive technologies to include diverse perspectives. In future work, we would like to validate our technology with both an increased number of participants and extended usage time to facilitate statistical analysis. We plan to conduct more in-depth human factors experiments to quantify ergonomics and comfort over longer periods.

Multiple speakers in group conversations can cause confusion in the transcription, whether a phone or HWD is used. While providing each speaker with a wireless microphone and displaying each speaker's captions in separate colors can improve the problem, the extra hardware is inconvenient to carry and keep charged. In future work, we are investigating whether a beam forming microphone on the HWD might help the user focus attention on one speaker when necessary by turning their head to that speaker. Combining this feature with an ambient microphone and different colors for the different conversational streams, might help disambiguate the attended speaker from others.

Jain et al. [16] prototyped world-stabilized captions, whereas Klose et al. [21] compared variations of head-locked versus body-locked AR text presentations during walking. In future work, we are interested in evaluating text placement with our prototype system and strategies to improve legibility and comprehension [4, 5]. We have recently implemented sound awareness, inspired by suggestions from related work and our studies. Our first diary study pilot participant, who used our device over four weeks, mentioned how useful it was to be able to "see" a door bell or their dogs barking. We have also enabled translation between different languages, which is another feature that could unlock benefits to a larger population. We hope to explore display architectures without active backlights to further reduce power consumption. We also plan to quantify the impact on the *phone's* battery life. An informal experiment suggests that cloud-based transcription in the eyewear (phone in ambient display mode) could extend the battery by 62% (suggesting battery life on par with the eyewear), compared to only using the phone with its display at full brightness.

Privacy issues are important for continuously captured audio. Our recent offline privacy-preserving ASR implementation on the phone still results in a 3% longer battery life when combined with the eyewear, compared to cloud-computed transcription displayed on the phone. Future technical privacy opportunities, such as beamforming, could constrain microphone direction and distance to match human perception, while UI and industrial design could improve transparency and conformity to social norms.

CONCLUSIONS

In this paper, we introduce Wearable Subtitles, a lightweight headworn prototype system for all-day hearing accessibility. By integrating a real-time pipeline that uses cloud speech recognition technologies with our low-power embedded system and a near-eye display, we enable continuous speech transcription, translation, and sound awareness in the user's private field-of-view. We describe our technical architecture and system evaluation which explain the strategies that enable up to 15 hours of active use, 54 g weight and compact electronics packaged into 3D-printed frames.

To validate our proposed approach for hearing accessibility, we conducted a pilot and two studies with 24 DHH participants who provided feedback on our prototypes in various scenarios and tasks. Our user research suggests that HWDs could greatly improve hearing accessibility through privately transcribed text, which can be used hands-free, in mobile contexts, and in socially acceptable interactions.

ACKNOWLEDGMENTS

We thank our many collaborators for contributing their expertise in hearing accessibility, in particular Dimitri Kanevsky, who is continuously helping us improve this vision while evangelizing the DHH perspective, and also Chet Gnegy, Pascal Getreuer and Dick Lyon; software development, Eric Bouchard; optics, human factors and hardware, especially Omar Negrete, Kiet Tang, Ozan Cakmakci, David Hoffman and Ella Zhang; mobile perception, especially Eunyoung Kim, Jeff Gilbert and Alec Go; user experience and product thinking, Robin Dua, Chelsey Fleming, JD Velasquez, Jin Kim, Leo Szrejter, and many others. We thank Mathieu Le Goc for valuable discussions and suggestions, and our reviewers and study participants for their useful feedback and perspectives. We thank Mike Burns for additional footage and feedback.

REFERENCES

[1] Matthias Böhmer, T. Scott Saponas, and Jaime Teevan. 2013. Smartphone use does not have to be rude. *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services (MobileHCI '13)*, 342–351. https://doi.org/10.1145/2493190.2493237

[2] Ozan Cakmakci, Oscar Martinez, and Jerry Carollo. 2019. Optical design of a thin curved lightguide and manufacturing using ophthalmic approaches. *Digital Optical Technologies 2019*. https://doi.org/10.1117/12.2530526

[3] Leah Findlater, Bonnie Chinh, Dhruv Jain, Jon Froehlich, Raja Kushalnagar, and Angela Carey Lin. 2019. Deaf and Hard-of-hearing Individuals' Preferences for Wearable and Mobile Sound Awareness Technologies. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19),* 1–13. https://doi.org/10.1145/3290605.3300276

[4] Michele Fiorentino, Saverio Debernardis, Antonio E. Uva, and Giuseppe Monno. 2013. Augmented Reality Text Style Readability with See-Through Head-Mounted Displays in Industrial Context. *Presence: Teleoperators and Virtual Environments 22*, 171–190. https://doi.org/10.1162/pres_a_00146

[5] Joseph L. Gabbard, J. Edward Swan, and Deborah Hix. 2006. The Effects of Text Drawing Styles, Background Textures, and Natural Lighting on Text Legibility in Outdoor Augmented Reality. *Presence: Teleoperators and Virtual Environments 15*, 16–32. https://doi.org/10.1162/pres.2006.15.1.16

