Towards Inclusive Streaming: Building Multimodal Music Experiences for the Deaf and Hard of Hearing

Thomas B. McHugh
mchugh@u.northwestern.edu
Northwestern University
Evanston, Illinois, USA

Abir Saha
abir@u.northwestern.edu
Northwestern University
Evanston, Illinois, USA

David Bar-El
davidbarel2021@u.northwestern.edu
Northwestern University
Evanston, Illinois, USA

Marcelo Worsley
marcelo.worsley@northwestern.edu
Northwestern University
Evanston, Illinois, USA

Anne Marie Piper
ampiper@uci.edu
University of California, Irvine
Irvine, California, USA

ABSTRACT

As online streaming becomes a primary method for music consumption, the various modalities that many people with hearing loss rely on for their enjoyment need to be supported. While visual or tactile representations can be used to experience music in a live event or from a recording, DRM anti-piracy encryption restricts access to audio data needed to create these multimodal experiences for music streaming. We introduce BufferBeats, a toolkit for building multimodal music streaming experiences. To explore the flexibility of the toolkit and to exhibit its potential use cases, we introduce and reflect upon building a collection of technical demonstrations that bring previous and new multimodal music experiences to streaming. Grounding our work in critical theories on design, making, and disability, as well as experiences from a small group of community partners, we argue that support for multimodal music streaming experiences will not only be more inclusive to the deaf and hard of hearing, but it will also empower researchers and hobbyist makers to use streaming as a platform to build creative new representations of music.

CCS CONCEPTS

• **Human-centered computing → Accessibility technologies**; • **Applied computing → Sound and music computing**.

KEYWORDS

music streaming, accessibility, deaf and hard of hearing, toolkit

ACM Reference Format:
Thomas B. McHugh, Abir Saha, David Bar-El, Marcelo Worsley, and Anne Marie Piper. 2020. Towards Inclusive Streaming: Building Multimodal Music Experiences for the Deaf and Hard of Hearing. In *CHI '21: ACM Conference on Human Factors in Computing Systems, May 08–13, 2021, Online Virtual Conference*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/1122445.1122456

CHI '21, May 08–13, 2021, Online Virtual Conference

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Many people consider music a core part of their cultural lives and personal identities. Despite "listening" or "hearing" being the primary language to describe experiencing music, which implicitly suggests that music is an auditory-only experience, music is, rather, a multimodal experience. People go to concerts to see their favorite artists perform music, they feel the vibrations caused by percussive instruments, and their love for watching music videos on platforms like MTV even re-energized a failing record industry [8]. We cannot ignore the existing hegemonic unawareness in our language and rhetoric regarding how we refer to non-dominant ways of experiencing life, which silences those who live through these experiences, particularly those with disabilities [29].

Embracing music as a multimodal experience allows its consumption to be accessible to those with hearing loss, and many deaf and hard of hearing people enjoy music through a combination of auditory, visual, and vibrational experiences [41]. The emphasis of modalities used to consume music is also unique to the individual because being deaf or hard of hearing represents a wide spectrum of hearing ability; a profoundly deaf person may rely on vibration or visuals [26], while a hard of hearing person may use visuals, such as lyric videos, in partnership with audio [34]. Being deaf or hard of hearing also does not immediately mean that one identifies as part of a Deaf cultural community [29], which has a diverse range of opinions on how to enjoy music and if music is even of any importance to Deaf culture [16].

The multimodal nature of music is often seen best in physical music environments like concerts. However, designs for most digital music interfaces do not support multiple experiential modalities for music consumption. Because of this, if one interacts with musical content through a smartphone, non-auditory modalities are often muted. Researchers have worked on designing digital music experiences for the deaf and hard of hearing through visualizations and haptic feedback [41]. Using a multitude of sound analysis algorithms [36], these representations of digital music can turn raw audio data into beautiful and contextual non-auditory music experiences.

Recently, online music streaming has become a central way for people to consume music [12]. As of December 2020, Spotify had approximately 345 million active users [7]. Digital streaming services come with many conveniences: access to a vast library of music, the ability to listen anywhere, personalized recommendations, and more. However, the drawbacks of these services include not owning the music one listens to and not controlling the interface one uses to get access to this music. Evaluating and improving the inclusivity of these music streaming platforms for people with hearing loss, though, has been a relatively under-explored area of research.

Conversations with our community partners and their desire to use music streaming services encouraged us to explore the possibility of designing an interface to support multimodal experiences for streaming. A key component in designing rich and accurate visualizations or haptic experiences is having access to raw audio buffers to analyze and transform the music. This presented a considerable challenge, since Digital Rights Management (DRM) encryption, used by almost all streaming services, obstructs access to the raw buffers on most software platforms [42].

In this paper, we present the ideation and implementation of BufferBeats, a toolkit for creating multimodal experiences for streaming services. To circumnavigate the obstacles we faced, we introduce a mobile digital intermediary that forwards raw audio buffers from music streaming services to a mobile app that can transform these buffers into multimodal experiences. Inspired by the ideas of our partners, we then designed and prototyped both a visual and haptic interface for experiencing streaming music on a smartphone to explore the capabilities of this new platform. Finally, we advocate for streaming services to build or expose multimodal and inclusive music expression and discovery interfaces.

