Deaf and Hard-of-Hearing Users' Preferences for Hearing Speakers' Behavior during Technology-Mediated In-Person and Remote Conversations

Matthew Seita
Rochester Institute of Technology
Rochester, NY, USA
mss4296@rit.edu

Sarah Andrew
Rochester Institute of Technology
Rochester, NY, USA
sa2941@rit.edu

Matt Huenerfauth
Rochester Institute of Technology
Rochester, NY, USA
matt.huenerfauth@rit.edu

ABSTRACT

Various technologies mediate synchronous audio-visual one-on-one communication (SAVOC) between Deaf and Hard-of-Hearing (DHH) and hearing colleagues, including automatic-captioning smartphone apps for in-person settings, or text-chat features of videoconferencing software in remote settings. Speech and non-verbal behaviors of hearing speakers, e.g. speaking too quietly, can make SAVOC difficult for DHH users, but prior work had not examined technology-mediated contexts. In an in-person study (N=20) with an automatic captioning smartphone app, variations in a hearing actor's enunciation and intonation dynamics affected DHH users' satisfaction. In a remote study (N=23) using a videoconferencing platform with text chat, variations in speech rate, voice intensity, enunciation, intonation dynamics, and eye contact led to such differences. This work contributes empirical evidence that specific behaviors of hearing speakers affect the accessibility of technology-mediated SAVOC for DHH users, providing motivation for future work on detecting or encouraging useful communication behaviors among hearing individuals.

CCS CONCEPTS

• Human-centered computing -> Accessibility design and evaluation methods; *Empirical studies in accessibility.*

KEYWORDS

Deaf, Hard of Hearing, Accessibility, Conversational Behaviors, Videoconferencing, Automatic Speech Recognition, COVID-19.

ACM Reference format:
Matthew Seita, Sarah Andrew, and Matt Huenerfauth. 2021. Deaf and Hard-of-Hearing Users' Preferences for Hearing Speakers' Behavior during Technology-Mediated In-Person and Remote Conversations. In *Proceedings of the 18th International Web for All Conference (W4A '21), April 19–20, 2021, Ljubljana, Slovenia.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3430263.3452430

W4A '21, April 19–20, 2021, Ljubljana, Slovenia

ACM ISBN 978-1-4503-8212-0/21/04...$15.00
https://doi.org/10.1145/3430263.3452430

1 INTRODUCTION

In the U.S., almost 20% of adults are Deaf or Hard-of Hearing (DHH) [5], 60% of whom are students or participate in the workplace [17], and many struggle with communication barriers in a hearing-dominated society. Many DHH people are not satisfied with their current experience with Synchronous Audio-Visual One-on-One Communication (SAVOC), in which the DHH user has access to the visual (and to some degree the audio, depending on the DHH user) behavior of their conversational partner, in real-time conversation. SAVOC may occur in-person meetings or remotely, e.g. videoconferencing. In prior research, DHH individuals have expressed a desire for better communication strategies, especially in the workplace [9], to improve their understanding of meetings and impromptu encounters with other employees and supervisors. Such challenges may lead to lower educational and professional opportunities: Studies have shown that those who are DHH are less likely to have graduated with a Bachelor's degree [11], more likely to be unemployed [10], and earn lower salaries than their hearing peers [43]. Participation in small-group meetings, in particular, has been shown to be correlated with academic and vocational success [2, 27]. Additionally, communication barriers often extend beyond the workplace: Many people in the DHH community have a difficult time communicating with hearing peers day-to-day, such as interacting with healthcare providers, business employees, or socializing with hearing friends [16, 25].

While paid, human services (e.g. a sign language interpreter) may be beneficial, they are not financially provided in all settings nor practical for impromptu meetings. Thus, many DHH individuals have turned to various technological tools to facilitate SAVOC with their hearing peers, e.g. automatic speech recognition (ASR) [23, 4, 8]. Mobile applications using ASR can be well-suited for in-person impromptu communication with a small number of hearing people, as they can be downloaded on a user's personal device. However, ASR applications are not without issues, including speed and accuracy concerns [3, 21] or the need for research on effective interfaces for these apps [26].

The use of videoconferencing apps, e.g. Zoom [46], has risen during the COVID-19 pandemic. These apps often provide text-chat, enabling participants to clarify words or phrases, or switch to text communication, e.g. [46, 13]. Some apps offer automatic captioning based on ASR, e.g. [13]. Increased use of videoconferencing for work may raise accessibility issues for

DHH users, as network factors can reduce call quality. Additionally, while text-chat or other features may increase accessibility, research is needed to understand this setting.

Technology that considers the behavior of hearing individuals is one promising avenue for supporting SAVOC with DHH users: Certain behaviors of hearing individuals may facilitate or impede communication with DHH individuals [35], and technology can support users in monitoring their behavior, e.g. [34, 7]. Before HCI researchers can investigate technology for encouraging hearing individuals to exhibit desirable behaviors in conversation, empirical evidence must first be obtained about which behaviors DHH users prefer and how they would prioritize them.

We investigate DHH individuals' preferences for how a hearing person should behave during technology-mediated SAVOC. DHH participants in our study had brief conversational exchanges with a hearing actor, who exhibited certain behaviors while speaking, including variations in: Speech Rate, Voice Intensity, Enunciation, Intonation, Eye Contact, Gesturing, and Intermittent Pausing. After each exchange with the hearing actor, DHH participants rated how satisfied they were with the actor's behaviors. In an in-person study, changes in the actor's enunciation and intonation affected the subjective preferences of the DHH participants. Motivated by COVID-19, a follow-up study was conducted with an additional 23 DHH participants in videoconferencing meetings, which revealed that changes in the hearing actor's speech rate, voice intensity, enunciation, intonation, and eye contact affected subjective preferences. Open-ended feedback revealed participants' views for in-person and virtual meetings.

Our paper has two main empirical contributions: (1) We present evidence that DHH individuals prefer particular communication behaviors (speech-related and non-verbal) from their hearing conversational partner during SAVOC meetings using ASR-based captioning apps. This is the first evidence from an experimental study, in which the both the preferences and the priorities of DHH users have been measured. (2) We also present evidence of DHH individuals' preferences and prioritization of these behaviors in SAVOC meetings using videoconferencing applications. Most notably, DHH users' preferences and priorities for such behaviors differ in the videoconferencing context. Future work would be needed to examine whether prompting hearing people to adjust their behavior would be beneficial, or might lead to awkwardness. Despite such a risk, the status quo inequitably places the burden on DHH individuals to remind their hearing partners to behave in ways to ensure communication. This paper provides guidance as to which behaviors may be fruitful to target in SAVOC settings.

2 PRIOR WORK

2.1 Speechreading

Speechreading, also called lipreading, is a technique in which DHH individuals observe the movements of the speaker's lips and face, to support conversation, and prior HCI research has supported users in developing these skills [14]. Other research has investigated methods of supplementing audio with additional information modalities to support speechreading, e.g. [18]. While speechreading is a valuable skill for many in the DHH community, speechreading alone is not sufficient for equal access to spoken information. A clear view of the speaker's face is required, and even skilled speech-readers perform below 50% accuracy in lip-reading recognition tasks [1]. Nevertheless, speechreading is an important tool when used in combination with other strategies, such as using context clues, gesturing, and technology-facilitated communication, particularly ASR technologies.

Several factors affect the speechreading abilities of DHH individuals, including behaviors of the hearing speaker, e.g. diction, body language, speech rate, and prosody. Such factors enumerated by prior researchers, e.g. [40], have guided the selection of variables investigated in our study. *Clear speech*, which consists of speaking at a slower rate and with good (but not exaggerated) enunciation, has also been shown to improve understandability [40], and other behaviors, e.g. avoiding eye contact or speaking too fast, can increase the difficulty of speechreading. Prior autoethnographic work by a hard-of-hearing researcher discussed speechreading challenges when a speaker's face is not visible (on phone calls), with lack of eye-contact (when the speaker is driving a car), or in a dimly lit environment [20]. Other work has revealed that speech recognition can be aided by speaking with variations in pitch [37] or repetition [19].