[6] Georgia Institute of Technology. 2014. Captioning on Glass: Enable Everyday Communication for the Hard-of-Hearing. *Georgia Institute of Technology.* Retrieved August 17, 2020 from http://cog.gatech.edu/

[7] Google. 2020. Analytics Solutions | Surveys. *Google Surveys.* Retrieved August 17, 2020 from https://surveys.withgoogle.com/

[8] Google. 2020. Opinion Rewards - It Pays to Share Your Opinion. *Google Surveys.* Retrieved August 17, 2020 from https://surveys.google.com/google-opinion-rewards/

[9] Google. 2020. Live Transcribe: Speech to Text App. *Android.* Retrieved August 17, 2020 from https://www.android.com/accessibility/live-transcribe/

[10] Google. 2020. Protocol Buffers - Google's data interchange format. *Github.* Retrieved August 17, 2020 from https://github.com/protocolbuffers/protobuf

[11] Greta App. 2020. Barrier-free cinema and inclusion. *Greta & Starks.* Retrieved August 17, 2020 from http://www.gretaundstarks.de/greta/GretaAndStarks

[12] Scharon Harding. 2019. North Focals Review: Stealthy, Stylish Smart Glasses. *Tom's Hardware*. Retrieved August 17, 2020 from https://www.tomshardware.com/reviews/north-focals-smart-glasses-ar,5968.html

[13] James D. House. 2016. Telecommunications for the Deaf and Hard of Hearing, Inc., Formation of. *The SAGE Deaf Studies Encyclopedia*, 960–961. https://doi.org/10.4135/9781483346489.n310

[14] Dhruv Jain, Bonnie Chinh, Leah Findlater, Raja Kushalnagar, and Jon Froehlich. 2018. Exploring Augmented Reality Approaches to Real-Time Captioning: A Preliminary Autoethnographic Study. In *Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems* (DIS '18 Companion), 7–11. https://doi.org/10.1145/3197391.3205404

[15] Dhruv Jain, Leah Findlater, Jamie Gilkeson, Benjamin Holland, Ramani Duraiswami, Dmitry Zotkin, Christian Vogler, and Jon E. Froehlich. 2015. Head-Mounted Display Visualizations to Support Sound Awareness for the Deaf and Hard of Hearing. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15),* 241–250. https://doi.org/10.1145/2702123.2702393

[16] Dhruv Jain, Rachel Franz, Leah Findlater, Jackson Cannon, Raja Kushalnagar, and Jon Froehlich. 2018. Towards Accessible Conversations in a Mobile Context for People who are Deaf and Hard of Hearing. *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '18),* 81–92. https://doi.org/10.1145/3234695.3236362

[17] Dhruv Jain, Angela Lin, Rose Guttman, Marcus Amalachandran, Aileen Zeng, Leah Findlater, and Jon Froehlich. 2019. Exploring Sound Awareness in the Home for People who are Deaf or Hard of Hearing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19),* 1–13. https://doi.org/10.1145/3290605.3300324

[18] Dhruv Jain, Kelly Mack, Akli Amrous, Matt Wright, Steven Goodman, Leah Findlater, and Jon E. Froehlich. 2020. HomeSound: An Iterative Field Deployment of an In-Home Sound Awareness System for Deaf or Hard of Hearing Users. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*, 1–12. https://doi.org/10.1145/3313831.3376758

[19] Michael Jones, M. Jeannette Lawler, Eric Hintz, Nathan Bench, Fred Mangrubang, and Mallory Trullender. 2014. Head mounted displays and deaf children. *Proceedings of the 2014 conference on Interaction design and children (IDC '14)* 317–320.

https://doi.org/10.1145/2593968.2610481

[20] Ichiro Kasai, Yasushi Tanijiri, Takeshi Endo and Hiroaki Ueda. A forgettable near eye display. *Digest of Papers. Fourth International Symposium on Wearable Computers*, 115–118. https://doi.org/10.1109/iswc.2000.888472

[21] Elisa Maria Klose, Nils Adrian Mack, Jens Hegenberg, and Ludger Schmidt. 2019. Text Presentation for Augmented Reality Applications in Dual-Task Situations. *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR '19)*, 636–644. https://doi.org/10.1109/vr.2019.8797992

[22] Marion Koelle, Swamy Ananthanarayan, and Susanne Boll. 2020. Social Acceptability in HCI: A Survey of Methods, Measures, and Design Strategies. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '19)*, 1–19. https://doi.org/10.1145/3313831.3376162

[23] Catherine Kudlick. 2011. Black Bike, White Cane: Nonstandard Deviations of a Special Self. *Disability Studies Quarterly, 31(1)*. https://doi.org/10.18061/dsq.v31i1.1373

[24] Walter S. Lasecki, Raja Kushalnagar, and Jeffrey P. Bigham. 2014. Helping students keep up with real-time captions by pausing and highlighting. *Proceedings of the 11th Web for All Conference on (W4A '14)*, 1–8. https://doi.org/10.1145/2596695.2596701

