See CC:
Improving Conversational Experiences for Hearing Loss

Kelden Lin and Erika Yasuda

Problem

Our Argument

According to the World Health Organization (WHO), over 5% of the world's population suffers from disabling hearing loss, which amounts to about 360 million people ("World Wide Hearing Loss: Stats from Around the World"). In developed countries like the U.S., "over half of the cases of hearing loss are due to genetic predisposition" ("World Wide Hearing Loss: Stats from Around the World").Thankfully, there is an abundance of technologies available to assist those with hearing impairments today, with the aim of helping these individuals get around the obstacles posed by their disability. However, while many of these devices are successful in bridging the communication gap, there are usually numerous intermediate steps that need to be satisfied to provide these individuals with "seamless" conversation; this puts them at a disadvantage because of the inconvenience of accessing and preparing these resources that are crucial to their ability to converse and listen.

In order to have a better understanding of what these intermediate steps were across different available solutions, we conducted some research on the current available resources for hearing impairments, and listed the setbacks of each:

- Cochlear Implant: An invasive surgical solution that implants a device that will send signals to the brain to aid with interpretation of sound waves from voices.
    - **Risk** – Surgery that has different effects on different people
    - **Expensive** – $30,000 and $50,000 for patients without insurance ("How Much Does a Cochlear Implant Cost?").
    - **Safety** – Vulnerable moving parts even when completed
- CART (Communication Access Realtime Translation): A service provided by the University of Washington, in which a transcriber is appointed to a student to transcribe a lecture word-for-word, which will be projected to a small monitor for the student to follow along with the lecturer in close to real-time ("What Is Real-Time Captioning?").
    - **Limited Availability** – Limited to use in the classroom setting and by appointment.
    - **Multi-tracking** – Need to keep looking at the monitor to read the spoken text, as well as the actual display if there are any visuals being explained alongside the spoken explanation.
- Specialized Courses: A service provided by many educational institutions, in which courses are specifically structured around the needs of the student's disability.
    - **Not Integrated** – Separated disabled students into a different class, which is not a seamless integration.

What we can infer from many of these current solutions is that they are either dependent on the cooperation of another individual, costly, invasive, or a combination of these three. While the benefits of these solutions may outweigh their respective complications for specific individuals, as a collective they fall short of providing a seamless and natural conversation experience for those with hearing impairments.

Given the results of our research on current solutions, as well as the statistics regarding the impact that hearing impairment has on a global scale, we want to design a product that can not only seamlessly fit into people's lives, but do so without the complications of those solutions. In fact, rather than a "solution" to change an individual's hearing impairment, we seek to create a product that simplifies the complications of current solutions, and serves as the single interface between a deaf user and the conversations happening around them.

Solution

Our Proposal

See CC is a hands-free smart glasses device, providing real-time closed captions of conversations for users with hearing impairments. It aims to remove the intermediate steps and complications that the current solutions involve, and help users focus on what's most important: being fully engaged in a conversation. We provide three modes to adapt to different listening situations: lecture, conversation, and custom. These are catered for listening to few speakers who are farther away, many speakers who are close by, and for custom situations respectively.

- In order to avoid
- Key features (but not in too much detail since we'll do that in the prototype section)
- How we plan to implement it
- How will we spread awareness of this idea? (e.g. social media?)

Information to Compute

Text-to-Speech

The most vital part of this design is to collect the auditory data coming in from the microphones, convert this into words, and ultimately project these words as text to the user's glasses. This calculation is crucial for delivering the desired behavior for a user with a hearing impairment, as the intended outcome is to simulate a conversation as someone who does not have a hearing disability.

Directional Data

Users will need to know where the person they are listening to is relative to themselves. This calculation is important because the glasses need to overlay the speech-to-text text bubble on top of where the speaker is displayed through the glasses. This will help explicitly tie the speaker with their respective speech.

Information to Store

Common Listening Modes

The device will not be storing any data pertaining to the conversations that the user engages in, besides when the conversation text is being displayed to the user. This is primarily for privacy purposes that come with preserving and recording conversations without consent. However, what will be important to store will be the settings for different "listening" modes, as listed below:

- *Lecture*
  This setting will be catered towards focused listening of the microphone at a farther distance.
- *Casual close-contact conversation*
  This setting will be catered towards a shorter range of microphone listening, but with the ability to listen to more than one individual at a time to accommodate group conversations.
- *Custom*

In the case that the above two modes are not sufficient to cater to a specific situation, the user has the ability to customize the settings that will best adhere to their specific situation. The device will save this setting so that the user can quickly switch between the other modes.