While prior research suggests guidelines for speaking behavior of hearing individuals [40, 18, 37, 19], prior work has not focused on technology-mediated contexts. Prior work often considers speechreading *accuracy* as a dependent variable, whereas our motivation is to understand the *preferences* of DHH participants, in regard to the behaviors of hearing speakers. Potentially, while some behaviors may aid understanding, DHH participants may subjectively prefer specific behaviors. No prior work had quantitatively examined DHH users' preferences as to how hearing individuals should behave in technology-mediated contexts (speech-recognition apps or video-conferencing). In the next section, we discuss some current technologies that are used to aid SAVOC between DHH and hearing partners.

2.2 Tech-Mediated DHH-Hearing SAVOC

For in-person SAVOC contexts, prior research has investigated the utility of various forms of *technology mediation* during conversations between DHH and hearing individuals, with much of this work investigating automatic captioning technologies based on ASR. While some studies indicated that participants seemed to be generally satisfied with ASR-based automatic captioning for conversations, e.g. [8], in studies with larger numbers of users, e.g. [21], participants expressed frustrations and asked for improvements in both accuracy and design. Berke et al. investigated such technology for one-on-one meetings [4] and how to convey visually when ASR was not performing well, so that users could adapt their communication strategy. Similar to the focus of our research, the Berke et al. [4] study revealed how limitations of access technology can be mitigated, in part, by encouraging hearing and DHH conversational partners to adjust their behavior. Despite limitations, automatic captioning technology for in-person SAVOC has moved beyond the lab, with several commercial services now offering ASR-based automatic captioning services on mobile devices, e.g. [12].

Furthermore, for in-person SAVOC settings, Wang and Piper have studied collaboration between DHH and hearing colleagues in dyads, some with a shared computer, to help communicate and collaborate with each other (e.g. typing on the computer so each person could understand one another) [44]. Their findings, through interviews and collaborative paired activities, provided insight into how DHH and hearing dyads strategize when communicating with each other, with or without technology present, including insight on what worked and didn't work well. Their research was primarily qualitative, however, and our research in this paper, has a quantitative focus on uncovering which specific behavior aspects are preferred by DHH individuals.

For remote SAVOC contexts, several videoconferencing platforms, e.g. [46, 13], also offer features that support DHH users in communicating with hearing colleagues, including: providing integration for human-powered accessibility services (e.g. sign-language interpreters or real-time captioning), text-chat features (which can serve an analogous role to handwritten notes in an in-person context), or automatic captioning. For instance, Microsoft Skype Translator uses ASR to transcribe the speech of hearing participants [24]. Some research has evaluated videoconferencing efficacy between DHH and hearing users [22, 33], including research in non-computing venues, e.g. case studies in specific domains like healthcare [15] or education [39]. Beyond direct audio-visual communication with a hearing partner, DHH users may use videophone services, typically to converse in sign-language. While video relay services enable communication between DHH and hearing users, there is typically no direct audio-visual link between them, with a sign-language interpreter linked to the DHH user via videoconferencing and to the hearing user via telephone. Prior accessibility research has also examined challenges DHH users encounter in this context [42].

2.3 Conversation Behavior Effects on SAVOC

In preliminary work, 8 DHH participants shared preferences for how hearing people should speak during in-person conversation [35]. Participants held short conversations with a hearing actor who exhibited variations in: Speech Rate, Voice Intensity, Over-Enunciating, Eye Contact, Gesturing, and Intermittent Pausing. Our design and methods of that prior study guided the two studies presented in this paper, e.g. a conversation between DHH participants and a hearing actor, the set of behaviors, and some response questions. Our preliminary study had focused on qualitative analysis of open-ended feedback comments as to why each behavior may be important [35]. That study had been too underpowered to compare quantitatively whether users preferred particular levels of these behaviors. This paper provides the first quantitative evidence of preferences among DHH individuals for particular levels of these behaviors, as well as a comparison of DHH users' preferences within in-person versus remote settings.

2.4 Influencing Behaviors in Hearing Speakers

Prior work has provided evidence that hearing people are capable of adjusting their communication patterns. Prior studies by Picheny et al. [30, 31, 32] found that hearing individuals are capable of making changes to their speech to increase its intelligibility for human listeners, but research is needed on methods for sustaining clear speech behaviors over time. Prior research has found that hearing individuals changed their speech characteristics when speaking to non-native speakers [36, 41], children [45], or infants [41]. Some prior work on hearing individuals speaking to ASR technology provides evidence that changes in speech behavior can also be induced in hearing speakers based on notifications generated by software systems. For instance, Oviatt et al. [28] studied hearing individuals who were speaking to software that was recognizing their speech, and they found that when participants were exposed to notifications indicating misrecognition errors in the system, they spoke more deliberately (with greater enunciation) and with more pausing. Stent et al. [38] also found that when hearing speakers were shown system misrecognition errors, they hyperarticulated. Burnham et al. [6] conducted a study where speakers spoke to a computer avatar and discovered that when the avatar indicated that it did not understand, speakers then spoke with greater articulation. One prior study investigated changes in the speaking behavior of hearing people when talking to *both* a DHH person and a technological system: Seita et al. [34] found that hearing participants spoke quicker, louder, with a higher harmonics-to-noise ratio (correlated to higher voice quality), and with non-standard articulation. These studies suggest when presented with a notification that their speech is not understood, including from technological systems, hearing individuals can change their speech behaviors to accommodate listeners.

3 RESEARCH QUESTIONS

Our analysis of prior work revealed that: there is a need for empirical research on technology-mediated conversations between DHH and hearing people, hearing speakers adjust their speaking behaviors when conversing to someone who is DHH with a technological system present, and it is possible for technological systems to prompt hearing people to change their behavior while speaking. Thus, we investigate which of these communication behaviors DHH individuals believe are beneficial for facilitating communication – so that future research may be conducted on technologies for mediating this communication which could prompt or encourage these beneficial behaviors.

Specifically, we investigate the preferences of DHH individuals in the context of a one-on-one conversation with a hearing actor speaking with a variety of different speech and non-verbal behaviors, including: speech rate (fast, medium, slow), speech intensity (loud, medium, quiet), enunciation (over-enunciating, natural, under-enunciating) speech intonation (dynamic pitch, natural, monotone), eye contact (frequent, natural, none), gesturing (2 or more gestures, one gesture, none), and intermittent pausing (long, natural, none). Research questions **RQ1** and **RQ2** examine in-person conversations with an ASR-based application that provides captions on a smartphone screen, and research questions **RQ3** and **RQ4** examine the context of a remote video-conference conversation on a computer.

RQ1: What speech and non-verbal behaviors do DHH individuals subjectively prefer for a hearing person to exhibit when conversing **in-person** when using an automatic captioning application on a smartphone?

RQ2: After experiencing **in-person** ASR-supported conversations, with the hearing individual exhibiting a variety of these behaviors, how would DHH individuals prioritize them?

RQ3: What speech and non-verbal behaviors do DHH individuals subjectively prefer for a hearing person to exhibit in **remote** conversation using a videoconferencing platform?

RQ4: After experiencing **remote** videoconferencing conversations, with the hearing individual exhibiting a variety of these behaviors, how would DHH individuals prioritize them?

