Understanding Social and Environmental Factors to Enable Collective Access Approaches to the Design of Captioning Technology

Emma J. McDonnell
University of Washington
ejm249@uw.edu

ABSTRACT

Oftentimes human computer interactions (HCI) accessibility research designs technology to support Deaf and disabled people in their existing social contexts. I, instead, propose an approach to accessible technology design that follows the disability justice principle of collective access, envisioning hearing and nondisabled people as key participants in making interactions accessible. Using captioning as a case study, I explore ways that technology could support accessible social norms, achieved by first paying close attention to the social, environmental, and technical factors that shape access for d/Deaf and hard of hearing (DHH) captioning users. My dissertation work will consist of four studies; 1) an exploration of the factors that shape DHH people's current experiences with and future preferences for captioning tools, 2) codesigning features to support accessible group communication with mixed groups of DHH and hearing people, 3) understanding TikTok captioning practices and their impact on DHH users, and 4) exploring the factors that influence professional captioners' work.

CCS CONCEPTS

• **Human-centered computing** → Empirical studies in accessibility.

KEYWORDS

Accessibility, Captioning, d/Deaf and Hard of Hearing, Disability Studies, Collective Access

1 INTRODUCTION

Human Computer Interactions (HCI) accessibility research has grown substantially as a field in recent decades [36], establishing itself as a core aspect of HCI practice [68]. HCI accessibility has also responded to critiques of over-medicalization [39] and has

begun to publish research that engages disability studies and activist thinking (e.g., [5, 6, 18, 50, 63]) as well as first-person accounts of disabled researchers' experiences (e.g., [18, 20, 23, 35, 65]). However, often when building accessible technology, the focus is on how technology could support Deaf and disabled people in their existing social contexts. I contend that many of the problems that HCI accessibility researchers build cutting-edge technology to address become more tractable if paired with efforts to alter social norms. My perspective is informed by current calls from the disability justice movement to prioritize collective access, or viewing access as a project all members of a group have a stake in creating [57]. Further, many in the Deaf community articulate a rejection of hearing world norms (e.g., Deafhood [33], Deaf gain [4]), providing impetus to imagine access outside of hearing and/or nondisabled expectations. My work uses captioning as a case study to explore an approach to accessible technology design that positions access as something that nondisabled/hearing people have a responsibility to, and envisions ways that technology could support the development of accessible social norms. To move toward collective access and group-based accessibility approaches, I argue we need to thoroughly understand the ways that social and environmental factors work alongside technical factors to determine the use and usability of accessible technologies [40]. My dissertation research focuses on the social, environmental, and technical factors that shape d/Deaf and hard of hearing (DHH) people's use of captioning and opportunities for current and future captioning technologies to be designed in ways that engage hearing people as active participants in crafting collective communication access.

2 RELATED WORK

My work is informed and motivated by work spanning HCI, Deaf studies, disability studies, and activist theorizing. Here I lay out relevant work on collective access and captioning technology.

2.1 Collective Access

Recent thinking in academic disability studies and activism has reconceptualized what access might look like and who is responsible for doing the work it takes to make things accessible. The disability justice movement, an activist movement that centers queer, trans, BIPOC disabled people and couples disability politics with broader social issues, has led the way in terms of reframing access as a collective responsibility [69]. Collective access is one of the 10 principles of disability justice, which, in essence, is the idea that *"we can share responsibility for our access needs"* [57] moves away from solutions that primarily support individual independence and toward interdependent approaches that question underlying ableist

norms [43]. Interdependence, which is beginning to be integrated into HCI accessibility work (e.g., [5, 35, 37]), stresses that everyone relies upon others and can be relied upon [44, 45, 62]. This reframing denaturalizes the idea that disabled people are uniquely dependent on others and highlights disabled people's competencies, creating opportunities to view access as for communities, rather than only serving disabled people [25]. I take up collective access as a primary value of my work, and seek to understand what kinds of technology emerge when designing for collective, rather than individual, access.

Deaf studies and Deaf cultural politics also inform my work. The relationship between Deafness and disability is often complicated, with many arguing that Deafness is more akin to membership in a cultural-linguistic minority than a disability [34], though others continue to investigate the motivation to separate Deaf identity from disability [51, 52, 57]. Deaf studies scholars define audism as systemic oppression on the basis of hearing ability, challenging ways that hearing norms serve to oppress Deaf people [3]. Furthermore, the Deaf community has worked to build community identity and values outside of expectations of the hearing world, such as Deafhood [33] and Deaf Gain [4]. This informs my work by driving my approach to technology design that seeks to alter hearing people's behavior and motivating me to be critically reflective around how scientific research can be pervasively audist [15].

2.2 HCI and Captioning

HCI accessibility researchers have explored the design and use of real-time captioning in a variety of manners. Much HCI captioning research utilizes experimental methods to assess novel interventions (e.g., [30, 47, 53]), leaving many opportunities to explore the qualitative and social aspects of captioning use, and I draw inspiration from Kawas et al. [26] and Wang and Piper's work [61], two deep, qualitative explorations of communication access. I briefly overview prior work that explores: key problems in captioning research, captioning use online, social contexts of captioning work, and the role of hearing people in captioned conversations.

