Audio Engineering for Podcasts by deaf and Hard of Hearing Creators

Keita Ohshiro
New Jersey Institute of Technology
Newark, New Jersey, USA
ko89@njit.edu

Mark Cartwright
New Jersey Institute of Technology
Newark, New Jersey, USA
mark.cartwright@njit.edu

Abstract

Podcasting has evolved as a significant medium where independent creators engage in various production stages, including audio engineering. However, research is limited on how deaf and hard of hearing (dHH) people create podcasts. To address this gap, we conducted interviews with four self-taught dHH podcasters engaged in audio engineering. Our findings reveal how they navigate hearing-related challenges through visual editing techniques, self-taught skill development, and reliance on peer support, while advocating for accessibility. For future directions, we discuss differences between podcast and music audio engineering and recommend integrating AI-powered tools and developing accessible learning resources and environments to support dHH creators.

CCS Concepts

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

Keywords

accessibility, deaf, hard of hearing, audio engineering, podcast

ACM Reference Format:
Keita Ohshiro and Mark Cartwright. 2025. Audio Engineering for Podcasts by deaf and Hard of Hearing Creators. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3706599.3719764

1 Introduction

Since its inception in 2004, podcasting has become a significant medium [69]. It contributes to diverse domains, including delivering news and entertainment [59], serving as a tool for education [27, 28], as well as supporting advocacy and community empowerment [16, 25, 73]. The rise of podcasting has been accompanied by a surge in amateur and independent podcasters [7], who often engage in multiple production stages — pre-production (e.g., creating stories), production (e.g., audio engineering), and post-production (e.g., distribution) [58, 66]. Technological advances, particularly in Digital Audio Workstations (DAWs), are often said to have democratized audio engineering [3, 31, 43].

CHI EA '25, April 26-May 1, 2025, Yokohama, Japan

ACM ISBN 979-8-4007-1395-8/2025/04
https://doi.org/10.1145/3706599.3719764

However, research on accessibility of audio engineering for people with disabilities — particularly for deaf, Deaf, and hard of hearing (DHH) people[1] [12, 40] — remains limited. With the growing DHH population [52] and increasing attention to accessibility in HCI [45] and audio technology [26], we have begun to see focused efforts to explore DHH individuals' experiences in audio engineering, primarily for music [50, 51] where the focus lies on sound elements. Yet, how they engage in audio engineering for podcast creation, which emphasizes storytelling, remains an underexplored area. Building on prior research on deaf and hard of hearing (dHH) audio engineers in music [51], this study aims to address the gap in understanding accessibility in podcast audio engineering through interviews with four self-taught dHH podcasters. This study contributes to a deeper understanding of accessible audio engineering by dHH individuals and highlights pathways to empower their representation in podcasting and broader creative sound production. To our knowledge, this is the first interview study focusing on dHH audio engineers in podcast production.

2 Related Work

2.1 Audio engineering and accessibility

Research has explored audio engineering and its associated technologies and tools. Regarded as a creative art form rather than a procedural task [48], audio engineering has evolved its workflow in digital production with the development of DAWs [5, 72]. Research has explored various interactions, including visual [19, 20], gesture [55], language-based [11], and multi-modal interfaces [49] as well as AI-powered tools for various tasks, such as mixing and mastering [17, 42, 68].

While this body of work is extensive, accessibility remains an emerging area of study. For blind and low-vision audio engineers, studies have examined their practices [54, 61, 63] and designed tools to enhance accessibility [36, 47, 53, 62, 71]. In contrast, research on DHH individuals in audio engineering is still in its early stages. From a pedagogical perspective, Cheng and McGregor proposed strategies to support DHH students in learning music and audio production [15]. Furthermore, studies involving participants have begun to explore this area, including a survey of 50 DHH individuals engaged in broader creative audio activities such as music playing and audio engineering [50], as well as interviews with eight dHH music audio engineers with formal training and education [51]. These studies revealed that dHH music audio engineers use technological (e.g., visualizations and manipulations of sound) and social workarounds (e.g., peer feedback) to address the lack of confidence

[1] People who identify as Deaf with a capital D share a cultural connection to the Deaf community, while those identifying as deaf with a lowercase d or hard of hearing refer to their audiological hearing levels.

Table 1: Participant Demographics and DAW Usage

ID	Identity	Self-Described Hearing	Hearing Device	DAW
P1	Hard of Hearing	65% hearing in left, worse in right	None	Audacity, Adobe Audition
P2	deaf	Moderate hearing loss	Hearing Amplifier	Audacity
P3	deaf	Moderate hearing loss	Hearing Aid	Audacity
P4	deaf	About 25 to 30% of normal hearing	Hearing Aid	Audacity

and the extra 'hearing work' required due to their hearing while leveraging their existing skills.

Building on this foundation, our study examines how self-taught dHH audio engineers engage in podcasting.

2.2 Podcast accessibility and creation

Research on podcast accessibility has focused primarily on enabling DHH audiences to engage with auditory content [13, 21, 30, 39]. Transcripts and captions are essential for providing access to both speech and non-speech audio elements in podcasts [14, 22, 29], as in videos [6, 41, 76]. Resources to help podcasters make their content accessible to DHH audiences are available [1, 2, 4], and podcast platforms have taken steps to improve accessibility with transcripts [32, 65].

While podcast creation by individuals without disabilities has been supported in non-academic contexts through various resources [35, 67], academic research remains relatively new. For instance, Rime et al. examined the iterative and collaborative nature of podcast production workflows [58] and explored the potential integration of emerging technologies [57]. In contrast, podcast creation by DHH individuals remains largely overlooked. Despite the presence of a few notable DHH podcasters [10, 37] and representations in podcast episodes [56, 64, 74], academic studies specifically examining podcast creation by DHH individuals in the context of audio engineering are scarce.

Our study addresses this gap, focusing on how self-taught dHH audio engineers engage in podcast creation.