4 METHODOLOGY

We conducted two experimental studies: one in-person and one remote. Given similarities in the methods of the two, we first discuss aspects that differ between the two studies in separate sections, **Section 4.1** and **Section 4.2**, one for each study. Then, an overview of the remainder of the methodology common among both studies is discussed in **Section 4.3.**

4.1 In-Person Experimental Study Details

A total of 20 participants were recruited for the in-person study by advertising on Facebook groups "NTID Community" and "RIT Cross-Registered Community". The former group consists of Facebook users who are part of the National Technical Institute for the Deaf (NTID) and the latter consist of DHH users who are supported by NTID (e.g. by interpreting or captioning services) but enroll at RIT. Participants were also encouraged to share information about the study via word-of-mouth. Of the 20 participants, 7 identified as male, 12 identified as female, and 1 identified as non-binary. The median age was 23, and ages ranged from 19 to 36. Regarding hearing identity, 7 identified as culturally Deaf [29], 5 identified as deaf and 8 identified as hard-of-hearing. Eleven reported congenital deafness, 8 reported acquired deafness, and 1 was unsure. Six used cochlear implants, 6 used hearing aids, two used hearing devices but unfortunately our records did not indicate which one, and 6 did not use any hearing devices. All but 3 were unfamiliar with Live Transcribe. Of our participants, 9 had a high school diploma, 7 had an associate degree and 4 had a bachelor's degree. It should be noted that all 10 participants with a high school diploma were in progress of obtaining another degree at a college or university. Demographic data appears at http://latlab.ist.rit.edu/w4a2021

The in-person study utilized an ASR-based app to support SAVOC between DHH participants and the hearing actor. On arrival, a DHH participant signed an informed consent form, and then they were asked demographic questions. Next, participants were handed a Google Pixel smartphone with the Google Live Transcribe app [12] pre-installed. In the next few minutes, the DHH participant was provided with a demonstration of how the Google Live Transcribe app works. Participants were shown that the words the hearing person speaks appear on the upper half of the screen, and they would be able to use the phone's keyboard to type any responses which would show up on the bottom half of the screen. A photo of the app used is shown in Figure 1 (a).

Participants were instructed that the ASR application was there to aid communication, if they did not require the app to engage in conversation, they did not have to use it. The goal was to try to encourage the DHH participants to converse with the hearing actor as they would normally in everyday life, as naturally as possible. Thus, all users had a phone with them with Live Transcribe running for the in-person study, but individuals may have differed in the degree to which they looked at the text on the phone. A photo of a DHH participant with the researcher conducting the interview as well as the hearing actor is shown in Figure 1 (b). From this point on, the in-person experimental study proceeded as outlined in **Section 4.3 Common Methodology**.

(a) (b)


Figure 1: (a) Screenshot of Google Live Transcribe [12] app used in in-person study. (b) Photo of a DHH participant (on the right) holding a smartphone with Google Live Transcribe [12] installed, communicating with a hearing actor (left). The researcher is sitting outside the picture to the left of the actor.

4.2 Remote Experimental Study Details

A total of 23 participants were again recruited primarily from the same Facebook groups. Of the 23 participants, only one participant also participated in the in-person study; the other 22 were a completely different set of participants. A similar advertisement as described in **Section 4.1.1** was used, except this time it was explained that the study would take place virtually via videoconferencing instead of in-person. Participants were also encouraged to spread information about the study via word-of-mouth. Of the 23 participants, 12 identified as male and 11 identified as female. The median age was 26, and ages ranged from 21 to 35. Regarding hearing identity, 9 identified as culturally Deaf [29], 6 identified as deaf and 6 identified as hard-of-hearing. Two participants identified as both Culturally Deaf *and* hard-of-hearing. Nineteen reported congenital deafness, 1 reported acquired deafness, and 2 were unsure. Five used cochlear implants, 7 used hearing aids, and 11 did not use any hearing devices. Demographic data also appears at http://latlab.ist.rit.edu/w4a2021. Of our participants, 2 had a high school diploma, 2 had an associate degree, 11 had a bachelor's degree, 7 had a master's degree and 1 had a Doctorate degree. Some notable differences in participant demographics between the in-person and remote studies include: a more even distribution between males and females (the in-person study had more females than males), the median age being 3 years higher in the remote study (26 versus 23), and more participants had completed higher education degrees in the remote study.

The remote study utilized a videoconferencing platform called Zoom [46] to enable face-to-face communication between DHH participants and the hearing actor, and DHH participants could use the built-in chat function to communicate. Zoom was selected for this experiment as it was popular and well-known. The chat functionality it provided was common among many platforms. Zoom's chat function was the non-speech technology-mediated communication method provided between participant and actor.
ASL interpretation was available only as a last-resort backup, to ethically avoid subjecting participants to the frustration of a Zoom meeting with someone they couldn't understand, and ASL translation was available only upon request of the participant and only after the actor had spoken. The DHH participant was emailed a consent form to read and sign virtually. After signing, participants were given a link to connect through Zoom. Once connected via videoconferencing, they were asked to answer demographic questions. Next, participants were asked to ensure that they were on a laptop and Zoom was set to Gallery mode so that they would be able to clearly see both the researcher conducting the interview and the hearing actor. An example of how the computer screen looked is shown in Figure 2; the participant's face is blocked to protect privacy. They were also asked to confirm that the video quality was clear, and that the audio was connected. From this point on, the in-person experimental study proceeded as outlined in Section 4.3.

Figure 2: A photo of one DHH participant (pictured top left), the researcher conducting the interview (pictured top right), and the hearing actor (pictured center bottom) connected via Zoom [46]. Gallery mode is enabled to ensure that all three faces can be made large enough and seen equally during the experiment.

4.3 Common Methodology

In both studies, technology-mediated synchronous conversations occurred between DHH participants and a hearing actor who played the role of a hearing conversational partner. Each study was IRB-approved, and a native ASL-signer was also present to moderate the session, ask the DHH participant interview questions, and record responses. Each 60-minute study primarily consisted of brief back-and-forth conversational exchanges between the DHH participant and the hearing actor. During each exchange, the hearing actor adjusted their behavior to exhibit a specific behavior. There were 7 behavioral categories, and each category has 3 ***sub-behaviors***, for a total of twenty-one specific behaviors. We scheduled a total of 21 conversational exchanges between the participant and actor, one for each behavior. To illustrate the actor's performance for these behaviors, a set of videos of the actor behaving as they did in the experiment are included at http://latlab.ist.rit.edu/w4a2021. Six of our conversational behaviors were the same as in the preliminary work of [35], and we added a seventh category: Intonation Dynamics (referred to as Intonation for brevity). For completeness, the 7 conversational behaviors and the 3 sub-behaviors we tested are outlined below:

1. **Speech Rate:** How fast the hearing actor should speak, with sub-behaviors: *Speech Rate – High, Speech Rate – Medium, and Speech Rate – Low*. The actor spoke faster than normal, at a typical pace, and slower than normal.
2. **Voice Intensity:** How loudly the hearing actor should speak, with sub-behaviors: *Intensity – High, Intensity – Medium, and Intensity – Low*. The actor spoke louder than normal, at a typical volume, and quieter than normal.
3. **Enunciation:** How the hearing actor should enunciate (the extent to which they move their lips while speaking), with sub-behaviors: *Enunciation – High, Enunciation – Medium, and Enunciation – Low*. The hearing actor over-enunciated, enunciated normally, and under-enunciated, respectively.
4. **Intonation:** How much tone-inflection and emotion the actor should exhibit when speaking, with sub-behaviors: *Intonation – High, Intonation – Medium, and Intonation – Low*. The hearing actor spoke with a great deal of tone inflection, spoke naturally, and spoke in monotone, respectively.
5. **Eye Contact:** How much eye contact the hearing actor should make with the DHH participant while speaking, with sub-behaviors: *Eye Contact – High, Eye Contact – Medium, and Eye Contact – Low*. The hearing actor spoke with constant eye contact, with naturally intermittent eye contact, and without any eye contact, respectively.
6. **Gesturing:** How much hand-gesturing the hearing actor performed, with sub-behaviors: *Gesturing – High, Gesturing – Medium, and Gesturing – Low*. The actor spoke while performing several hand gestures, while performing only one gesture, and without any gesturing at all, respectively.
7. **Intermittent Pausing:** How long the actor paused between sentences or concepts, with sub-behaviors: *Intermittent Pausing – High, Intermittent Pausing – Medium, and Intermittent Pausing – Low*. The actor spoke while pausing for at least one second between each sentence or concept, while pausing naturally for about half a second between each sentence or concept, and with minimal pausing, respectively.