Design Rationale
In order to improve conversational experiences for those with hearing loss while minimizing cost and vulnerability (no surgeries, not add-ons). By implementing transcription right on the screen, the user can directly understand what is being said without actually hearing, but using their sense of sight – the rest is up to the program itself. We've made some design decisions to assist edge cases where there may be more than one person talking, or somebody you're listening to but not looking at.

High-Fidelity Prototype
https://docs.google.com/presentation/d/118MgLDYwzhzcbe1hw0VwvmyMv61TFHxgsLtd-yVIOYs/edit#slide=id.g473db0b229_0_6

Calibration

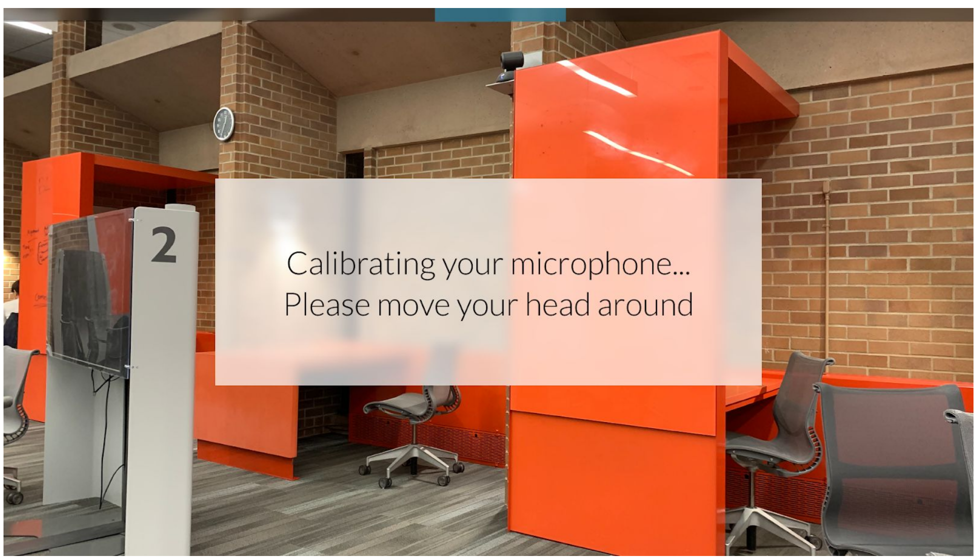	Figure 1

Figure 2

Figure 1. We wanted to make calibration a very easy and straightforward task. Turning on the glasses for the first time will prompt the calibration screen. To finish the calibration, the user needs to look around. As you look around, the the blue bar on the top of the screen fills up based on how much of the calibration you've completed.

Figure 2. Once completed, the screen then briefly displays a quick "Calibration completed!" message, and a gesture indicator on the top right of the screen. It teaches the user how to lock and unlock onto people talking.

Enabling Lock/Unlock using "Eyeover"

Figure 3

Figure 4

Figure 3. The bright green tint on the unlock/lock indicates that that current person is highlighted (or “eyeovered” – mouseover but with your eyes). This happens because the user is facing her and looking at her to enable the eyeover. The female speaker is currently unlocked, but able to be locked as she is now selected.

Figure 4. Now the bright green tint is on the male speaker, indicating that his speech can now be unlocked (as it is locked right now). What that means will be explained in the next section – “Locking/Unlocking a speaker”.