A key concern when designing captioning systems is limiting visual dispersion, or the need to attend to multiple visuals at once, and a wide range of approaches have been explored in prior work, including using head mounted displays, integrating captions into classroom lecture configurations, and annotating captions [1, 12, 21, 22, 30, 31, 42, 46–48]. Other design efforts work to communicate non-speech elements of conversation in captioning [16, 17, 59, 70]. An additional focus of recent captioning research is how to convey error rates and uncertainty in automatic captioning [7–10, 48, 49, 53].

Most captioning tools have been developed for in-person conversations, with a burgeoning body of work on online captioning use. Kushalnagar and Vogler published teleconference best practices for communication with DHH people [32], and Seita et al. have explored methods for remote design activities with DHH captioning users [56], but research has not yet explored the particular design and interaction considerations for online captioning applications, a more pressing context since COVID 19 altered work and learning practices. Other accessibility research has explored disabled people's teleconferencing experiences broadly, including DHH participants, finding particular challenges around how audio-driven speaker identification and video-feed prioritization limited DHH people's ability to follow and contribute to a conversation [35, 58].

Captioning is most often studied in classroom or lab settings, with limited focus on small group conversations. Many captioning tools are designed for use during classroom lectures [2, 11, 12, 14, 27–30, 60], which differs significantly from the unstructured communication dynamics present in small-group conversations. Exploratory research has assessed the viability of phone-based automatic speech recognition paired with typed responses for communication between Deaf and hearing participants [13, 38], with promising results. Additionally, head-mounted displays have been assessed for use in small-group conversation, finding that participants valued being able to see captions in the same field of view as their communication partner(s) [22, 47].

Most related to my work is an ongoing effort by Seita et al. [53–56], who have been exploring how hearing people's behavior changes when using automatic captioning, behavioral impact on DHH communication partners, and methods for codesigning captioning tools with DHH/hearing dyads. Seita et al.'s work focuses on specific behaviors and has been able to quantify the impact that different behavioral variations, such as non-standard articulation or rapid speech rate, can have on captioning and DHH caption users. They primarily explore dyadic conversation in experimental contexts. My work is in-conversation with Seita et al.'s work, providing deep qualitative context to their experimental findings.

3 POSITIONALITY

I approach this work as a hearing person and work to center DHH people's perspectives on captioning because, while I believe hearing people have a key role in communication access, any technology ultimately has to be rooted in DHH people's wants and needs. My perspective on access has evolved greatly over the last three years and is not solely academic, as I acquired a disabling chronic illness in 2019.

4 CURRENT AND PROPOSED WORK

My dissertation will consist of four projects, two of which I have completed and two that I propose here. Throughout all these projects I am exploring how social and environmental factors impact people's use of captioning and look towards design guidelines and future technologies that leverage captioning's context of use to make accessible communication a more communal responsibility.

4.1 Social, Environmental, and Technical

I began my work in this space with a combination interview and design probe study focused on d/Deaf and hard of hearing captioning users' experiences of small-group captioned conversations and their preferences for future captioning technologies. This work has been published at CSCW 2021 [40]

Contrary to much of prior research's focus on captioning in relatively structured settings, such as classroom lectures and one-on-one conversations, this research explored the myriad factors that impact communication access during captioned small-group conversations. It was guided by the research questions:

1. What social, environmental, and technical factors impact the use and usefulness of captioning in small groups?

2. What opportunities exist to design captions and caption displays in ways that support more accessible group communication practices?

To answer these questions, we recruited 15 DHH participants with experience using real-time captioning during small group conversations. Each participant completed a 90-minute combination interview and design probe activity. The interview portion of the study focused on how participants use captioning in their daily lives and how the hearing people they communicate with help or hinder effective captioning. Following the interview, we began a design probe activity, focusing on the role of environmental configurations on captioned conversations and potential for additional features that could provide greater conversational context.

From this process, we identified themes in participants' experiences of small-group captioning. A key theme we identify is the way that social dynamics can impact captioning use and performance. Some participants described adaptive practices interlocutors would take on designed to ensure accessible communication while others explained how unwitting or actively inaccessible behaviors exclude them from conversations, even when captioning is present. Another main finding is that captioning is often poorly suited to interactive conversations (as opposed to, for example, one-way lectures), as factors such as lag, overlapping conversation, and a lack of expressive capacity for signers can limit captioning users' ability to join in conversation. Finally, participants reflected on the way that online conversation presented new access barriers (e.g., a lack of spatial information about who is speaking), but also presented new opportunities for conversation access (e.g., strong norms around turn taking and omnipresent text channels for clarification).

We also identify design considerations around captioning displays and future features that could be integrated into captioning tools. Though participants questioned if displaying captions for all in-person conversation participants would hurt or help accessible conversation dynamics, they were broadly enthusiastic about making captioning available to all when meeting online. Participants had varied reactions to adding features to captioning displays - they were widely enthusiastic about speaker identification and overlap alerts for their own use and considered that other people they were interacting with may benefit from feedback as to their speech rate, volume, and caption lag.