The purpose of the IRB-approved study was explained at the beginning, and participants were informed that they would have brief back and forth conversations with a hearing actor, in which they would ask the hearing actor some questions (from a list provided). The actor would respond (displaying one of the sub-behaviors), and the DHH participant would provide a rating of how satisfied they were with actor's conversational behaviors for each response. An overview of the back-and-forth structure of this interaction between the DHH participant and hearing actor was explained to the participant in advance. In order to determine which specific behaviors and sub-behaviors were preferred by the DHH community, each behavior was tested according to the following structure, similar to that of [35]:

1. The native-ASL-signer researcher instructed the DHH participant to ask a specific question ("please ask the question: *What is your favorite cuisine and why?*") to the hearing actor either verbally or by typing into the ASR application (**for the in-person study**), or either verbally or using the chat function on the Zoom application (**for the remote study**). Participants used the option that felt most comfortable for them. **The list of questions used in this study is detailed later in this section of the paper**.
2. The hearing actor responded to the question with a specific sub-behavior, e.g. *Gesturing – High*.
3. The DHH participant observed this response, and if needed to aid understanding, utilized the ASR application (**for the in-person study**) or sign-language interpretation from the researcher conducting the study (**for the remote study**) to understand what the hearing actor said.
4. The DHH participant was asked questions about their perspective regarding this behavior, and responses to these questions are used to answer **RQ1** and **RQ3**. **The set of questions asked are listed in a later section**.
5. The DHH participant was instructed that the hearing person would repeat their answer a second time but change their behavior when speaking, and steps 2-4 were repeated for a different sub-behavior, e.g. *Gesturing – Medium*.
6. The DHH participant was instructed that the hearing person would repeat their answer a third and final time, again changing their behavior when speaking. Steps 2-4 were repeated for the third sub-behavior which had not yet been tested, e.g. *Gesturing – None*.
7. After this third repeated answer, participants were informed of what the behavioral differences were in each response.
8. Steps 1-7 are repeated for each of the other behavior categories, e.g. if *Gesturing* was just completed, then the steps are repeated for *Intermittent Pausing*, and so on. The order of speech behaviors and question prompts was rotated among participants using a Graeco-Latin scheduling to balance assignment of conditions and avoid ordering effects.
9. Participants were asked more questions after experiencing each behavior, including **open-ended feedback** and which behaviors hearing people should prioritize focusing on. Their responses for these questions are answer our research questions **RQ2** and **RQ4**. More detail about what questions were asked is explained later in this section of the paper.
10. The study was then concluded, and participants were paid $40 in compensation and thanked for their time.

In preparation for the experiment, the researcher and hearing actor practiced extensively. During experiments, the researcher monitored the actor's behavior, and in those few cases (fewer than 5 total) in which the actor did not perform correctly, the actor was asked to repeat. We have video recordings for the in-person sessions, but our rapid conversion of our IRB protocol for the Zoom/remote study did not allow us to include video recording for the remote study. After the conclusion of the experimental study, we had a hearing researcher do a spot check of six participants' run-throughs of the experiment for the in-person case. For each of the six participants, the hearing actors' behavior was analyzed and graded based on whether the hearing actor exhibited each behavior at an *Adequate* (acceptable, each of the three behavior levels are distinct) level, or if there was a *Possible Issue* (the behavior levels displayed are not be distinct enough). Each participant experienced 21 interactions (7 questions and three levels per question) so each of the six spot checks had 21 grades total. Of the six randomly chosen participants, three of them had 21 out of 21 Adequate scores, and three of them had 20 out of 21 Adequate scores. Overall, 3 out of 126 interactions yielded a possible issue with the exhibition of speaking behavior, thus 97.62% of interactions were Adequate.

The conversation prompts DHH participants asked the hearing actor were designed to elicit responses of only a few sentences. The question prompts used during the study were the same as asked in [35] and are repeated below:

- What is your favorite cuisine and why?
- What do you like about the city of Rochester?
- What do you NOT like about the city of Rochester?
- Where do you want to travel to next, and why?
- What were your plans for the recent winter holidays?
- What sports do you play, and which is your favorite?
- How many family members do you have, and who?

After each of the 21 brief conversational exchanges, the DHH participant was asked "How satisfied were you with the speech/conversational behavior of the hearing person?" responses indicated on a scale of 1 (very unsatisfied) to 10 (very satisfied). Participants were also asked open-ended feedback questions regarding the reasoning behind their responses. The responses to this question are used for **RQ1** and **RQ3 – Determining which behaviors DHH participants subjectively prefer to see from their hearing conversational partners**.

After experiencing the 7 main categories of behaviors from the hearing actor', participants assigned **Priority Scores** for the relative importance of the hearing person exhibiting each behavior correctly. Participants gave each of the 7 categories a priority score between 1 and 7, with 1 indicating low and 7 indicating high priority. These priority scores are used when answering research questions **RQ2** and **RQ4 – Determining which behaviors DHH people want hearing people to prioritize in conversation**.

The reason for having only brief interactions was twofold. Firstly, we did not want to cognitively overload participants with information and stimuli. Since we had 21 total behaviors to test, having longer interactions was infeasible due to time constraints. During our experiments, each case only required about 2 minutes each to complete. Participants answered questions about their experience immediately following each interaction, so they didn't have to remember prior interactions. Thus, all participants were able to finish the whole experiment in under 60 minutes and no participant reported any issues of fatigue or cognitive overload.

Secondly, the hearing actor already had to practice extensively to exhibit appropriate behaviors while controlling for other behaviors. Having a longer, more natural dialogue would make this even more difficult for the actor to maintain consistency, as not only would they have to think about how to exhibit behaviors but also about what to say. This could also potentially increase time for the participant to complete the experiment as this dialogue by nature would last longer, and natural dialogue would result in different conversations taking place for each participant, adding another variable of unpredictability.

5 ANALYSIS AND RESULTS

In this section, we analyze and discuss our results. The analysis of our results for the in-person study and the remote study are discussed separately. The in-person study addresses research questions **RQ1 and RQ2**: which behaviors DHH participants subjectively prefer and want the hearing person to prioritize when communicating in person, while the remote study addresses questions **RQ3 and RQ4:** which behaviors DHH participants subjectively prefer and want the hearing person to prioritize when communicating remotely via videoconferencing.

Recall that for each behavior's sub-behavior, participants were asked the question: "How satisfied were you with the speech/conversational behavior of the hearing person?" and responded on a scale of 1 (very unsatisfied) to 10 (very satisfied). For each behavior, the three sub-behaviors were compared using a Friedman test to see if participants strongly preferred one sub-behavior over another. If the test returned a significant result ($p<0.05$), post-hoc Wilcoxon signed rank tests were performed to reveal any pairwise significant differences. A Bonferroni correction was applied to the pairwise comparisons, so the threshold for significance became $p<0.05/3 = p<0.0167$. Participants had also been asked to assign a priority score, from 1 (low) to 7 (high), to each behavior based on how important they thought an appropriate level of each behavior was in helping them understand a hearing person in conversation. A Bonferroni correction was applied, since we made 21 comparisons, and the threshold for significance became $p<0.05/21 = 0.0024$.