3 Methods

3.1 Procedure

We conducted remote semi-structured interviews with four dHH audio engineers engaged in podcasting (Table 1). Each session lasted about an hour and was conducted in spoken English via Zoom or Google Meet, with automated captions enabled for P2 and P3. We recorded the interviews and provided a $20 Amazon gift card as compensation. The study received IRB approval from the New Jersey Institute of Technology.

The interview included questions on participants' demographics, hearing-related topics, and audio engineering experiences such as practices, challenges, workarounds, and learning strategies. Participants also shared their screens and guided us through the processes of their past work[2].

We analyzed the data with Braun and Clarke's Reflexive Thematic Analysis [8], emphasizing its flexibility and focus on researchers' subjectivity and reflexivity. After transcribing the video recordings, we conducted coding, generated themes, and iteratively refined them.

[2]Screenshots from P3's workflow were unavailable due to technical issues.

3.2 Participants

Participants lived in the United States and were fluent in English. They included two non-binary individuals, one transgender male, and one female. Their ages ranged two between 25-34, one between 35-44, and one between 45-54. To preserve anonymity, we use they/them pronouns and avoid linking age and gender details.

P1 and P2 became dHH later in life, while P3 and P4 were uncertain about the timing due to the time gap in initial awareness, the doctor's diagnosis, and its progression over time. Only P1 started podcasting and audio engineering during their transition to being dHH and the others began after becoming dHH. All participants were self-taught, starting between 2017 and 2019, and consistently published episodes or contributed to others' podcasts. Beyond audio engineering, they took on various roles such as writing, directing, voice acting, distributing, and marketing.

4 Findings

We present our findings on the current practices of dHH individuals in podcast audio engineering: using visual editing (Section 4.1); developing audio engineering skills (Section 4.2); relying on peer support (Section 4.3); and advocating for accessibility (Section 4.4). To provide context, we first outline the processes and tasks in podcast audio engineering:

- `Recording`: Capturing sounds, primarily speech (e.g., interviews or voice actors' dialogues). In some cases, voice actors record independently and send tracks to the audio engineer.
- `Editing`: Managing dialogue and speech through tasks such as track selection, track alignment (ensuring natural timing and flow), and noise reduction. This also includes tasks common in the mixing process in music audio engineering, such as adjusting sound levels and applying equalization (EQ) and effects.
- `Sound Design`: Adding sounds not present in the recorded tracks, such as background music and sound effects (e.g., dishes clinking). Some tasks, like track alignment and sound level adjustment, overlap with editing.
- `Rendering`: Exporting finalized tracks, typically without the mastering process done in music audio engineering.

4.1 Visual editing: strengths and challenges

4.1.1 Efficiency and hearing conservation through visual editing. Participants employed various visuals in their workflows, including

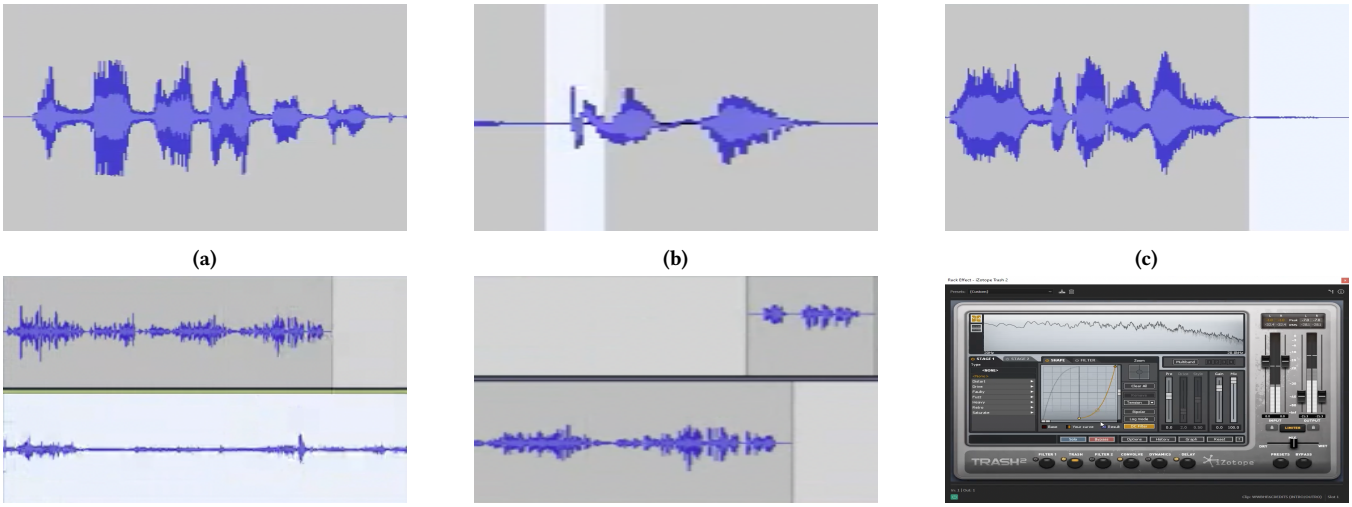

Figure 1: Screenshots from the interviews illustrating visual editing with waveform: (a) Fishbone shape indicating laughter. (b) Sudden spike, highlighted in light blue, indicating a plosive consonant sound. (c) Small up-and-down pattern, highlighted in light blue, indicating background noise. (d) P2 adjusting the sound levels, making the background sound (bottom) smaller than the foreground speech (top). (e) P2 adjusting the position of the waveform above without listening, saying, "*[The waveform above] can overlap here because [it] definitely said 'um'.*" (f) P1 evaluating the intensity of an effect with visuals.

waveforms, numeric adjustments, and textual references. Waveforms were central, providing information about sound often before listening. They used waveform shapes to identify speech patterns such as laughter (Figure 1a) and plosive sounds (Figure 1b), and to support tasks like noise reduction (Figure 1c), volume adjustment (Figure 1d), and track alignment (Figure 1e). P3 emphasized the usefulness of this approach, stating, "*I can look at the waveform for audio and tell [if] I'm going to need to do noise reduction... [or] need to amplify that... That's definitely helpful.*"