Synthesizing these findings, we highlight the need to account for social, environmental, and technical factors on small group captioning and envision a future of captioning tool design for group, rather than solely DHH individuals', use. The conclusions we draw from our empirical findings serve as the basis for the argument that I plan to make in my dissertation; that captioning technology must be understood in its social and environmental context and can be designed to target accessible group behavior.

4.2 Codesigning Online Captioning Tools with Mixed Groups of Hearing and DHH People

The next project in my dissertation work builds on the findings in my first paper to codesign captioning tools with mixed groups of hearing and DHH people who have used captioning when meeting together online. My team and I have completed all study sessions and aim to submit this work to CHI 2023.

Motivated by my prior work, we sought to understand how mixed hearing and DHH groups approach using captioning together and to codesign tools to support more accessible communication. It is guided by the research questions:

1. How do mixed DHH and hearing groups think about, interact with and react to captions during online conversations?
2. What kinds of and designs for real-time feedback during online captioned conversation interest mixed groups of DHH and hearing people?

To answer these questions, we recruited small groups (3-6 people) of participants with experience communicating using captioning online, requiring that at least one person identify as DHH and at least one person identify as hearing. We recruited three groups that fully met these criteria, totaling 13 participants (7 DHH, 6 hearing). To triangulate our data, we also conducted the study with a group (3 DHH, 1 hearing) with less experience communicating together with captions, but with a valuable perspective on how to better support signers.

Each group joined our research team for three consecutive study sessions. In the first session participants began by playing a few rounds of the game Twenty Questions using automatic captioning, while researchers observed groups' communication styles and management of potentially inaccurate or difficult-to-follow captions (e.g., when speakers overlapped or ASR miscaptioned key words). Researchers then conducted a group interview, focused on particular behaviors we observed and participants' thoughts on captioning broadly. This session ended with an introduction to the idea of designing tools for group use and some brainstorming around what kinds of information or features they might be interested in.

The second study session began by revisiting the idea of designing features to use while having a captioned conversation online, and participants reviewed what they had brainstormed in the previous session, and added any new ideas. Then group members sketched out their design ideas for features they were most interested in. Each participant presented their sketch and other group members reacted to their designs and provided feedback. The group then discussed the range of presented ideas and selected their favorites. Some groups (N=2) ended by collectively directing a researcher acting as a "design" scribe to illustrate their top ideas.

After completing the second session, researchers created video prototypes of each group's top three features, illustrating each feature individually and showing all three in context of a short conversation. During the third study session, participants watched each of the videos explaining their feature designs, followed by a chance to comment and react. They then watched the longer video placing those features into the context of conversation and reflected on how they'd feel about using them. Following this, participants then watched and assessed the short feature videos from each of the other three groups participating in the study, focusing on how other groups' ideas may or may not be useful to them.

Our preliminary findings focus on group practices and design considerations for future tools. Groups had different levels of planning around accessibility; some trusted that DHH members would raise any access concerns while others established explicit conversation access approaches. Participants stressed the role of relationships in effective communication – when communicating with

people unfamiliar with DHH conversation norms or in professional contexts, they reported having fewer options to clarify or adapt behavioral expectations. When identifying what made captioning difficult to use, participants consistently highlighted overlapping speech, inaccurate captions, a lack of speaker identification in most automatic captioning tools, and more. Participants approached the design of future tools with many shared values; minimizing visual and cognitive overload, the necessity of user control and customizability, and using established, learnable interaction paradigms when building new features. Several groups put forward the idea that while technology cannot force people to change their behavior, it could provide information and structure to guide people towards more captioning-friendly norms.

This work contributes an empirical account of mixed hearing ability groups' use of captioning, a set of collaboratively designed features and priorities for future captioning tools to be used by DHH and hearing individuals alike, and design considerations for captioning in online environments. I planned this study to explore the core arguments of my dissertation work; exploring what groups would be interested in using to support accessible communication and how technology can leverage the specifics of the environment (videoconferencing) and intervene in social factors that shape captioned conversations.

4.3   TikTok Captioning

For my next study, I will lead research into captioning practices on the social video platform TikTok. While my prior work explored real-time captioned conversations and possible interventions, this work will examine social and environmental factors that shape pre-generated captions and how collective access approaches could shape how TikTok creators caption their content. This work is still in planning stages but will begin in earnest later this summer. A likely target timeline and venue for this work is submission as a full paper to ASSETS 2023.

TikTok is a social media platform where users submit up to 10-minute videos (though platform norms encourage shorter-form content). Prior to April 2021, it did not support integrated in-platform captioning, though now allows users to autocaption (and, if desired, correct captions) their videos [66]. Other approaches to captioning on the platform include adding text overlays to the video, manually timed to serve as captions and burned into the video. TikTok creators have developed stylized approaches to captioning (e.g., adding emoji, placing captions in a way that conveys speaker identity). They also integrate word substitutions and altered spellings, used across the platform to avoid algorithmic consequences, into captions (e.g., captioning the word "lesbian" as "le$bian" or "le dollar bean") [71]. Many TikToks are uncaptioned, and TikToks often contain both viral "sounds" and user-generated audio, and often users don't caption all audio streams.