[25] Tara Matthews, Scott Carter, Carol Pai, Janette Fong, and Jennifer Mankoff. 2006. Scribe4Me: evaluating a mobile sound transcription tool for the deaf. In *Proceedings of the 8th international conference on Ubiquitous Computing (UbiComp'06)*, 159–176. https://doi.org/10.1007/11853565_10

[26] Abby McCormack and Heather Fortnum. 2013. Why do people fitted with hearing aids not wear them? *International Journal of Audiology 52*, 360–368. https://doi.org/10.3109/14992027.2013.769066

[27] MediaTek. MT2523D for Wearables: Highly integrated and efficient wearable-focused chip with power management unit and dual-mode Bluetooth. *MediaTek*. Retrieved August 17, 2020 from http://www.mediatek.com/products/wearables/mt2523d

[28] Microsoft. 2020. Microsoft Translator for Consumers. *Microsoft*. Retrieved August 17, 2020 from https://www.microsoft.com/en-us/translator/

[29] 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. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*, 1909–1915. https://doi.org/10.1145/3027063.3053117

[30] Shalini Misra, Lulu Cheng, Jamie Genevie, and Miao Yuan. 2016. The iPhone Effect: The Quality of In-Person Social Interactions in the Presence of Mobile Devices. *Environment and behavior* 48, 2: 275–298. https://doi.org/10.1177/0013916514539755

[31] National Association of the Deaf. 2020. CART: Communication Access Realtime Translation. *National Association of the Deaf*. Retrieved August 17, 2020 from https://www.nad.org/resources/technology/captioning-for-access/communication-access-realtime-translation/

[32] Yi-Hao Peng, Ming-Wei Hsi, Paul Taele, Ting-Yu Lin, Po-En Lai, Leon Hsu, Tzu-Chuan Chen, Te-Yen Wu, Yu-An Chen, Hsien-Hui Tang, and Mike Y. Chen. 2018. SpeechBubbles: Enhancing Captioning Experiences for Deaf and Hard-of-Hearing People in Group Conversations. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*, 1–10. https://doi.org/10.1145/3173574.3173867

[33] Halley Profita, Reem Albaghli, Leah Findlater, Paul Jaeger, and Shaun K. Kane. 2016. The AT Effect. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 4884–4895. https://doi.org/10.1145/2858036.2858130

[34] Rachel Rood. 2013. New Closed-Captioning Glasses Help Deaf Go Out To The Movies. *NPR.org*. Retrieved August 17, 2020 from https://www.npr.org/sections/alltechconsidered/2013/05/12/183218751/new-closed-captioning-glasses-help-deaf-go-out-to-the-movies

[35] Chris Schipper and Bo Brinkman. 2017. Caption Placement on an Augmented Reality Head Worn Device. *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '17)*, 365–366. https://doi.org/10.1145/3132525.3134786

[36] Kristen Shinohara and Jacob O. Wobbrock. 2011. In the shadow of misperception. *Proceedings of the 2011 annual conference on Human factors in computing systems (CHI '11)*, 705–714. https://doi.org/10.1145/1978942.1979044

[37] Mark B. Spitzer, Noa M. Rensing, Robert McClelland, and Paul Aquilino. Eyeglass-based systems for wearable computing. *Digest of Papers. First International Symposium on Wearable Computers*, 48–51. https://doi.org/10.1109/iswc.1997.629918

[38] Mark Spitzer. Personal Communication. 2011.

[39] Texas Instruments. 2020. BQ24230: USB Friendly 0.5A Li-ion charger. *Texas Instruments*. Retrieved August 17, 2020 from https://www.ti.com/product/BQ24230

[40] Aaron Toney, Barrie Mulley, Bruce H. Thomas, and Wayne Piekarski. 2003. Social weight: designing to minimise the social consequences arising from technology use by the mobile professional. *Personal and Ubiquitous Computing 7*, 309–320. https://doi.org/10.1007/s00779-003-0245-8

[41] Hubert W. Upton. 1968. Wearable eyeglass speechreading aid. *American annals of the deaf* 113, 2: 222–229.

[42] Mordekhai Velger. 1998. *Helmet-mounted Displays and Sights*. Artech House Publishers. ISBN-13: 978-0890068229.

[43] Stefan Werrlich, Phuc-Anh Nguyen, and Gunther Notni. 018. Evaluating the training transfer of Head-Mounted Display based training for assembly tasks. *In Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference (PETRA '18)*, 297–302. https://doi.org/10.1145/3197768.3201564

[44] What is CART and FAQ's – Find a CART Captioning Provider. 2012. *Captioning Activism and Community*. Retrieved August 17, 2020 from http://ccacaptioning.org/faqs-cart/

[45] Winbond. 2020. W25N01GV: (1G-bit) Serial SLC NAND Flash Memory. *Winbond.* Retrieved August 17, 2020 from https://www.winbond.com/resource-files/w25n01gv%20revg%20032116.pdf