Numeric adjustments provided precision and consistency for tasks such as track alignment and volume adjustment. For example, P4 relied on numerical fields in DAWs to align tracks when inserting promotions or ads, explaining, "*Being hard of hearing, I've had to figure out how to do these steps visually and using the actual number count... I know that there are [hearing] podcasters who can rely on [their ears]... I can't do that... because I can't necessarily tell if music is fading out. I don't want to cut off somebody's promo [and] somebody else's ad just because I can't hear it.*"

Textual references, such as scripts and memos, were also essential in editing. In scripted podcasts like fictional dramas, scripts helped participants follow the sound to perform necessary tasks. P3 highlighted their usefulness in sound design, saying, "*I usually have written in the script what sound effects I want at different places... I will generally have the script pulled up as I am working... That definitely helps me process everything that I'm hearing.*" For unscripted podcasts, such as interviews, memos played a similar role. P4 wrote down the editing points they recognized, such as noise and talk overlap, with timestamps while interviewing their guests, saying, "*I start every recording with a blank piece of paper... When I'm done, I have a list of timestamps for things. That speeds up the editing a lot.*"

These visual editing techniques allowed participants to perform audio engineering tasks effectively while saving their hearing. P2 explained their approach, saying, "*[First] I do the entire thing like you just saw [visual editing], and then I go back and listen. I don't listen to the whole thing at all until I have a complete product, just so that I don't overuse my ears and so that, by the time [to listen], I'm still fresh.*"

4.1.2 Limitations of visual editing. While visual editing offered significant advantages, it posed challenges for tasks requiring nuanced auditory perception. Participants working on fictional podcasts (P1, P2, and P3) noted difficulties in applying effects. While showing Figure 1f, P1 noted, "*Sometimes it doesn't directly translate to what it looks like to the intensity of the effect on it... That's the thing I have to listen to.*" However, listening was not always sufficient, as P2 explained, "*It can sometimes just be hard for me to tell audio quality [of] the sound effects because I have trouble hearing.*"

They also faced challenges in selecting and applying sound effects. They frequently relied on free sound effect websites (e.g., Freesound [24]), but finding the right sound effect to match scenes was often difficult. For example, P3 shared, "*[With] a free sound effects website, I often can't be super picky with the sounds I picked... If I'm engineering walking sound effects, I will often try to think about what kind of shoes would these characters wear because different shoes have different sound, like a pair of heels sounds very different than a pair of combat boots.*" Similarly, P2 described the difficulty of adding what they called "life sound", saying, "*Little things like setting dishes on a table, that's kind of a quiet subtle sound, but it needs to be just right, that part is so hard to do because I will listen to the same sound of a mug being put on a table until I feel like 'I don't know what mugs or tables sound like anymore.'*"

4.2 Learning and growth as dHH audio engineers

4.2.1 Adapting to hearing limitations. All participants emphasized the extra hearing energy required as dHH audio engineers. P3 explained, "*Average hearing person does not have to put active effort into understanding things that they are hearing... But for most deaf people, like me, hearing is something that takes active effort and active energy... That is something that does get really exhausting.*"

Acknowledging and adapting to their hearing limitations was fundamental to their work. P1 described the importance of pacing, saying, "*A lot of it is learning how to pace myself... and making sure that I'm not listening too loud and being very self-aware of the way that I am working... [and] that I don't lose my hearing halfway through an episode.*" Similarly, P4 discussed avoiding perfectionism, saying, "*I don't make myself crazy by perfectionism because I'm deaf. If somebody has a problem with imperfect sound from a deaf podcaster, they have problems so that I can't begin to address.*" Flexibility in scheduling was also crucial, as P1 saying, "*Just in case my hearing deteriorates a little bit faster than I was planning,*" Otherwise, they would have to request last-minute accommodations for schedule changes.

4.2.2 Self-teaching through trial and error. All participants were self-taught, learning through trial and error. P3 described this approach as a lot of "*sitting down and trying to figure it out for myself and then googling any specific questions.*" Similarly, P4 reflected, "*You learn by screwing it up, and then at one point, you realize, 'Oh, this is a lot easier.'*"

Participants often questioned whether they were learning correctly. P2 said, "*When I was first learning, I didn't know what I was doing. I was just teaching myself. I was doing a lot of googling and I thought 'What I'm doing? Is this right?' I wasn't sure.*" P1 explained how their hearing influenced their learning approach, saying, "*I could learn how to do it quote-unquote wrong as I went in order to make up for the fact that I sometimes couldn't hear things.*"

P3 and P4 highlighted the inaccessibility of learning materials, such as videos without captions and podcasts without transcripts, compounded the difficulties of self-teaching, forcing them to rely on text-based resources. P3 expressed frustration, saying, "*I would see people recommend a resource and say 'This is invaluable... Everyone who is an audio engineer should check it out.' and then I gonna look into it and discover [they were inaccessible]... I was, honestly, limited to use purely text-based tutorials, which cut down on a lot, because that's obviously a really small percentage of the resources.*"

4.2.3 Challenges to skill development. Participants' adaptability in self-teaching often led to the development of unique methods. For instance, P4 adopted a backward editing approach to avoid disrupting timestamps, saying, "*I learned early on the easiest thing to do is edit backwards, start with the latest timestamp and work toward the front... If I were to start at the beginning, I've skewed all of my time stamps.*" While effective, such methods often deviated from standard practices. P4 acknowledged this, saying, "*I don't know anyone else who edits the way that I do.*" Similarly, P2 noted, "*There's still some stuff that I like don't 100% know... But I know what it does and that's good for me.*"

However, this self-reliant learning style also presented challenges when participants attempted to adopt new tools or workflows. For example, P4 described a stressful attempt to transition to a more advanced DAW, saying, '*A former friend pressured me to try to learn Reaper instead of Audacity... I tried hard and I reached a point where I would just come to sit in my computer and cry because it was so stressful trying to relearn all of these things. I'm sure I have bad habits with editing, but those bad habits aren't hurting anybody else.*"

We also observed participants' reliance on manual editing over automation tools. P1, the only participant who used Adobe Audition alongside Audacity, explained their reluctance to rely on automation, saying, "*I find Audacity can't really automate anything besides noise reduction... And Adobe Audition, I find that because everyone's voices are so different and have different pitches and frequencies, and also background noise, it is not really worth it to automate any of the focused effect... You can definitely automate like compression and EQ and stuff, [though] sometimes you have to fiddle with it a bit. But if you're trying to get down to the nitty gritty, it's best to do it by hand.*" Other participants, who exclusively used Audacity, did not mention automation tools, which may reflect a lack of familiarity or the limitations of the software.