2 RELATED WORK

This is the first study, to the best of our knowledge, to explore the inclusivity and accessibility of music streaming services. Additionally, we present the design of two multimodal prototype interfaces specifically for music streaming. In this work, we built upon prior research in two main fields. First, the accessibility of media streaming services. Secondly, the development of non-auditory digital music interfaces.

2.1 Accessibility in Media Streaming Services

In our own examination of the different features of music interaction present on Apple Music and Spotify on the iOS software platform, we noticed these applications have a limited set of unique non-auditory interfaces. Both systems have recently implemented their own lyric presentation interfaces when playing a song. Spotify allows users to view all lyrics in the current song through scrolling or the SoundHound app. Apple Music implemented in the iOS 13 update live scrolling lyrics built directly into the app. Both platforms also display album artwork of a song. Moreover, Spotify will display a music video of a song, if a popular video is available, in the song player interface. However, Apple Music does not have this feature. Despite progress seen through updates to streaming services, Morreale and Eriksson highlight that these updates can sometimes degrade the accessibility of music streaming, given the lack of agency and ownership that these platforms provide to users [35].

Little work has been done to explore the accessibility of the music consumption component of music streaming, especially work aimed at improving access for those with hearing loss. Choi [13] argued that full captioning and lyrics are required to ensure equal access in streaming services for the deaf. An investigation of the accessibility of online streaming services by Granberg explored the accessibility of Netflix and Spotify for different disabilities [31]. However, this work only examined whether a user with profound hearing loss could navigate the user interface of Spotify and not whether the user could actually experience the music the service offers. The present paper aims to begin filling this gap, by designing multimodal interfaces for digital music streaming.

2.2 Non-Auditory Digital Music Interfaces

Our work towards designing a multimodal accessible music streaming experience is informed by a growing body of literature on non-auditory modalities (such as haptics, vibration, and visualization) to design music experiences for people with hearing loss.

2.2.1 Vibrotactile & Haptic Interfaces. Prior work has highlighted that different musical properties like rhythm and tempo can be experienced through vibration [19]. Merchel et al. [32] found that vibrations perceived on body surface play a key role in music perception and the experiential components of concerts. Karam and Fels designed Model Human Cochlea (MHC), a sensory substitution technique to translate musical emotions through an audio-haptic display [25]. Building upon this work, Karam et al. used MHC technique to design a chair with embedded voice coils that provide vibrotactile translation of music to deaf users [26]. Jack, McPhearson and Stockman [24] developed audio-tactile furniture for deaf users and reported that deaf users found the furniture experience positive in cases of highly rhythmic music. Nanayakkara et al.[37] developed a multimodal music experience consisting of a visual display and a haptic chair for people with hearing loss. Their findings suggest that users found the haptic chair to enhance their music experience. While many of these findings show these these tools to be useful for experiencing music for deaf and hard of hearing people, many of these tools are large and not portable. This lack of portability limits the applicability of such tools for music streaming experiences, which people usually engage with via smartphones.

There has been a number of different systems and devices built that explore the representation of music through haptic interfaces. Petry et al. performed a comprehensive review and comparison of available visual and vibrotactile interfaces for digital music representation in the work to develop Muss-Bits, a sensor-display pair that can provide visual and vibrotactile feedback to enable exploration of musical sound [41]. They suggest in their review of current multimodal music interfaces that the existing tools aimed at improving the music experience for the deaf and hard of hearing have not become frequently used because they are not adaptable and simple for the user. While their work frames their contribution as sensory substitution, instead of challenging the notion that music is in its purest form not just an auditory experience [20], the platform was able to be used to explore music with multiple components such as voice, hand clapping and bass drum. A controlled study with deaf children found that Muss-Bits made the children feel more confident in rhythm discrimination and production [40]. Trivedi, Alqasemi and Dubey designed a vibrotactile sleeve with bone conduction speakers [44]. A study evaluating the system suggests that it can convey basic music information. Despite this, relatively little work has been done to explore native haptic or vibrotactile music support for mobile phones.

2.2.2 Visual Interfaces. A body of research has also examined the use of visuals to convey music. The purpose of these visualizations aims to improve the accessibility of music and to improve the Music Information Retrieval (MIR) field. Isaacson evaluated a number of different music visualizations used to improve music analysis such as time plot graphs, spectrograms, and tonal landscapes [23]. Bergstrom et al. developed Isochords, a visualization of music that improves upon the classification of musical structures compared to a standard staff-based notation [9]. Gumulia et al. presented a novel method for comparing different songs using a visualization of the perceived speed of the musical composition [21]. Their work goes beyond analyzing music through tempo by integrating the number of notes per time unit and rhythms of the composition into the visualization.

Improving upon the existing literature on mapping sound to colors, Ciuha et al. presented a method for visualizing a group of concurrent tones with a single color by mapping the circle of thirds to a color wheel [14]. Weinel introduced Cyberdream VR, a tool to generate visualizations for rave and vaporwave music [47]. Outram used existing literature on mapping sound to colors to generate contextual visualizations that could also be explored in a virtual space through an 'orbital mode' [38].