We clarify here that no prior experimental work has provided a quantitative basis for understanding such preferences of DHH users in technology-mediated contexts. Some participant comments were shared to complement the quantitative data, but we do not claim a rigorous qualitative analysis as a contribution.

5.1 In-Person Experimental Study Results

5.1.1 Subjective Preference for Each Behavior (RQ1)

The results for our statistical analysis of satisfaction scores for the in-person study are shown in Figure 3(a). The behaviors **Enunciation (p=0.0051)** and **Intonation (p=0.0045)** both yielded significant omnibus Friedman test results. Pairwise post-hoc Wilcoxon signed-rank tests are presented below:

1. __**Enunciation**__: Post-hoc pairwise tests revealed a significant difference between a **Medium** and **Low** level (**p=0.0003**, with **Medium** ranked higher). There was consensus in open-ended responses that the actor was hard to understand when they under-enunciated. P14 said "*When there's not enough lip enunciation, it's too difficult for me to understand. I liked the normal amount of enunciation.*"
2. __**Intonation**__: Post-hoc pairwise tests revealed a significant difference between a **High** and **Low** level (**p=0.0032**, with **High** ranked higher). Participants preferred dynamic intonation, as this helped the flow of conversation, kept them engaged, or supported a connection with the speaker. P13 explained "*I like it when the participant speaks when using lots of tone inflections over monotone. It makes the conversation more interesting and engaging for me.*" Another participant (P12) mentioned that "*I like high tone inflections because I feel like it will result in a better back and forth conversation and I feel more connected with the person I am talking to.*"

(a)

(b)

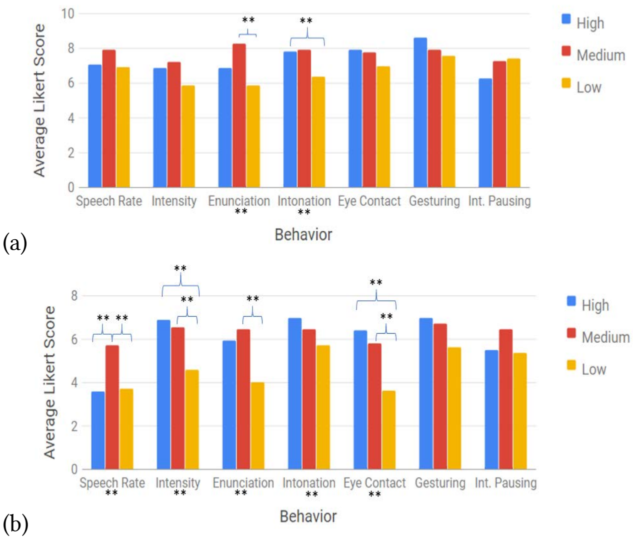

Figure 3: Satisfaction scores in (a) in-person and (b) remote study. For significant Friedman test results, a double asterisk (**) appears below the behavior on the x-axis. Pairwise significant differences are indicated with a double asterisk (**) above the graph.

The behaviors Speech Rate (p=0.089), Intensity (p=0.14), Eye Contact (p=0.062), Gesturing (p=0.14), and Intermittent Pausing (p=0.26) did not yield significant differences. For these, participants did not strongly prefer one any one sub-behavior over another (e.g. for Speech Rate, there was no strong preference for any one of: speaking fast, speaking at a medium pace, or speaking slowly). Since there were no significant results for these behaviors, pairwise post-hoc tests were not performed.

5.1.2 Prioritization of Behaviors (RQ2)

A Friedman test was applied on the Priority response data with a significant result (p=0.004). Post-hoc Wilcoxon signed-rank tests were done on each pair to determine any pairwise significant differences. The pairs **Enunciation and Intermittent Pausing** (p=0.0014, with **Enunciation** ranked higher) and **Eye Contact and Intermittent Pausing** (p=0.0023 with **Eye Contact** ranked higher) were significant. The behaviors Enunciation and Eye Contact were both consistently rated higher than Intermittent Pausing, indicating that DHH participants thought that hearing people should focus on appropriately exhibiting Enunciation and Eye Contact, rather than focusing on Intermittent Pausing, in this in-person conversation context.

5.2 Remote Experimental Study Results

5.2.1 Subjective Preference for Each Behavior (RQ3)

The results for our statistical analysis of satisfaction scores for the remote study are shown in Figure 3(b). The behaviors Speech Rate (p=0.0077), Intensity (p=0.00098), Enunciation (p=0.00059), Intonation (p=0.043), and Eye Contact (p=0.00002) all yielded significant omnibus Friedman test results. A summarization of pairwise post-hoc Wilcoxon signed-rank tests, for each behavior, are presented in the following list:

1. **Speech Rate**: Post-hoc pairwise tests revealed a significant difference between a **High** and **Medium** level (**p=0.0012**, with **Medium** ranked higher), as well as a **Medium** and **Low** level (**p=0.0071**, with **Medium** ranked higher). Some expressed that fast speech was difficult to understand and slow speech was unhelpful, with one participant (P9) saying "*I prefer when the speech rate is not too fast or too slow. Too quickly makes it harder to lip-read which I rely on and speaking too slowly makes the speaker look stupid to me.*"
2. **Intensity**: Post-hoc tests revealed a significant difference between a **Medium** and **Low** level (**p=0.0035**, with **Medium** ranked higher), as well as a **High** and **Low** level (**p=0.00038**, with **High** ranked higher). Participants preferred louder speech, with P8 saying "*I can understand louder voices more easily and it's easier for me to lipread. Softer voices are harder to hear.*" Even some participants who could not hear the volume differences preferred louder intensity, e.g. P5 said "*Due to the larger mouth movements when people are speaking louder, that's why I ended up preferring that one.*"
3. **Enunciation**: Post-hoc pairwise tests revealed a significant difference between a **Medium** and **Low** level (**p=0.00026**, with **Medium** ranked higher). Participants agreed that enunciating normally was better than under-enunciating due to the difficulty of understanding the speaker in the latter case, with P1 saying "*It's easier to follow what's being said when enunciation is normal. When it's too low, it is the most difficult to understand. Over-exaggeration may be somewhat alright but it's just better to speak normally.*"
4. **Intonation**: Post-hoc tests did not reveal significant pairwise differences; however, open-ended feedback revealed some participants' views. P14 explained that although they have trouble discerning tone, they preferred tone inflections: "*I had a hard time telling the difference with the tone levels but I still prefer the hearing person speak with a dynamic tone because emotions and feelings show in your face and body language. If you're speaking monotone you have a blank expression and I can't really work with that.*"
5. **Eye Contact**: Post-hoc pairwise tests revealed a significant difference between a **Medium** and **Low** level (**p=0.00084**, with **Medium** ranked higher), as well as a **High** and **Low** level (**p=0.00008**, with **High** ranked higher). Participants preferred that the hearing actor speak with either a maintained or a natural amount of eye contact, and many explained that eye-contact was important due to many DHH people being more visually oriented. P15 commented: "*Especially as a deaf person, eye contact is important to show recognition and that you're paying attention to the person. But perhaps you shouldn't use full-on eye contact, being casual and glancing around at your surroundings is normal as well.*"

The behaviors Gesturing (p=0.054) and Intermittent Pausing (p=0.30) did not yield significant differences. Participants did not strongly prefer one any one sub-behavior over another, e.g. for Gesturing, there was no strong preference for any one of speaking with >1 gesture, speaking with only one gesture, or speaking without any gesturing. Since there were no significant results for these behaviors, pairwise post-hoc tests were not performed.