4.3 The role of peer support in dHH audio engineering

4.3.1 Peer feedback for quality and growth. All participants emphasized the importance of peer feedback to ensure quality and address uncertainties due to their hearing. As P2 described, "*Visually, I can check and immediately see [some issues]... And sometimes I'm not quite sure. That's when I usually get someone to listen for me and tell me what they think.*, further noted, "*I almost always have someone double-checking.*"

Beyond quality assurance, peer feedback helped participants uncover gaps in their self-taught knowledge. P3 described how peers introduced new concepts, saying, "*[My friends said,] 'You describe this thing happening here and you don't have any kind of sound for it...' and I'm like, 'That makes a sound?'*" Similarly, P2 said, "*I didn't know for the longest time for forever that when you use audacity, you can turn on a little thing that will show you when something clips... [Before I learned this,] people would tell me like, 'There's some clipping in this' and I was like, 'What are you talking about?... I can't hear it.'*" Further, P4 credited peer input with helping balancing volume, saying, "*Early on, people would say, 'your music is too loud.' I took that feedback and I learned how to level out the music with the speech.*"

4.3.2 Learning to navigate peer support and accommodation. Participants were mindful of not over-relying on peers for support. They sought help only when necessary, expressing concerns about burdening others. P3 explained, "*I really try to figure it out myself, if I can at all. I only ask for help if I really can't get it myself.*" At times, participants declined offers of help if the task required significant effort. For instance, P4 declined free transcription services from friends, saying, "*Before my hearing got as bad as it is, I used to work as a transcriptionist, so I know it's work. And I've had friends offer to do it for free, and I'm like 'Absolutely not. This is work. You need to be compensated.' ... A third of my episodes are not transcribed yet because I haven't been able... to afford it yet.*"

All participants emphasized the importance of becoming better at asking for help. Particularly, those working on fiction podcasts (P1, P2, and P3) highlighted their progress in requesting accommodations in collaborative settings. P1 described when asking accommodations during read-throughs and recording sessions, saying, "*I've gotten a lot better at asking about using services that has transcripts or using closed captions. When I first started out is very much like, 'Oh, don't worry about it... I need to suffer through it.' ... And now I don't mind [asking].*" Similarly, P2 highlighted the importance of clear communication when flexibility was required, stating, "*[When I got sick,] it really affected my hearing... I couldn't hear enough to do it, and I had to delay a lot of stuff... It's so important to communicate that as soon as you can... [It's important that] you're up front about it and you really clear about things that you need and the things that you expect.*"

4.4 Advocating accessibility from within

For each participant, podcasting and audio engineering held personal significance as dHH individuals. For example, P4 shared how podcasting helped them reconnect with the world after a period of isolation, saying, "*I have been on full disability... [After starting podcasts,] what I was enjoying was feeling functional again, having a connection with the outside world again, and just feeling smart again after spending a lot of time feeling isolated.*"

Beyond personal fulfillment, three participants (P1, P2, and P3) described their activity as advocacy for accessibility. P1 emphasized, "*I'm definitely always trying to advocate for better accessibility options... I think a lot of people don't realize the fact that deaf or hard-of-hearing folks like to listen to podcasts.*" Similarly, P2 highlighted the potential impact of collective advocacy, saying,"*As kind of indie creators, we have this pretty small but very vocal group. If we can get that ball rolling, it will hopefully reach people with a bigger audience and a bigger platform and hopefully become the standard.*"

5 Discussion

We presented the current practices of dHH individuals in audio engineering for podcasting. Participants relied on visual editing techniques and peer feedback to address auditory challenges and ensure the quality of their work. As self-taught learners, they navigated limitations in learning resources through trial and error, supported by peer input, and adapted their workflows with unique methods. Despite these adaptations, challenges remain in expanding their skills. Beyond personal fulfillment, participants used their work to advocate for greater accessibility in the field.

In this section, we build on prior research on audio engineering in music by dHH individuals [51] and discuss opportunities to enhance accessibility in podcast audio engineering. Specifically, we examine the differences between podcasting and music audio engineering, highlight the potential of AI-powered technologies, and propose creating accessible learning resources and fostering inclusive communities for dHH individuals in this field.

5.1 Audio engineering for podcasting

Our findings revealed differences between audio engineering for podcasting and music among dHH individuals. Our participants were self-taught, whereas those in music [51] had formal training and education. Also, podcasting emphasizes storytelling with speech and contextually appropriate sounds, while music focuses on the sound elements themselves. Despite similarities such as the use of visuals and peer feedback, these differences shape the tasks and workflows in podcast audio engineering. For instance, participants did not engage in processes common in music production, such as stereo imaging (e.g., panning and phase interference), precise EQ adjustments (e.g., frequency-specific boosts or cuts), EQ automation (e.g., controlling EQ over time), and mastering. This likely reflects podcasting's focus on narrative delivery rather than the detailed sound manipulation, as well as the participants' self-taught backgrounds.

Podcasting also involves unique collaborative workflows that require participants to engage in various production processes [58]. These workflows often include tasks like read-throughs, recording sessions, interviews, and project management. As a result, participants need not only technical audio engineering expertise but also strong communication and coordination skills to collaborate effectively with others. Recognizing these distinctions, we propose recommendations to enhance the accessibility of podcast audio engineering for dHH individuals.