To my knowledge, to date, captioning practices and prevalence on TikTok have not been comprehensively described nor investigated in terms of their impact on DHH platform users. With over one billion monthly users [67], TikTok is a key public space that hosts content that is widely inaccessible to DHH users. The first step to being able to address this inaccessibility is to define and

quantify TikTok captioning practices and determine their impact on DHH users.

TikTok captioning practices merit targeted study, due to myriad platform-specific considerations. TikTok's unique level of user control has created emergent norms around the appearance, placement, styling, and content of captions, a fascinating context in which to explore captioning design. Because generating captions is currently up to TikTok creators, not governed by the platform or third parties, it also poses an interesting collective access problem – any migration toward more accessible norms must include action on the part of majority-hearing creators. Further, TikTok as a platform disincentives accurate captions due to their aggressive censorship and demonitization practices, which often target minoritized creators (e.g., queer people [72], Black people [73]) or people discussing sensitive topics (e.g., mental health crises [64], sexual health education [19]). Therefore understanding and providing guidance for future TikTok captioning practices must be firmly located in the context of the environment that is the TikTok platform and social factors that shape which conversations can be fully captioned without consequence.

I propose three research questions to guide this work:

1. What trends and norms are present in TikTok captions?
2. How do varied approaches to captioning impact DHH TikTok users?
3. How do TikTok creators think about captioning their videos?

I, along with a cross-institution team that I will lead, plan to run a three phase study to investigate the answers to these questions.

To begin, we plan to systematically analyze a corpus of TikTok videos to quantify the prevalence of captioning, types and frequency of caption substitutions and edits, variety of stylistic elements added to captions, how non-standard speech (e.g., toddler speak) and non-speech (e.g., dog barking) are captioned, and how much of the audio in the video is captioned (e.g., only user-added content or user content and viral "sounds"). I believe that quantitative analysis is the best approach to gain this information, as systematic coding of a carefully selected sample will allow us to empirically describe trends, data which has not yet been published around TikTok captioning. Further, having determined which practices occur frequently across TikTok will provide focus for the next stages of this project.

After gathering this empirical data, we plan to survey and interview DHH TikTok users. We will circulate a survey around common TikTok captioning practices, asking DHH captioning users to rate them on a series of relevant metrics (e.g., impact on comprehension, favorability). Interested survey-takers will be invited to participate in an interview, which will focus on participants' experiences using TikTok, what makes a video particularly accessible or inaccessible to them, their preferred captioning approaches, their perspectives on algorithmic censorship and captioning, and other relevant topics.

In parallel with our engagement with DHH TikTok users, we will also send out a survey to TikTok creators who have experience captioning TikToks. This survey will focus on their captioning process, their feelings about captioning TikToks, and inquire as to which captioning approaches identified in our quantitative analysis they have used. Depending on the state of the research at this point, we may also interview TikTok creators.

I envision the contributions of this work to be a quantification of captioning practices on TikTok, empirical data around how various choices impact DHH TikTok users, and recommendations for practice around TikTok captioning (and, more broadly, user-generated captioning) informed by DHH users' perspectives. Though this study differs from my prior focus on how small group captioning tools could be designed for real-time use, it is an exciting additional context in which to explore my dissertation argument. Understanding existing TikTok captioning practices, and the social and environmental factors that shape TikTok creators' choices and DHH users' preferences will provide the opportunity to explore what collective responsibility for access can look like in this context.

4.4 Exploring the Practices of Professional Captioners

After completing the TikTok study, the next aspect of captioning I plan to explore is professional captioners' experiences and perspectives on the factors that shape their work. I will also compare professional captioners' work to state-of-the-art automatic captioning engines and survey DHH captioning users around the impact of different captioning approaches. This complements my first two projects by exploring how social behavior impacts real-time captioning and builds on my TikTok work focused on non-professional captioning practices to also consider the factors that shape professional captioning. This work could be feasibly submitted to CHI, CSCW, or ASSETS during the 2023-2024 academic year, depending on other factors such as job search, teaching responsibilities, or internships.

The current preferred and highest-quality real-time captioning is human-generated, often referred to as Communication Access Realtime Translation, or CART. CART captioners acquire professional certification, which benchmarks competency at the ability to caption speech at 180 words per minute with 96% accuracy [74]. As the people responsible for generating captions, they have a valuable and under-considered perspective around the factors that make a conversation easy or difficult to caption. Additionally, while qualitative analysis methods often stress that a text transcript is a subjective and incomplete record of a conversation [41], within captioning research there is often an assumed ground-truth transcript of a conversation that human and machine captioning is assessed against [24]. Drawing from my experience working with captioners, I hypothesize that there are myriad decisions and social inventions that human captioners make which shape the text of what is actually captioned. Further, the set of decisions and considerations that professional captioners make could be an immensely valuable guide for the development of better automatic captioning. I propose a study that explores professional captioners' perspectives on and approaches to captioning, assesses how those approaches compare to current state-of-the-art automatic captioning, and identifies how different approaches impact DHH captioning users' experiences viewing content.

This work will be guided by the following research questions:

1. How do social, environmental, and technical factors impact CART writers' experience of providing real-time captioning?

2. What, if any, differences exist between individual CART writers' captions and between human-generated and automatic captions?