5.2.2 Prioritization of Behaviors (RQ4)

A Friedman test was applied on the Priority responses, with a significant result (p=0.00031). Post-hoc Wilcoxon signed-rank tests were done on each pair to determine any pairwise significant differences. Overall, two pairs yielded significant results: **Speech Rate and Eye Contact** (**p=0.00054** with **Eye Contact** rated higher priority) and **Eye Contact and Intermittent Pausing** (**p=0.0010** with **Eye Contact** rated higher priority). The behavior Eye Contact was rated significantly higher than both Speech Rate and Intermittent Pausing, indicating that DHH participants felt that hearing people should focus on appropriately exhibiting Eye Contact, rather than focusing on either Speech Rate or Intermittent Pausing, in this remote context.

6 DISCUSSION

In this section, we discuss our observations in relation to the results of both studies[1]. Overall, our participants preferred either the high or medium range over the low range for behaviors. This was the case for the behaviors Enunciation and Intonation in the in-person study, and the behaviors Intensity, Enunciation, and Eye Contact in the remote study. The medium range was preferred to both the high and low ranges only for Speech Rate in the remote study. This indicates that generally participants preferred either a high or a normal level of each behavior over a low level. Our data does not indicate whether a high or a normal level is preferable for most of the behaviors analyzed. When considering DHH user's satisfaction with the hearing actor's behavior, we observed a greater number of significant omnibus and post-hoc results in our remote study (speech rate, intensity, intonation, eye contact, enunciation) than in the in-person study (enunciation and intonation). Two behaviors (enunciation and intonation) had a significant effect on DHH user's satisfaction in both studies, and in the remote study, we observed significant effects for eye contact, intensity, and speech rate.

[1] It is important to note that our in-person and remote studies were conducted independently, with our remote study inspired by world events relating to COVID-19, after the data-collection and statistical analysis for our in-person study had already been concluded. We therefore disclaim that no statistical significance testing was conducted upon a merged dataset across both studies. In principle, a meta-analysis could be conducted, i.e. treating *in-person* and *remote* as two levels of a single factor, using a non-parametric two-factor "mixed" (between, within) difference testing.

- **Eye Contact:** In regard to why eye contact had a significant effect in the remote study, but not in our in-person study, we speculate this was due to lower visual dispersion in the remote study. The in-person study occurred in a three-dimensional real-world setting, and the DHH user held a smartphone device with the captioning while sitting at a table with a hearing actor. Thus, there was a rather large difference in visual angle between the actor's face and the device, and participants may have spent some of the time looking down on the phone, thereby missing rapid changes in eye contact. Whereas, in the remote study, the DHH user was looking at a computer screen with the videoconferencing software, with the chat panel of the interface adjacent to the video of the actor's face, with a smaller visual angle between these two areas of interest.
- **Intensity:** Voice intensity (loudness) also had a significant effect on satisfaction responses in our remote study. We speculate that in the remote context, users were able to control the volume level on their computer, to set it to a desired level for the meeting, which may have made it easier for users to discern differences in voice intensity along the range of the actor's voice. In addition, we noticed a slightly higher percentage of participants in the remote study were wearing assistive listening devices (e.g. hearing aids or cochlear implants), when compared to the in-person study (about 70% versus 65%). These users may have been better able to differentiate among the voice intensity sub-behaviors.
- **Speech Rate:** Speech Rate also had a significant effect in the remote study but not the in-person study. We speculate that perhaps the speech rate, like eye contact, was impacted by the lower visual dispersion in the remote study, as people may have relied on observing how fast the lips were moving to help deduce the speech rates of the hearing actor.

6.1 Discussion of Relevant Cultural Factors

Considering several cultural factors among the Deaf community may illuminate some of our findings: For instance, while no significant difference in DHH users' satisfaction was observed between high or medium levels of Enunciation, many DHH individuals have strong opinions of this behavior. It is a common sentiment within the DHH community that over-enunciating makes the hearing person seem rude and condescending. Thus, discretion must be used even though the statistics don't favor one sub-behavior over another, and in this particular example, care should be taken to avoid appearing to over-enunciate.

Given the importance of hand movements and body language in DHH culture, it was surprising that Gesturing had no significant effect on DHH user's satisfaction, in either study. There may be diverging opinions on the effectiveness of hearing people using gesturing during conversations. It is possible that some in the DHH community would benefit from some gesturing from the hearing person. However, it is also possible that the hearing person may gesture in a manner that is confusing or misleading, or even attempt to use sign language incorrectly. Analysis of priority scores revealed strong preferences for prioritizing eye contact. This is unsurprising due to the heavy importance of eye contact in Deaf Culture, as ASL is a very visual language involving hand movements and facial expressions.

6.2 Discussion of Prior Work

Our findings align with some prior speechreading work [40, 37, 19] discussed in Section 2.1, which found speechreading is easier when the speaker talks at a slower rate with non-exaggerated enunciation, dynamic pitch, and maintained eye contact. These behaviors that benefit speechreading also align with the DHH community's subjective preferences with enunciation (participants preferred enunciating normally than under-enunciating), dynamic pitch (participants preferred a dynamic tone over speaking monotonously), and eye contact (participants preferred maintained or natural eye contact over low eye contact). Our data, however, did not enable us to differentiate whether over-enunciating was better or worse than enunciating normally, nor whether speaking with highly dynamic tone was better or worse than speaking with regular conversational pitch.

On the other hand, our results differed somewhat from prior work on speechreading in that our participants strongly preferred speaking at a medium rate over a slower rate. We speculate that while slower speech may make speechreading easier, there is a risk that it may be perceived as rude or patronizing. Overall, our data revealed that while good practices for speechreading often align with DHH preferences, there may be cases in which some behaviors that are ideal for speechreading are not preferred by the DHH community in this setting. Furthermore, while there may be some generally accepted guidelines for speaking behavior, there will always be varying subjective preferences for each behavior among specific individuals, and it is important to not exclude those DHH individuals who disagree with majority preferences.

The work presented in this paper also ties back to the paper by Wang and Piper [44] which had found that sometimes DHH and hearing colleagues, in collaborative settings, find themselves in positions where one or both of them are in unfamiliar territory and one of them has to adjust to fit the needs and expectations of the other. In practice, this burden of adjustment often falls upon the shoulders of DHH individuals, in large part due to DHH people living in a largely hearing world, whereas a typical hearing person has little to no experience with the DHH community. Wang and Piper put it well when they said, quote "*... while signed communication is often more comfortable for Deaf signers...[however] Deaf professionals may sacrifice accuracy and comfort by choosing to lipread instead of burdening their hearing collaborators to change their behavior.*" [44] This highlights the need for the kind of research presented in this paper: future HCI researchers can help ease the burden and the pressure on the DHH community to adapt to hearing norms, in the form of designing software and applications that are able to encourage beneficial behaviors so that DHH individuals do not need to devote time and energy in teaching and acclimating their hearing colleagues. To help accomplish that goal, the research in this paper provides a list of behaviors that should be encouraged in hearing speakers, based on the subjective preferences of our DHH participants.

6.3 Implications of these Findings

The findings of this study are important for establishing that certain behaviors of hearing people during technology-mediated conversation are more desirable by the DHH community, and our findings inform guidelines as to which behaviors hearing people should be encouraged to exhibit while speaking to DHH people in such settings. Thus, our findings may be useful for providing training or recommendations to hearing individuals who work or collaborate with DHH colleagues in educational or employment settings, during technology-mediated conversation.

As discussed in Sections 1 and 3, our research findings suggest an agenda for future HCI research. Since prior work has found that hearing individuals do change their behaviors when interacting with DHH individuals and technology, and since prior work has established that prompting or notification technologies can support behavior changes, the findings from our study suggest a new avenue of research: Future research could examine prompting or notification systems to influence hearing speakers so that they communicate in a manner that is more easily understandable by DHH users. The findings of our study suggest which categories of speech and non-verbal behavior may be most fruitful for such research. DHH individuals currently bear much of the responsibility during conversations to indicate when their hearing partner may be engaging in behaviors that impede communication, such technology could more equitably distribute such responsibility among all parties in conversation. Of course, future research would need to determine whether such prompting or notification tools lead to awkwardness or if they do provide benefits for DHH users, thereby providing DHH users with better accessibility and more independence in impromptu conversations.