5.2 Technologies to empower dHH audio engineers

5.2.1 Transcript-based audio editing tools. Transcript-based editing tools — such as Adobe Podcast [33], Adobe Premiere Pro [34], and Descript [18] — enable users to edit audio using text rather than waveforms [60]. While these tools are not specifically designed for dHH audio engineers and were not used by participants, they emphasized the value of transcripts in their workflows to reduce listening effort and enhance efficiency. This presents an opportunity to evaluate these tools with dHH audio engineers, investigating how the interactions and features can be designed for them.

5.2.2 Generative AI for sound design. Sound design posed challenges for participants, particularly in sourcing sound effects that matched specific scenes. They often relied on free sound effect websites, which offered limited options and sometimes required settling for imperfect matches. Generative AI tools, such as Stable Audio [44] and ElevenLabs [23], present a promising alternative by enabling users to generate custom sounds through text prompts. Future research could explore how these tools enhance accessibility and streamline the sound design process for dHH audio engineers.

5.2.3 Audio captioning technologies. Audio engineering often requires repeated listening to capture subtle details, posing challenges for dHH individuals. Audio captioning technologies, such as Automated Audio Captioning [46, 75] to describe non-speech sounds in natural language and Audio Difference Captioning [38, 70] to caption differences between two recordings, could help dHH audio engineers better perceive and interpret sounds. However, their applications in creative sound activities remain largely unexplored. Tailoring these tools to the specific needs of dHH audio engineers, such as identifying subtle differences in sound effects, have the potential to enhance accessibility.

5.3 Accessible learning resources and environments

Developing accessible learning resources tailored for dHH audio engineers is crucial to empowering podcast creation. Participants primarily relied on self-teaching through trial and error, a process that required significant emotional and cognitive effort. This approach often resulted in gaps in foundational audio engineering skills and reinforced 'strategy adaptation' [9], where they favored familiar methods over exploring potentially more efficient alternatives. Developing structured and accessible resources can provide dHH audio engineers with a solid foundation in audio engineering and podcasting while exposing them to diverse techniques. Such resources would also reduce the effort required to navigate scattered materials and lower barriers for novices, creating a more inclusive learning environment.

Additionally, strengthening community networks for dHH audio engineers can foster knowledge exchange, mentorship, and skill development. While three participants (P1, P2, and P3) mentioned small, niche communities for dHH podcasters, these primarily focused on indie fictional podcasting. Expanding these networks to encompass broader fields, such as music production where formal training and education are more common [51], could create a more robust and inclusive community. Such networks could serve as platforms for sharing innovative techniques, advocating for accessibility, and fostering a sense of belonging. To support this vision, we have initiated efforts to connect dHH audio engineers across podcasting and music production, aiming to build a supportive and inclusive community.

6 Limitations

This study is limited by its small sample size and the self-taught backgrounds of all participants. Additionally, the hearing statuses of the participants represented a narrow range, excluding individuals such as Deaf and single-sided DHH individuals, as well as cochlear implant users.

Moreover, our research team does not include any DHH members. Both authors have backgrounds in music and audio engineering, the first as an amateur musician and the second with formal training and professional expertise in the field. However, our perspectives on these topics are inevitably influenced by our own experiences as hearing individuals.

Future research should include a more diverse group of participants to provide a more comprehensive understanding of accessibility in podcasting and audio engineering for DHH individuals.

7 Conclusion

This study examined the accessibility of audio engineering in podcasting through interviews with four self-taught dHH audio engineers. Participants relied on visual editing techniques and peer support to address insecurities in sound perception, perform tasks efficiently, and conserve hearing energy. As self-taught learners, they developed their skills through trial and error, often encountering inaccessible educational materials. While this led to the creation of unique workflows, it sometimes limited their ability to adopt new, potentially more efficient methods. To guide future research, we discussed the distinct aspects of podcast audio engineering compared to music, the integration of AI-powered tools, and the development of accessible learning resources and environments.

Acknowledgments

We thank all participants, organizations, and communities that contributed to this study.

References

[1] Podcast Accessibility. 2022. *Podcast Accessibility.* Podcast Accessibility. Retrieved January 22, 2025 from https://podcast-accessibility.com/

[2] Louise Aldridge. 2023. *5 tips for creating accessible podcasts | AbilityNet.* AbilityNet. Retrieved January 22, 2025 from https://abilitynet.org.uk/news-blogs/5-tips-creating-accessible-podcasts

[3] Adam Patrick Bell. 2018. *Dawn of the DAW: The Studio as Musical Instrument.* Oxford University Press, Oxford, UK. https://doi.org/10.1093/oso/9780190296605.001.0001

[4] AJ Beltis. 2021. *6 Tips for Improving Podcast Accessibility.* 3Play Media. Retrieved January 22, 2025 from https://www.3playmedia.com/blog/6-tips-for-improving-podcast-accessibility/

[5] Joe Bennett. 2018. Songwriting, Digital Audio Workstations, and the Internet. In *The Oxford Handbook of the Creative Process in Music.* Oxford University Press, Oxford, UK. https://doi.org/10.1093/oxfordhb/9780190636197.013.28 arXiv:https://academic.oup.com/book/0/chapter/355803461/chapter-ag-pdf/57927633/book_42047_section_355803461.ag.pdf

[6] Larwan Berke, Matthew Seita, and Matt Huenerfauth. 2020. Deaf and hard-of-hearing users' prioritization of genres of online video content requiring accurate captions. In *Proceedings of the 17th International Web for All Conference* (Taipei, Taiwan) *(W4A '20)*. Association for Computing Machinery, New York, NY, USA, Article 3, 12 pages. https://doi.org/10.1145/3371300.3383337

[7] Richard Berry. 2006. Will the iPod Kill the Radio Star? Profiling Podcasting as Radio. *Convergence* 12, 2 (2006), 143–162. https://doi.org/10.1177/1354856506066522 arXiv:https://doi.org/10.1177/1354856506066522

[8] Virginia Braun and Victoria Clarke. 2021. *Thematic analysis: A practical guide.* SAGE, London, UK.

[9] Sarah A Brown, David Menendez, and Martha W Alibali. 2019. Strategy adoption depends on characteristics of the instruction, learner, and strategy. *Cognitive Research: Principles and Implications* 4, 1 (2019), 9.