3. If present, what impact does variance in captioning style have on DHH captioning users' comprehension of and satisfaction with captioned content?

To answer these questions I plan to run a three phase study: interviews and captioning activities with CART writers, a comparative analysis between CART writers' and ASR engines' captions, and, if relevant, a survey of DHH captioning users' perspectives on differences between varied captioning approaches.

First, I propose running study sessions with professional CART writers, consisting of an interview and captioning activity. The semi-structured interviews will focus on CART writers' experiences working as captioners, the social dynamics that impact their work, what makes for an easy or hard to caption conversation, the role technology plays in their work, and ways they might intervene in a conversation to be able to provide better captions. Following the interview, I will request that the CART writers caption several (3-5) short clips selected to include ambiguous or difficult-to-caption scenarios (e.g., overlapping speakers, rapid speech). After each clip, I will ask CART writers for reflections or reactions to the exercise.

After completing study sessions with all recruited CART writers, I propose using the state-of-the-art, popular automatic speech recognition engines to caption the same set of clips used during study sessions. Then I plan to compare all sets of captions to assess similarities and differences. Likely metrics for this analysis include Word Error Rate (WER), Automated Caption Evaluation (ACE) [24], caption lag, and presence and accuracy of speaker identification.

The potential next stage of this work is dependent on the findings in comparative analysis. If there is no significant difference between individual CART writers, ASR engines, or human vs machine approaches to captioning, that would be an interesting finding in and of itself, and likely would not motivate a follow up survey. I, however, hypothesize that there will be relevant differences that would be interesting to explore in terms of their impact on DHH captioning users. In that case, I plan to identify representative examples of the captioning differences we found to be most salient. This would serve as the basis for a survey wherein DHH participants would view paired approaches to the same video clip and be asked to rate each approach in terms of comprehension and favorability, and to select which they preferred.

From this work, I envision contributing an empirical account of the social, environmental, and technical factors that shape CART writers' work, a quantitative analysis of how different human and machine captioning approaches result in similarities and differences in the captions themselves, empirical data around DHH people's preferred approaches, and design considerations for future captioning practices. I see this study as in line with the broader narrative of my dissertation due to its exploration of the way that social and environmental factors impact a service that is often considered to be objective and its surfacing of how people's behavior shapes access. I'm still pondering how to make this study clearly focus on collective access approaches, since CART writers are operating as service providers rather than conversation participants. I suspect their in depth knowledge of what shapes their ability to provide

captions is a crucial insight to guide collective access recommendations, but would appreciate discussing how to emphasize collective access in this study during the Doctoral Consortium.

5 CONCLUSION

In conclusion, my work explores how attention to the social and environmental aspects of captioning use can drive the development of technology that supports collective access approaches to accessible communication. I propose doing so via four projects, two of which I've done and two that I will do in the fourth and fifth years of my PhD. My work will contribute to the broader HCI accessibility community in the following ways:

- Demonstrating the need to consider social, environmental, and technical factors in concert when seeking to understand captioning's use and usefulness, explored with DHH people, hearing people, and professional captioners
- Providing concrete design recommendations and priorities for captioning tools to be used by groups having conversations, rather than solely for DHH captioning users
- Design considerations for online captioning tools
- An empirical account of captioning practices on TikTok, and, more broadly, current user-generated captioning approaches
- Recommendations for best practices around user-generated captioning, informed by DHH captioning users' perspectives

ACKNOWLEDGMENTS

REFERENCES

[1] Akhter Al Amin, Saad Hassan, Sooyeon Lee, and Matt Huenerfauth. 2022. Watch It, Don't Imagine It: Creating a Better Caption-Occlusion Metric by Collecting More Ecologically Valid Judgments from DHH Viewers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (CHI '22), 1–14. https://doi.org/10.1145/3491102.3517681

[2] Keith Bain, Sara H. Basson, and Mike Wald. 2002. Speech recognition in university classrooms: liberated learning project. In *Proceedings of the fifth international ACM conference on Assistive technologies* (Assets '02), 192–196. https://doi.org/10.1145/638249.638284

[3] H.-Dirksen L. Bauman. 2004. Audism: Exploring the Metaphysics of Oppression. *The Journal of Deaf Studies and Deaf Education* 9, 2: 239–246. https://doi.org/10.1093/deafed/enh025

[4] H.-Dirksen L. Bauman and Joseph J. Murray. 2014. *Deaf Gain: Raising the Stakes for Human Diversity*. U of Minnesota Press.