Locking/Unlocking a speaker

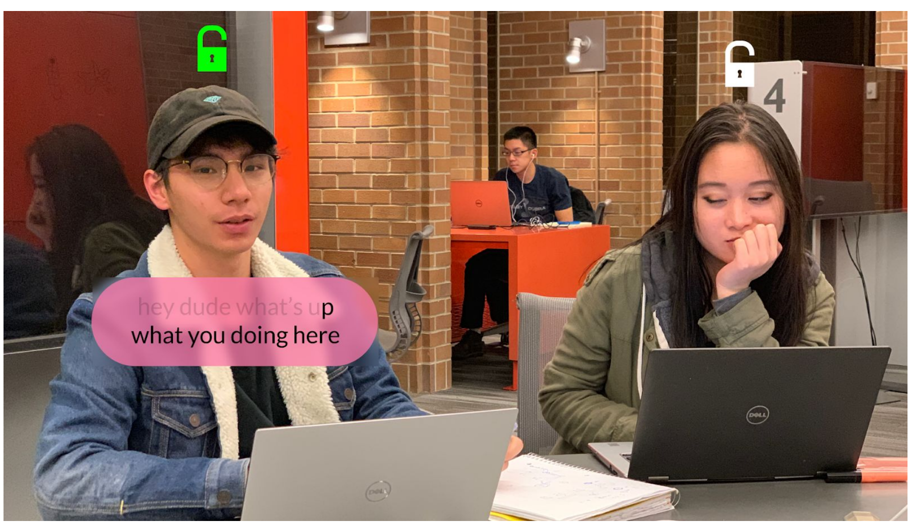

Figure 5

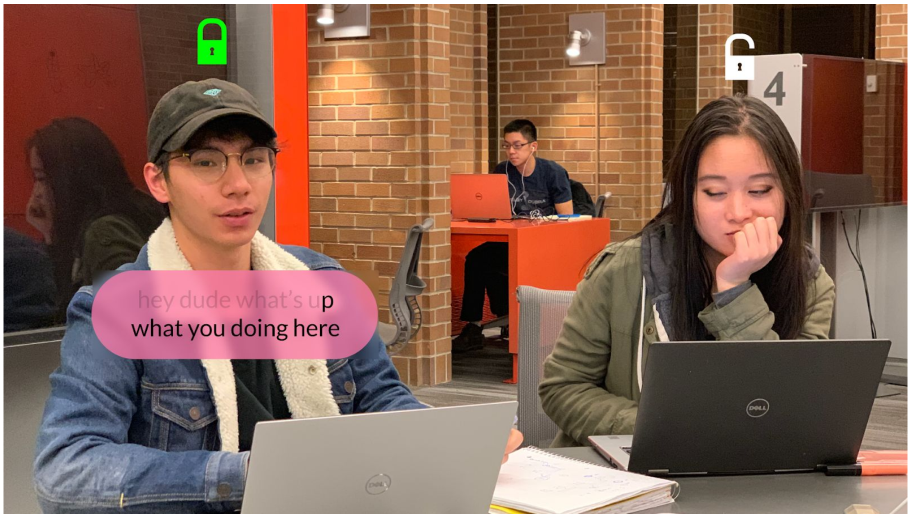	Figure 6
	Figure 7

Figure 5. To enable locking or unlocking, the user must “Eyeover” onto the speaker they’re interested in locking or unlocking on. In this figure, the user is eyeovering on the male speaker, and since he’s unlocked, tapping on the side of the right side of the glasses will lock the speaker.

Figure 6. In this figure, the male speaker has been locked on. This means that the transcription of that speaker will always be available to the user even if the user isn’t looking at them.

Figure 7. When the user looks away from this speaker, the transcription of his speech appears on the bottom left or right, showing which way the speaker is relative to the user.

Evaluation

We chose to use the GenderMag Walkthrough as our design method in order to recognize and rectify the possible breakdowns for a typical user, represented by the personas of our design.

We tested the goal of “locking and unlocking on a speaker” because it is a vital feature in our product that can make or break a lot of edge cases. In cases where there many people talking, where you’re not necessarily looking at the speaker, or where you’re in lecture writing notes, locking is a key feature for the user to know because it makes everything listed above much easier to do.

Limitations

References

“How Much Does a Cochlear Implant Cost?” *CostHelper*, 2018, health.costhelper.com/cochlear-implant.html.

“What Is Real-Time Captioning?” *DO-IT Disabilities, Opportunities, Internetworking, and Technology*, University of Washington, 15 Sept. 2017, www.washington.edu/doit/what-real-time-captioning.

“World Wide Hearing Loss: Stats from Around the World.” *Audicus*, Audicus, 1 Feb. 2016, www.audicus.com/world-wide-hearing-loss-stats-from-around-the-world/.