[10] Cole Burkhardt. 2021. *Most De(a)f Representation in Podcasting - Discover the Best Podcasts | Discover Pods.* DiscoverPods, LLC. Retrieved January 22, 2025 from https://discoverpods.com/deaf-representation/

[11] Mark Cartwright and Bryan A Pardo. 2013. Social-EQ: Crowdsourcing an equalization descriptor map. In *Proceedings of the 14th International Society for Music Information Retrieval Conference, ISMIR 2013 (Proceedings of the 14th International Society for Music Information Retrieval Conference, ISMIR 2013)*, Alceu de Souza Britto, Fabien Gouyon, and Simon Dixon (Eds.). International Society for Music Information Retrieval, 395–400. Funding Information: This work was supported by grant by National Science Foundation Grant No. IIS-1116384. Publisher Copyright: © 2013 International Society for Music Information Retrieval.; 14th International Society for Music Information Retrieval Conference, ISMIR 2013 ; Conference date: 04-11-2013 Through 08-11-2013.

[12] Anna Cavender and Richard E. Ladner. 2008. Hearing Impairments. In *Web Accessibility: A Foundation for Research*, Simon Harper and Yeliz Yesilada (Eds.). Springer London, London, 25–35. https://doi.org/10.1007/978-1-84800-050-6_3

[13] Amelia Chelsey. 2021. Is There a Transcript? Mapping Access in the Multimodal Designs of Popular Podcasts. In *Proceedings of the 39th ACM International Conference on Design of Communication* (Virtual Event, USA) *(SIGDOC '21)*. Association for Computing Machinery, New York, NY, USA, 46–53. https://doi.org/10.1145/3472714.3473622

[14] Amelia Chelsey. 2021. Is There a Transcript? Mapping Access in the Multimodal Designs of Popular Podcasts. In *Proceedings of the 39th ACM International Conference on Design of Communication* (Virtual Event, USA) *(SIGDOC '21)*. Association for Computing Machinery, New York, NY, USA, 46–53. https://doi.org/10.1145/3472714.3473622

[15] Lee Cheng and Iain Mcgregor. 2024. Pedagogical Approaches in Music and Audio Education for Deaf and Hard-of-Hearing Students. *Organised Sound* 29, 2 (2024), 188–196. https://doi.org/10.1017/S1355771824000074

[16] Stacey Copeland. 2018. *A Feminist Materialisation of Amplified Voice: Queering Identity and Affect in The Heart.* Springer International Publishing, Cham, 209–225. https://doi.org/10.1007/978-3-319-90056-8_11

[17] Brecht De Man, Ryan Stables, and Joshua D Reiss. 2019. *Intelligent Music Production.* Routledge, Abingdon, UK.

[18] Descript. 2025. *Descript: Edit Videos & Podcasts Like a Doc | AI Video Editor.* Descript. Retrieved January 22, 2025 from https://www.descript.com/

[19] Christopher Dewey and Jonathan P. Wakefield. 2016. Novel designs for the audio mixing interface based on data visualisation first principles. In *140th Audio Engineering Society International Convention 2016*. Audio Engineering Society, United States, 215–224. http://www.aes.org/events/140/ Cant' find an ISBN or ISSN. Accepted before 1st April 2016. HN 25/10/2017; 140th Audio Engineering Society International Convention, AES 2016 ; Conference date: 04-06-2016 Through 07-06-2016.

[20] Christopher Dewey and Jonathan P Wakefield. 2016. Audio interfaces should be designed based on data visualisation first principles. In *Proceedings of the 2nd AES Workshop on Intelligent Music Production*. Audio Engineering Society, United Kingdom.

[21] Becca Dingman, Garreth W. Tigwell, and Kristen Shinohara. 2021. Designing a Podcast Platform for Deaf and Hard of Hearing Users. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, USA) *(ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 59, 4 pages. https://doi.org/10.1145/3441852.3476523

[22] Becca Dingman, Garreth W. Tigwell, and Kristen Shinohara. 2021. Designing a Podcast Platform for Deaf and Hard of Hearing Users. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, USA) *(ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 59, 4 pages. https://doi.org/10.1145/3441852.3476523

[23] ElevenLabs. 2025. *Free Text to Speech & AI Voice Generator | ElevenLabs*. ElevenLabs. Retrieved January 22, 2025 from https://elevenlabs.io/

[24] Universitat Pompeu Fabra. 2005. *Freesound*. Freesound.org. Retrieved January 22, 2025 from https://freesound.org/

[25] Sarah Florini. 2015. The Podcast "Chitlin' Circuit": Black Podcasters, Alternative Media, and Audio Enclaves. *Journal of Radio & Audio Media* 22, 2 (2015), 209–219. https://doi.org/10.1080/19376529.2015.1083373 arXiv:https://doi.org/10.1080/19376529.2015.1083373

[26] Emma Frid. 2019. Accessible digital musical instruments—a review of musical interfaces in inclusive music practice. *Multimodal Technologies and Interaction* 3, 3 (2019), 57.

[27] Mark Frydenberg. 2008. Principles and Pedagogy: The Two P's of Podcasting in the Information Technology Classroom. *Information Systems Education Journal* 6, 6 (2008). http://isedj.org/6/6/ Also appears in the Proceedings of ISECON 2006: §3354 (ISSN: 1542-7382)..

[28] Campbell Gardner. 2005. There's something in the air: podcasting in education. *EDUCAUSE Review* 40, 6 (2005), 32–47.

[29] Nicole Hennig. 2017. Podcast literacy: Educational, accessible, and diverse podcasts for library users. *Library Technology Reports* 53, 2 (2017), 1–42.

[30] Nicole Hennig. 2017. Podcasts in context. *Library Technology Reports* 53, 2 (2017), 30–38.

[31] Brian J Hracs. 2012. A creative industry in transition: the rise of digitally driven independent music production. *Growth and Change* 43, 3 (2012), 442–461.