[5] Cynthia L. Bennett, Erin Brady, and Stacy M. Branham. 2018. Interdependence as a Frame for Assistive Technology Research and Design. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS '18*, 161–173. https://doi.org/10.1145/3234695.3236348

[6] Cynthia L. Bennett, Daniela K. Rosner, and Alex S. Taylor. 2020. The Care Work of Access. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI '20), 1–15. https://doi.org/10.1145/3313831.3376568

[7] Larwan Berke. 2017. Displaying confidence from imperfect automatic speech recognition for captioning. *ACM SIGACCESS Accessibility and Computing*, 117: 14–18. https://doi.org/10.1145/3051519.3051522

[8] Larwan Berke, Khaled Albusays, Matthew Seita, and Matt Huenerfauth. 2019. Preferred Appearance of Captions Generated by Automatic Speech Recognition for Deaf and Hard-of-Hearing Viewers. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3290607.3312921

[9] 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*, 155–164. https://doi.org/10.1145/3132525.3132541

[10] Larwan Berke, Sushant Kafle, and Matt Huenerfauth. 2018. Methods for Evaluation of Imperfect Captioning Tools by Deaf or Hard-of-Hearing Users at Different Reading Literacy Levels. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–12. https://doi.org/10.1145/3173574.3173665

[11] Janine Butler, Brian Trager, and Byron Behm. 2019. Exploration of Automatic Speech Recognition for Deaf and Hard of Hearing Students in Higher Education Classes. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '19), 32–42. https://doi.org/10.1145/3308561.3353772

[12] Anna C. Cavender, Jeffrey P. Bigham, and Richard E. Ladner. 2009. ClassInFocus: enabling improved visual attention strategies for deaf and hard of hearing students. In *Proceeding of the eleventh international ACM SIGACCESS conference on Computers and accessibility - ASSETS '09*, 67. https://doi.org/10.1145/1639642.1639656

[13] 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*, 405–406. https://doi.org/10.1145/3132525.3134798

[14] Lisa Elliot, Michael Stinson, Donna Easton, and Jennifer Bourgeois. 2008. College students' learning with C-Print's educational software and automatic speech recognition. In *American Educational Research Association Annual Meeting*.

[15] Lance Forshay, Kristi Winter, and Emily M Bender. 2016. Sign Aloud Open Letter. Retrieved from http://faculty.washington.edu/ebender/papers/SignAloudOpenLetter.pdf

[16] Benjamin Gorman, Michael Crabb, and Michael Armstrong. 2021. *Adaptive Subtitles: Preferences and Trade-Offs in Real-Time Media Adaption*. https://doi.org/10.1145/3411764.3445509

[17] Michael Gower, Brent Shiver, Charu Pandhi, and Shari Trewin. 2018. Leveraging Pauses to Improve Video Captions. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '18), 414–416. https://doi.org/10.1145/3234695.3241023

[18] Megan Hofmann, Devva Kasnitz, Jennifer Mankoff, and Cynthia L Bennett. 2020. Living Disability Theory: Reflections on Access, Research, and Design. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '20), 1–13. https://doi.org/10.1145/3373625.3416996

[19] Anna Iovine. 2021. Why is TikTok removing sex ed videos? *Mashable*. Retrieved July 11, 2022 from https://mashable.com/article/tiktok-sex-education-content-removal

[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), 236–248. https://doi.org/10.1145/3308561.3353800

[21] 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. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15), 241–250. https://doi.org/10.1145/2702123.2702393

[22] 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. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, 81–92. https://doi.org/10.1145/3234695.3236362

[23] Dhruv Jain, Venkatesh Potluri, and Ather Sharif. 2020. Navigating Graduate School with a Disability. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '20), 1–11. https://doi.org/10.1145/3373625.3416986

[24] Sushant Kafle and Matt Huenerfauth. 2017. Evaluating the Usability of Automatically Generated Captions for People who are Deaf or Hard of Hearing. https://doi.org/10.1145/3132525.3132542

[25] Devva Kasnitz. 2020. The Politics of Disability Performativity: An Autoethnography. *Current Anthropology* 61, S21: S16–S25. https://doi.org/10.1086/705782

[26] 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*, 15–23. https://doi.org/10.1145/2982142.2982164

[27] Richard Kheir and Thomas Way. 2007. Inclusion of deaf students in computer science classes using real-time speech transcription. 261–265. https://doi.org/10.1145/1269900.1268860

[28] R S Kushalnagar, W S Lasecki, and J P Bigham. 2014. Accessibility Evaluation of Classroom Captions Accessibility Evaluation of Classroom Captions Accessibility evaluation of classroom captions. *ACM Trans. Access. Comput* 5, 7: 23. https://doi.org/10.1145/2543578

[29] Raja Kushalnagar and Poorna Kushalnagar. 2014. Collaborative Gaze Cues and Replay for Deaf and Hard of Hearing Students. In *Computers Helping People with Special Needs* (Lecture Notes in Computer Science), 415–422. https://doi.org/10.1007/978-3-319-08599-9_63

[30] Raja S. Kushalnagar, Gary W. Behm, Aaron W. Kelstone, and Shareef Ali. 2015. Tracked Speech-To-Text Display: Enhancing Accessibility and Readability of Real-Time Speech-To-Text. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (ASSETS '15), 223–230. https://doi.org/10.1145/2700648.2809843

[31] Raja S. Kushalnagar, Anna C. Cavender, and Jehan-François Pâris. 2010. Multiple view perspectives: improving inclusiveness and video compression in mainstream classroom recordings. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '10*, 123. https://doi.org/10.1145/1878803.1878827

[32] Raja S. Kushalnagar and Christian Vogler. 2020. Teleconference Accessibility and Guidelines for Deaf and Hard of Hearing Users. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '20), 1–6. https://doi.org/10.1145/3373625.3417299