7 CONCLUSIONS AND FUTURE WORK

We investigated the perspectives and preferences among DHH individuals as to how hearing people should behave while in conversation, including which behaviors they should prioritize, for in-person conversations using ASR and remote conversations using videoconferencing. Our in-person and remote studies revealed that participants had significant preferences for sub-behaviors of Enunciation and Intonation. The remote case had additional significant results, revealing preferences for sub-behaviors for Speech Rate, Voice Intensity, and Eye Contact.

One limitation of this study is that we only examined three levels for each category of behavior in the study (High, Medium, and Low), but in cases in which users preferred the high or low option, we do not know whether their ideal preference is actually outside of this range. Further, the three-level design did not enable fine-grained comparisons of subtle differences in these behaviors, which could fall in-between these three positions in the spectrum. While the levels of sub-behaviors had been intended to provide a wide range for DHH users to experience, without going so far as to be upsetting for the participant or the hearing actor, e.g. requiring the actor to scream at a participant or to over-enunciate their speech to a degree that would appear insulting. Furthermore, while the hearing actor had rehearsed extensively how to replicate these various performances, since the hearing actor answered each question live, their exhibition of each of sub-behavior varied to some degree. While we could have shown recordings of the hearing actor speaking, to control for such variations, we were concerned that this would feel too impersonal, and unlike a real interaction with a live person –which could have affected the validity of the responses from participants. Finally, the same hearing actor was used throughout this study, but it is possible that variations in the actors' demographics (age, gender, etc.) may influence perceptions of the DHH participants.

One limitation is that our study did not investigate combinations of multiple factors simultaneously. We tested each behavior category individually and the hearing actor did their best to focus on one behavior at a time while acting as naturally as possible for the remaining behaviors. A multi-factor study could reveal potential interaction effects. Another limitation is that while we focused on behaviors of the hearing actor, behaviors of DHH participants were not analyzed, and future work could examine this. Another limitation is that conversations were very structured in order to test each of the individual behaviors in a controlled manner and to reduce the cognitive load of the hearing actor. Future work is needed to determine whether our results generalize to settings in which people communicate in less structured manners, e.g., in impromptu meetings at the workplace. Such studies could take place in field studies where researchers observe individuals interacting with each other in real-life work settings or virtual meetings. One more thing to consider is that we did not test whether the audio and visual were perfectly synced up on the participants' end. We received no complaints, but there is a slight possibility that participants did not report issues with audio-visual sync.

In addition, our study did not investigate a control group of people who are not DHH, and as such we cannot state whether the preferences of DHH individuals overlap with those of hearing people, which can be investigated in future work. Finally, there is diversity within the DHH community, which motivates future work to gather judgments from a wider cross-section of the community, to compare findings across different sub-groups based on some aspect of their hearing or conversational abilities. Simply categorizing DHH participants based on their self-identified hearing status can be problematic, as there can be diversity in the actual level of hearing loss among people who identify themselves similarly. In addition, users' preferences may vary depending upon whether they are wearing their hearing aids or cochlear implants. Our sample size of 20 and 23 participants did not provide sufficient power for such sub-group comparison, and a future study employing assigning DHH participants to several sub-groups based on conversational abilities, with a large enough sample size for each sub-group, could reveal interesting differences in users' preferences for the behavior of their hearing conversational partner.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation (NSF) under award No. 1954284 and the NSF Graduate Research Fellowship under award No. 1746056.

REFERENCES

[1] Altieri NA, Pisoni DB, Townsend JT. Some normative data on lip-reading skills (L). J Acoust Soc Am. 2011;130(1):1-4. doi:10.1121/1.3593376

[2] S. Antia, D. Sabers, M. Stinson. 2007. Reliability and validity of the classroom communication questionnaire. Journal of Deaf Studies and Deaf Education. 12, 2, (March 2007), 158-171. DOI: http://dx.doi.org/10.1093/deafed/enl028.

[3] Jon P. Barker, Ricard Marxer, Emmanuel Vincent, Shinji Watanabe. 2017. The CHiME challenges: Robust speech recognition in everyday environments. In: Watanabe S., Delcroix M., Metze F., Hershey J. (eds.), New Era for Robust Speech Recognition. Springer, Cham, 327-344. https://doi.org/10.1007/978-3-319-64680-0_14.

[4] Larwan Berke, Christopher Caulfield, and Matt Huenerfauth. 2017. Deaf and Hard-of-Hearing Perspectives on Imperfect Automatic Speech Recognition for Captioning One-on-One Meetings. In Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '17). ACM, NY, NY, USA, 155–164. https://doi.org/10.1145/3132525.3132541.

[5] Debra L Blackwell, Jacqueline W Lucas, and Tainya C Clarke. 2014. Summary Health Statistics for U.S. Adults: National Health Interview Survey, 2012. National Center for Health Statistics. https://www.ncbi.nlm.nih.gov/pubmed/24819891.

[6] Denis Burnham, Sebastian Joeffry, Lauren Rice. 2010. Computer- and Human-Directed Speech Before and After Correction. In Proceedings of the 9th Speech Science and Technology Conference 2010, Melbourne, Australia. Australian Speech Science and Technology Association.

[7] Herbert H. Clark, Meredyth A. Krych, Speaking while monitoring addressees for understanding, Journal of Memory and Language, Volume 50, Issue 1,2004, Pages 62-81, ISSN 0749-596X, https://doi.org/10.1016/j.jml.2003.08.004.

[8] Lisa B. Elliot, Michael Stinson, Syed Ahmed, and Donna Easton. 2017. User Experiences When Testing a Messaging App for Communication Between Individuals Who Are Hearing and Deaf or Hard of Hearing. In Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '17). ACM, NY, NY, USA, 405–406. https://doi.org/10.1145/3132525.3134798

[9] Lisa B. Elliot, Michael Stinson, James Mallory, Donna Easton, and Matt Huenerfauth. 2016. Deaf and Hard of Hearing Individuals' Perceptions of Communication with Hearing Colleagues in Small Groups. In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '16). ACM, NY, NY, USA, 271–272. https://doi.org/10.1145/2982142.2982198.

[10] C. Garberoglio, S. Cawthon, M. Bond. 2016. Deaf People and Employment in the United States. Retrieved August 4, 2017 from https://www.nationaldeafcenter.org/sites/default/files/resources/Deaf%20Employm ent%20Report_final.pdf.

[11] Carrie Lou Garberoglio, Stephanie Cawthon, and Adam Sales. 2017. Deaf People and Educational Attainment in the United States: 2017. Technical Report. Washington, DC.

[12] Google Live Transcribe. 2019. Introducing Live Transcribe. Retrieved May 7, 2020 from https://www.android.com/accessibility/live-transcribe/

[13] Google Meet. Retrieved May 7, 2020. https://meet.google.com/

[14] Benjamin M. Gorman and David R. Flatla. 2018. MirrorMirror: A Mobile Application to Improve Speechreading Acquisition. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18). Association for Computing Machinery, New York, NY, USA, Paper 26, 1–12. DOI:https://doi.org/10.1145/3173574.3173600

[15] Gournaris, M. J., & Leigh, I. W. (2019). Comparison of Face-to-Face and Video-Mediated Communication with Deaf Individuals: Implications for Telepsychotherapy. JADARA, 37(2). Retrieved from https://repository.wcsu.edu/jadara/vol37/iss2/5

[16] Jennifer Harris & Claire Bamford (2001) The Uphill Struggle: Services for Deaf and hard of hearing people - issues of equality, participation and access, Disability & Society, 16:7, 969-979, DOI: 10.1080/0968759012009785

[17] Hearing Loss Association of America. 2017. Basic Facts About Hearing Loss. Retrieved December 17, 2017 from http://www.hearingloss.org/ content/basic-facts-about-hearing-loss.