[32] Apple Inc. 2024. *Apple introduces transcripts for Apple Podcasts - Apple*. Apple Inc. Retrieved January 22, 2025 from https://www.apple.com/newsroom/2024/03/apple-introduces-transcripts-for-apple-podcasts/

[33] Adobe Inc. 2025. *Adobe Podcast | AI audio recording and editing, all on the web*. Adobe Inc. Retrieved January 22, 2025 from https://podcast.adobe.com/

[34] Adobe Inc. 2025. *Transcribe video to text with Speech to text - Adobe*. Adobe Inc. Retrieved March 5, 2025 from https://www.adobe.com/products/premiere/speech-to-text.html

[35] Podcast Insights. 2024. *How To Start A Podcast: A Complete Step-By-Step Tutorial*. Paid Insights, LLC. Retrieved January 22, 2025 from https://www.podcastinsights.com/start-a-podcast/

[36] Aaron Karp and Bryan Pardo. 2017. HaptEQ: A Collaborative Tool For Visually Impaired Audio Producers. In *Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences* (London, United Kingdom) *(AM '17)*. Association for Computing Machinery, New York, NY, USA, Article 39, 4 pages. https://doi.org/10.1145/3123514.3123531

[37] Colin J Kelly. 2020. *Podcasts with d/Deaf/Hard of Hearing Creators/Contributors | Podcast List on Podchaser*. Podchaser. Retrieved January 22, 2025 from https://www.podchaser.com/lists/podcasts-with-ddeafhard-of-hearing-creatorscontributors-107a4Qtr61

[38] Tatsuya Komatsu, Yusuke Fujita, Kazuya Takeda, and Tomoki Toda. 2024. Audio Difference Learning for Audio Captioning. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, Piscataway, NJ, USA, 1456–1460. https://doi.org/10.1109/ICASSP48485.2024.10448085

[39] Jacobien Korver. 2023. *The Accessibility of Podcasts for Deaf and Hard of Hearing Users: A Survey-Based Study*. Master's thesis. Utrecht University.

[40] Raja Kushalnagar. 2019. Deafness and Hearing Loss. In *Web Accessibility: A Foundation for Research*, Yeliz Yesilada and Simon Harper (Eds.). Springer London, London, 35–47. https://doi.org/10.1007/978-1-4471-7440-0_3

[41] Raja S. Kushalnagar, Walter S. Lasecki, and Jeffrey P. Bigham. 2013. Captions versus transcripts for online video content. In *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility* (Rio de Janeiro, Brazil) *(W4A '13)*. Association for Computing Machinery, New York, NY, USA, Article 32, 4 pages. https://doi.org/10.1145/2461121.2461142

[42] m. nyssim lefford, gary bromham, györgy fazekas, and david moffat. 2021. context-aware intelligent mixing systems. *journal of the audio engineering society* 69, 3 (march 2021), 128–141. https://doi.org/10.17743/jaes.2020.0043

[43] Andrew Leyshon. 2009. The Software Slump?: digital music, the democratisation of technology, and the decline of the recording studio sector within the musical economy. *Environment and planning A* 41, 6 (2009), 1309–1331.

[44] Stability AI Ltd. 2025. *Stable Audio - Generative AI for music & sound fx*. Stability AI Ltd. Retrieved January 22, 2025 from https://www.stableaudio.com/

[45] 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* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 371, 18 pages. https://doi.org/10.1145/3411764.3445412

[46] Xinhao Mei, Xubo Liu, Mark D. Plumbley, and Wenwu Wang. 2022. Automated audio captioning: an overview of recent progress and new challenges. *EURASIP Journal on Audio, Speech, and Music Processing* 2022, 1 (Oct. 2022), 26. https://doi.org/10.1186/s13636-022-00259-2

[47] Oussama Metatla, Fiore Martin, Adam Parkinson, Nick Bryan-Kinns, Tony Stockman, and Atau Tanaka. 2016. Audio-haptic interfaces for digital audio workstations: A participatory design approach. *Journal on Multimodal User Interfaces* 10 (2016), 247–258.

[48] Virgil Moorefield. 2005. *The Producer as Composer: Shaping the Sounds of Popular Music*. The MIT Press, Cambridge, MA, USA. https://doi.org/10.7551/mitpress/5606.001.0001

[49] Joshua Mycroft, Tony Stockman, and J. D. Reiss. 2018. A Prototype Mixer to Improve Cross-Modal Attention During Audio Mixing. In *Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion* (Wrexham, United Kingdom) *(AM '18)*. Association for Computing Machinery, New York, NY, USA, Article 3, 7 pages. https://doi.org/10.1145/3243274.3243290

[50] Keita Ohshiro and Mark Cartwright. 2022. How people who are deaf, Deaf, and hard of hearing use technology in creative sound activities. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (Athens, Greece) *(ASSETS '22)*. Association for Computing Machinery, New York, NY, USA, Article 66, 4 pages. https://doi.org/10.1145/3517428.3550396

[51] Keita Ohshiro and Mark Cartwright. 2024. Audio Engineering by People Who Are deaf and Hard of Hearing: Balancing Confidence and Limitations. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 35, 13 pages. https://doi.org/10.1145/3613904.3642454

[52] World Health Organization. 2021. *World Report on Hearing*. World Health Organization, Geneva, Switzerland. https://www.who.int/publications/i/item/9789240020481 Accessed: January 22, 2025.