[33] Paddy Ladd. 2005. Deafhood: A concept stressing possibilities, not deficits. *Scandinavian Journal of Public Health* 33, 66_suppl: 12–17. https://doi.org/10.1080/14034950510033318

[34] Harlan Lane. 2005. Ethnicity, Ethics, and the Deaf-World. *The Journal of Deaf Studies and Deaf Education* 10, 3: 291–310. https://doi.org/10.1093/deafed/eni030

[35] Kelly Mack, Maitraye Das, Dhruv Jain, Danielle Bragg, John Tang, Andrew Begel, Erin Beneteau, Josh Urban Davis, Abraham Glasser, Joon Sung Park, and Venkatesh Potluri. 2021. Mixed Abilities and Varied Experiences: a group autoethnography of a virtual summer internship. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '21), 1–13. https://doi.org/10.1145/3441852.3471199

[36] Kelly Mack, Emma McDonnell, Dhruv Jain, Lucy Lu Wang, Jon E. Froehlich, and Leah Findlater. 2021. What Do We Mean by "Accessibility Research"?: A Literature Survey of Accessibility Papers in CHI and ASSETS from 1994 to 2019. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. https://dl.acm.org/doi/10.1145/3411764.3445412

[37] Kelly Mack, Emma McDonnell, Venkatesh Potluri, Maggie Xu, Jailyn Zabala, Jeffrey Bigham, Jennifer Mankoff, and Cynthia Bennett. 2022. Anticipate and Adjust: Cultivating Access in Human-Centered Methods. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (CHI '22), 1–18. https://doi.org/10.1145/3491102.3501882

[38] James R. Mallory, Michael Stinson, Lisa Elliot, and Donna Easton. 2017. Personal Perspectives on Using Automatic Speech Recognition to Facilitate Communication between Deaf Students and Hearing Customers. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, 419–421. https://doi.org/10.1145/3132525.3134779

[39] Jennifer Mankoff, Gillian R. Hayes, and Devva Kasnitz. 2010. Disability studies as a source of critical inquiry for the field of assistive technology. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility* (ASSETS '10), 3–10. https://doi.org/10.1145/1878803.1878807

[40] Emma J. McDonnell, Ping Liu, Steven M. Goodman, Raja Kushalnagar, Jon E. Froehlich, and Leah Findlater. 2021. Social, Environmental, and Technical: Factors at Play in the Current Use and Future Design of Small-Group Captioning. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2: 434:1-434:25. https://doi.org/10.1145/3479578

[41] Caitlin McMullin. 2021. Transcription and Qualitative Methods: Implications for Third Sector Research. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*. https://doi.org/10.1007/s11266-021-00400-3

[42] Dorian Miller, Karl Gyllstrom, David Stotts, and James Culp. 2007. Semi-transparent video interfaces to assist deaf persons in meetings. In *Proceedings of the 45th annual southeast regional conference* (ACM-SE 45), 501–506. https://doi.org/10.1145/1233341.1233431

[43] Mia Mingus. 2011. Changing the Framework: Disability Justice. *Leaving Evidence*. Retrieved March 20, 2020 from https://leavingevidence.wordpress.com/2011/02/12/changing-the-framework-disability-justice/

[44] Mia Mingus. 2017. Access Intimacy, Interdependence and Disability Justice. *Leaving Evidence*. Retrieved February 5, 2020 from https://leavingevidence.wordpress.com/2017/04/12/access-intimacy-interdependence-and-disability-justice/

[45] Michael Oliver. 1986. Disability and Dependency. A Creation of Industrial Societies? In *Disability and Dependency*, 7–22.

[46] Alex Olwal, Kevin Balke, Dmitrii Votintcev, Thad Starner, Paula Conn, Bonnie Chinh, and Benoit Corda. 2020. Wearable Subtitles: Augmenting Spoken Communication with Lightweight Eyewear for All-day Captioning. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. Association for Computing Machinery, New York, NY, USA, 1108–1120. Retrieved July 11, 2022 from http://doi.org/10.1145/3379337.3415817

[47] 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.

[48] Agnès Piquard-Kipffer, Odile Mella, Jérémy Miranda, Denis Jouvet, and Luiza Orosanu. 2015. Qualitative investigation of the display of speech recognition results for communication with deaf people. In *Proceedings of SLPAT 2015: 6th Workshop on Speech and Language Processing for Assistive Technologies*, 36–41. https://doi.org/10.18653/v1/W15-5107

[49] Kevin Rathbun, Larwan Berke, Christopher Caulfield, Michael Stinson, and Matt Huenerfauth. 2017. Eye Movements of Deaf and Hard of Hearing Viewers of Automatic Captions. *California State University, Northridge*. Retrieved August 3, 2020 from http://dspace.calstate.edu/handle/10211.3/190208

[50] Kathryn E. Ringland, Jennifer Nicholas, Rachel Kornfield, Emily G. Lattie, David C. Mohr, and Madhu Reddy. 2019. Understanding Mental Ill-health as Psychosocial Disability: Implications for Assistive Technology. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '19), 156–170. https://doi.org/10.1145/3308561.3353785

[51] Octavian Robinson. 2010. We Are of a Different Class: Ableist Rhetoric in Deaf America, 1880-1920.