[18] Hnath-Chisolm T, Boothroyd A. Speechreading enhancement by voice fundamental frequency: the effects of Fo contour distortions. J Speech Hear Res. 1992;35(5):1160-1168. doi:10.1044/jshr.3505.1160

[19] IJsseldijk, F. J. (1992). Speechreading performance under different conditions of video image, repetition, and speech rate. Journal of Speech & Hearing Research, 35(2), 466–471. https://doi.org/10.1044/jshr.3502.466

[20] Dhruv Jain, Audrey Desjardins, Leah Findlater, and Jon E. Froehlich. 2019. Autoethnography of a Hard of Hearing Traveler. In The 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19). Association for Computing Machinery, New York, NY, USA, 236–248. DOI:https://doi.org/10.1145/3308561.3353800

[21] Saba Kawas, George Karalis, Tzu Wen, and Richard E. Ladner. 2016. Improving Real-Time Captioning Experiences for Deaf and Hard of Hearing Students. In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '16). ACM, NY, NY, USA, 15–23. https://doi.org/10.1145/2982142.2982164.

[22] Konstantinidis, B. and Fels, D. (2006). Hand waving apparatus for effective turn-taking (HWAET) using video conferencing for deaf people, in the Proceedings of the 3rd Cambridge Workshop on Universal Access and Assistive Technology (CWUAAT 2006), 10–12 April 2006, Cambridge, U.K., pp. 77–82. Cambridge, U.K.: Cambridge University Press.

[23] Raja Kushalnagar, Walter Lasecki, Jeffrey Bigham. 2014. Accessibility evaluation of classroom captions. ACM Trans. Access. Comput. 5, 3, Article 7 (Jan. 2014), 24 pp. http://dx.doi.org/10.1145/2543578

[24] William Lewis. 2015. Skype Translator: Breaking Down Language and Hearing Barriers. Proceedings of the 37th Conference Translating and the Computer, pages 58–65, London, UK, November 26-27, 2015

[25] Lezzoni LI, O'Day BL, Killeen M, et al. Communicating about Health Care: Observations from Persons Who Are Deaf or Hard of Hearing. Ann Intern Med. 2004;140:356–362. doi: https://doi.org/10.7326/0003-4819-140-5-200403020-00011.

[26] J. Mallory, M. Stinson, L. Elliot, D. Easton. 2017. Personal perspectives on using automatic speech recognition to facilitate communication between deaf students and hearing customers. To appear in Proceedings of the 19th Annual International Conference on Computers and Accessibility (ASSETS '17), ACM, New York, NY.

[27] A. O'Donnell. 2006. The role of peers and group learning. In P. Alexander & P. Winne (Eds.), Handbook of educational psychology. 781-802. Mahwah, NJ: Lawrence Erlbaum Assoc.

[28] Sharon Oviatt, Gina-Anne Levow, Elliott Moreton, and Margaret MacEachern. 1998. Modeling Global and Focal Hyperarticulation during Human–Computer Error Resolution. J. Acoust. Soc. Amer. 104, 3080– 3098. https://doi.org/10.1121/1.423888.

[29] Carol Padden, Tom Humphries, and Carol Padden. 2009. Inside deaf culture. Harvard University Press.

[30] Picheny MA, Durlach NI, Braida LD. Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. J Speech Hear Res. 1985 Mar; 28(1):96-103

[31] Picheny MA, Durlach NI, Braida LD. Speaking clearly for the hard of hearing. II: Acoustic characteristics of clear and conversational speech. J Speech Hear Res. 1986 Dec; 29(4):434-4.

[32] Picheny MA, Durlach NI, Braida LD. Speaking clearly for the hard of hearing. III: An attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech. J Speech Hear Res. 1989 Sep; 32(3):600-3

[33] Anne Marie Piper and James D. Hollan. 2008. Supporting medical conversations between deaf and hearing individuals with tabletop displays. In Proceedings of the 2008 ACM conference on Computer supported cooperative work (CSCW '08). Association for Computing Machinery, New York, NY, USA, 147–156. DOI:https://doi.org/10.1145/1460563.1460587

[34] Matthew Seita, Khaled Albusays, Sushant Kafle, Michael Stinson, and Matt Huenerfauth. 2018. Behavioral Changes in Speakers who are Automatically Captioned in Meetings with Deaf or Hard-of-Hearing Peers. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '18). Association for Computing Machinery, New York, NY, USA, 68–80. DOI:https://doi.org/10.1145/3234695.3236355.

[35] Matthew Seita and Matt Huenerfauth. 2020. Deaf Individuals' Views on Speaking Behaviors of Hearing Peers when Using an Automatic Captioning App. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20). Association for Computing Machinery, New York, NY, USA, 1–8. DOI:https://doi.org/10.1145/3334480.3383083.

[36] Rein Ove Sikveland. 2006. How do We Speak to Foreigners? — Phonetic Analyses of Speech Communication between L1 and L2 Speakers of Norwegian. Working Papers 52, 109–112. Centre for Language and Literature, Lund University, Sweden.

[37] Shen J, Souza PE. On Dynamic Pitch Benefit for Speech Recognition in Speech Masker. Front Psychol. 2018;9:1967. Published 2018 Oct 22. doi:10.3389/fpsyg.2018.01967

[38] Amanda J. Stent, Marie K. Huffman, and Susan E. Brennan. 2008. Adapting Speaking After Evidence of Misrecognition: Local and Global Hyperarticulation. Speech Commun. 50, 3 (March 2008), 163–178. https://doi.org/10.1016/j.specom.2007.07.005.

[39] Allan Sutherland & Tessa Padden (1999) Videoconferencing for Deaf people: a case study of on-line education for Deaf people, Deafness & Education International, 1:2, 114-120, DOI: 10.1179/146431599790561352

[40] Tye-Murray, N., & Clark, W. (1998). Foundations of aural rehabilitation: Children, adults, and their family members. San Diego: Singular Pub. Group.

[41] M. Uther, M.A. Knoll, D. Burnham, Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech, Speech Communication, Volume 49, Issue 1, 2007, Pages 2-7, ISSN 0167-6393, https://doi.org/10.1016/j.specom.2006.10.003.

[42] Vogler, C., Tucker, P., & Williams, N. (2013). Mixed local and remote participation in teleconferences from a deaf and hard of hearing perspective. ASSETS '13. DOI:10.1145/2513383.251703.

[43] Gerard G Walter. 2010. Deaf and Hard-of-Hearing Students in Transition: Demographics with an Emphasis on STEM Education. National Technical Institute for the Deaf, Rochester, NY.

[44] Emily Q. Wang and Anne Marie Piper. 2018. Accessibility in Action: Co-Located Collaboration among Deaf and Hearing Professionals. Proc. ACM Hum.-Comput Interact. 2, CSCW, Article 180 (November 2018), 25 pages. DOI:https://doi.org/10.1145/3274449

[45] Puisan Wong and Kelly Wing Sum Ng. Testing the Hyperarticulation and Prosodic Hypotheses of Child-Directed Speech: Insights From the Perceptual and Acoustic Characteristics of Child-Directed Cantonese Tones. Journal of Speech Language and Hearing Research. Volume 61, Issue 8. August 2018. Pages: 1907-1925. https://doi.org/10.1044/2018_JSLHR-S-17-0375

[46] Zoom. Retrieved May 7, 2020. https://zoom.us/