[53] William Payne, Fabiha Ahmed, Michael Gardell, R. Luke DuBois, and Amy Hurst. 2022. SoundCells: designing a browser-based music technology for braille and print notation. In *Proceedings of the 19th International Web for All Conference* (Lyon, France) *(W4A '22)*. Association for Computing Machinery, New York, NY, USA, Article 14, 12 pages. https://doi.org/10.1145/3493612.3520462

[54] William Christopher Payne, Alex Yixuan Xu, Fabiha Ahmed, Lisa Ye, and Amy Hurst. 2020. How Blind and Visually Impaired Composers, Producers, and Songwriters Leverage and Adapt Music Technology. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, Greece) *(ASSETS '20)*. Association for Computing Machinery, New York, NY, USA, Article 35, 12 pages. https://doi.org/10.1145/3373625.3417002

[55] Jarrod Ratcliffe. 2014. Hand and Finger Motion-Controlled Audio Mixing Interface. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Goldsmiths, University of London, London, United Kingdom, 136–139. https://doi.org/10.5281/zenodo.1178911

[56] Paul Riismandel. 2018. *Podcast #129 - Deaf Accessibility for Podcasts & Radio - Radio Survivor*. Radio Survivor. Retrieved January 22, 2025 from https://www.radiosurvivor.com/2018/02/podcast-129-deaf-accessibility-for-podcasts-radio/

[57] Jemily Rime, Alan Archer-Boyd, and Tom Collins. 2023. How Will You Pod? Implications of Creators' Perspectives for Designing Innovative Podcasting Tools. *ACM Trans. Multimedia Comput. Commun. Appl.* 20, 3, Article 69 (Oct. 2023), 25 pages. https://doi.org/10.1145/3625099

[58] Jemily Rime, Jon Francombe, and Tom Collins. 2022. How Do You Pod? A Study Revealing the Archetypal Podcast Production Workflow. In *Proceedings of the 2022 ACM International Conference on Interactive Media Experiences* (Aveiro, JB, Portugal) *(IMX '22)*. Association for Computing Machinery, New York, NY, USA, 11–18. https://doi.org/10.1145/3505284.3529977

[59] Jemily Rime, Chris Pike, and Tom Collins. 2022. What is a podcast? Considering innovations in podcasting through the six-tensions framework. *Convergence* 28, 5 (2022), 1260–1282. https://doi.org/10.1177/13548565221104444 arXiv:https://doi.org/10.1177/13548565221104444

[60] Steve Rubin, Floraine Berthouzoz, Gautham J. Mysore, Wilmot Li, and Maneesh Agrawala. 2013. Content-based tools for editing audio stories. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St.

Andrews, Scotland, United Kingdom) *(UIST '13)*. Association for Computing Machinery, New York, NY, USA, 113–122. https://doi.org/10.1145/2501988.2501993

[61] Abir Saha, Darren Gergle, and Anne Marie Piper. 2023. Understanding Peer-to-Peer Instructional Support in an Online Community for Blind Audio Producers. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, Article 48, 15 pages. https://doi.org/10.1145/3597638.3608399

[62] Abir Saha, Thomas Barlow McHugh, and Anne Marie Piper. 2023. Tutoria11y: Enhancing Accessible Interactive Tutorial Creation by Blind Audio Producers. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 220, 14 pages. https://doi.org/10.1145/3544548.3580698

[63] Abir Saha and Anne Marie Piper. 2020. Understanding Audio Production Practices of People with Vision Impairments. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, Greece) *(ASSETS '20)*. Association for Computing Machinery, New York, NY, USA, Article 36, 13 pages. https://doi.org/10.1145/3373625.3416993

[64] Seen and Not Heard. 2020. *Bonus: Let Deaf People Say F**k with Cole Burkhardt.* Spotify. Retrieved January 22, 2025 from https://open.spotify.com/episode/77bVJX1m0sgabW5AbKtRa6

[65] Spotify. 2023. *Everything In Store for Podcast Listeners and Creators This International Podcast Day — Spotify.* Spotify. Retrieved January 22, 2025 from https://newsroom.spotify.com/2023-09-28/international-podcast-day-transcripts-chapters-show-pages-global/

[66] Spotify. 2023. *Podcast production: Tips for every stage of making a podcast.* Spotify. Retrieved January 22, 2025 from https://creators.spotify.com/resources/create/podcast-production?utm_source=chatgpt.com

[67] Spotify. 2024. *How to start a podcast: The creator's step-by-step guide.* Spotify. Retrieved January 22, 2025 from https://creators.spotify.com/resources/how-to/start-a-podcast

[68] Christian J. Steinmetz, Jordi Pons, Santiago Pascual, and Joan Serrà. 2021. Automatic Multitrack Mixing With A Differentiable Mixing Console Of Neural Audio Effects. In *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Toronto, ON, Canada). IEEE, Piscataway, NJ, USA, 71–75. https://doi.org/10.1109/ICASSP39728.2021.9414364

[69] John L Sullivan. 2019. The platforms of podcasting: Past and present. *Social media+ society* 5, 4 (2019), 2056305119880002.

[70] Daiki Takeuchi, Yasunori Ohishi, Daisuke Niizumi, Noboru Harada, and Kunio Kashino. 2023. Audio Difference Captioning Utilizing Similarity-Discrepancy Disentanglement. arXiv:2308.11923 [eess.AS] https://arxiv.org/abs/2308.11923

[71] Atau Tanaka and Adam Parkinson. 2016. Haptic Wave: A Cross-Modal Interface for Visually Impaired Audio Producers. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 2150–2161. https://doi.org/10.1145/2858036.2858304

[72] Michael Terren. 2019. *The grain of the digital audio workstation.* Ph. D. Dissertation. Edith Cowan University. https://ro.ecu.edu.au/theses/2201

[73] Rosemary Thackeray and MaryAnne Hunter. 2010. Empowering Youth: Use of Technology in Advocacy to Affect Social Change. *Journal of Computer-Mediated Communication* 15, 4 (07 2010), 575–591. https://doi.org/10.1111/j.1083-6101.2009.01503.x arXiv:https://academic.oup.com/jcmc/article-pdf/15/4/575/22211187/jjcmcom0575.pdf

[74] What the Deaf?! 2023. *Home | What The Deaf?!* What the Deaf?! Retrieved January 22, 2025 from https://www.whatthedeaf.com/

[75] Xuenan Xu, Zeyu Xie, Mengyue Wu, and Kai Yu. 2024. Beyond the Status Quo: A Contemporary Survey of Advances and Challenges in Audio Captioning. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 32 (2024), 95–112. https://doi.org/10.1109/TASLP.2023.3321968

[76] Alexandros Yeratziotis and Panayiotis Zaphiris. 2018. A heuristic evaluation for deaf Web user experience (HE4DWUX). *International Journal of Human–Computer Interaction* 34, 3 (2018), 195–217.