[52] Octavian E. Robinson and Jonathan Henner. 2017. The personal is political in *The Deaf Mute Howls*: deaf epistemology seeks disability justice. *Disability & Society* 32, 9: 1416–1436. https://doi.org/10.1080/09687599.2017.1313723

[53] 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), 68–80. https://doi.org/10.1145/3234695.3236355

[54] 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*, 1–12. https://doi.org/10.1145/3430263.3452430

[55] 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), 1–8. https://doi.org/10.1145/3334480.3383083

[56] Matthew Seita, Sooyeon Lee, Sarah Andrew, Kristen Shinohara, and Matt Huenerfauth. 2022. Remotely Co-Designing Features for Communication Applications using Automatic Captioning with Deaf and Hearing Pairs. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (CHI '22), 1–13. https://doi.org/10.1145/3491102.3501843

[57] Sins Invalid. 2019. Skin Tooth and Bone: The Basis of Movement is Our People, a Disability Justice Primer. Sins Invalid.

[58] John Tang. 2021. Understanding the Telework Experience of People with Disabilities. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1: 30:1-30:27. https://doi.org/10.1145/3449104

[59] Máté Ákos Tündik, György Szaszák, G. Gosztolya, and A. Beke. 2018. User-centric Evaluation of Automatic Punctuation in ASR Closed Captioning. In *INTERSPEECH*. https://doi.org/10.21437/Interspeech.2018-1352

[60] M. Wald. 2005. Using Automatic Speech Recognition to Enhance Education for All Students: Turning a Vision into Reality. In *Proceedings Frontiers in Education 35th Annual Conference*, S3G-22-S3G-25. https://doi.org/10.1109/FIE.2005.1612286

[61] Emily Q. Wang and Anne Marie Piper. 2018. Accessibility in Action: Co-Located Collaboration among Deaf and Hearing Professionals. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW: 180:1-180:25. https://doi.org/10.1145/3274449

[62] Glen W. White, Jamie Lloyd Simpson, Chiaki Gonda, Craig Ravesloot, and Zach Coble. 2010. Moving from Independence to Interdependence: A Conceptual Model for Better Understanding Community Participation of Centers for Independent Living Consumers. *Journal of Disability Policy Studies* 20, 4: 233–240. https://doi.org/10.1177/1044207309350561

[63] Rua M. Williams, Kathryn Ringland, Amelia Gibson, Mahender Mandala, Arne Maibaum, and Tiago Guerreiro. 2021. Articulations toward a crip HCI. *Interactions* 28, 3: 28–37. https://doi.org/10.1145/3458453

[64] WIRED. 2022. Are TikTok algorithms changing how people talk about suicide? *Ars Technica*. Retrieved July 11, 2022 from https://arstechnica.com/gaming/2022/05/are-tiktok-algorithms-changing-how-people-talk-about-suicide/

[65] Anon Ymous, Katta Spiel, Os Keyes, Rua M. Williams, Judith Good, Eva Hornecker, and Cynthia L. Bennett. 2020. "I am just terrified of my future" — Epistemic Violence in Disability Related Technology Research. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI EA '20), 1–16. https://doi.org/10.1145/3334480.3381828

[66] 2019. Introducing auto captions. *Newsroom | TikTok*. Retrieved July 11, 2022 from https://newsroom.tiktok.com/en-us/introducing-auto-captions

[67] 2022. TikTok Statistics - Everything You Need to Know [Apr 2022 Update]. *Wallaroo Media*. Retrieved July 11, 2022 from https://wallaroomedia.com/blog/social-media/tiktok-statistics/

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

[68] Weaving CHI - Top Keyword Topics. *Tableau Software.* Retrieved May 13, 2020 from https://public.tableau.com/views/WeavingCHI-TopKeywordTopics/TopKeywordTopics?:embed=y&:display_count=yes&publish=yes:showVizHome=no

[69] What is Disability Justice? *Sins Invalid.* Retrieved July 11, 2022 from https://www.sinsinvalid.org/news-1/2020/6/16/what-is-disability-justice

[70] Ava - All-in-One Click Captions for All Conversations. Retrieved April 14, 2021 from https://ava.me/

[71] Internet 'algospeak' is changing our language in real time, from 'nip nops' to 'le dollar bean.' *Washington Post.* Retrieved July 11, 2022 from https://www.washingtonpost.com/technology/2022/04/08/algospeak-tiktok-le-dollar-bean/

[72] TikTok Apologizes After Reportedly Censoring LGBTQ+ Users | Them. Retrieved July 11, 2022 from https://www.them.us/story/lgbtq-users-reportedly-being-censored-by-tiktok

[73] Months after TikTok apologized to Black creators, many say little has changed. *NBC News.* Retrieved July 11, 2022 from https://www.nbcnews.com/pop-culture/pop-culture-news/months-after-tiktok-apologized-black-creators-many-say-little-has-n1256726

[74] Certified Realtime Captioner (CRC) | NCRA. Retrieved July 11, 2022 from https://www.ncra.org/certification/NCRA-Certifications/certified-realtime-captioner