Kim et al. developed Seen Music [27] as a tangible and visual representation of music for children with hearing loss. Seen Music represents sounds of a violin through a tangible flower pot, where the flowers move, bloom or shut off depending on the notes played on a violin. Hiraga and Kato found that accompanying drum sounds with visuals helped people with and without hearing loss to identify the emotions conveyed in music [22]. Vy et al. developed EnACT, a software tool that generated animated and colorful lyric captions, to convey the emotions and tone of a song in a video [46]. Participants in a study evaluating EnACT found these animated text-based visuals improved their understanding of the portrayed emotions of a song. While these tools showcase innovative ways of expressing music through new modalities, these tools do not integrate well with the new streaming modal of music distribution, limiting what music can be experienced through them.

Our review of prior work revealed that, while an extensive body of research has examined the use of visual and haptic modalities in designing accessible music interfaces for the deaf and hard of hearing, there has been little work on exploring the use of non-auditory modalities in online streaming. Therefore, there is a need to understand how these many digital music interfaces can work with existing music streaming services. Our paper explores the current challenges in enabling non-auditory experiences in online streaming and presents our implementation of a digital intermediary that works around these challenges.

3 FORMATIVE WORK

While a deeper evaluation of the power dynamics of the digitization of music expression is needed, it was paramount for us to approach our design work as partners with community members whose relevant experiences are critical to understanding the problems of equity in the music expression space. Through informal interviews to ground our need-finding process, we spoke with two youth music educators, one who has taught at a music education institution for deaf youth musicians and one who is a high school band instructor who assists a deaf family member. We also conducted the same process of interviews with one deaf and one hard of hearing adult, who were both self-described active music consumers and who enjoy music. The interviews were mostly guided by the participants' specific musical interests. We used an inductive coding method to analyze the interviews and to synthesize themes and patterns in our participant's discussions about music and accessibility. Through our conversations with these community partners, three common discussion themes emerged and shaped the direction of our need-finding process.

3.0.1 Physical vs. Digital Music Environments. All of our community members mentioned that they saw physically-supported music environments as more supportive of the way they enjoyed music compared to digital music players. Examples from themselves, their students, and their family members included holding onto an instrumentalist to feel the song being playing, going to a concert, and rocking out to the vibrations of a speaker.

3.0.2 Mobile Audio Player Limitations. We also noted from our discussions that our partners found digital experiences, such as using headphones or mobile phone speakers, as particularly unsupportive of the ways they experience music. Despite this, many of our partners expressed a deep interest in using digital music streaming services.

3.0.3 Music Streaming Inaccessibility. Finally, our partners also described to us that, for many of their friends and peers without hearing loss, platforms like Spotify and Apple Music had become the home to social interactions like building collaborative playlists that they often felt excluded from because they could not fully experience the services being used. One particular case that our music educator partners described to us was that they often wanted their students to listen to recordings of songs to benefit their lesson. However, getting access to these songs with an inclusive interface was difficult and, thus, in turn limits the number of opportunities to listen to pieces of music that would benefit a young musician.

Using these experiences and stories as our grounding, we began to explore the accessibility problems of the digital music streaming service space. We sought to create a contextual visualization interface that would support all songs on the Apple Music and Spotify platforms. We chose to design for the iPhone platform as that was the smartphone platform that all of our partners used. Our partners wished for such a device to be used in public areas and to be inconspicuous. They also disliked the idea of wearing an electronic armband or other device that would seem to them unnatural. As we found through our prototyping process, building such a device required inventing a new architecture and toolkit to overcome barriers to extending streaming interfaces, which is where this work's contribution lies.

4 ENABLING ACCESSIBLE STREAMING

As we began working on a prototype of our interface, we realized that the iOS APIs for audio transformation and analysis all failed to work when playing music from streaming services. As we investigated, we identified that Apple's FairPlay [1] media encryption service, to protect against piracy of copyrighted streaming content, prevents all audio analysis interfaces from accessing the raw audio buffers from these streaming platforms. Unlike some open software platforms that provide interfaces for capturing all sounds coming from the smartphone speaker, iOS's sandboxing protections prevent this.

When investigating Google's Android system, we found they do provide a programming interface to access low-quality buffers to generate visualizations [6], but it is unclear whether the quality degradation would impair advanced haptic-based representations of digital music and if the system would even work for streaming music that is encrypted using Widevine [3], Google's digital rights encryption module.

4.1 Overcoming Content Encryption

We investigated multiple approaches to bypassing the encryption obstacle, including cracking the encryption algorithm. This proved unfeasible as none of the major content decryption services have been able to implement a decrypting method for the SAMPLE-AES encryption algorithm used by most content encryption services.

Having failed to find a solution to the encryption obstacles, we considered implementing our system using the "analog hole" [43], which simply records the audio using the built-in system microphone. While this would return a low-quality audio buffer that might be usable, it would also include any other noises in the environment since it is recording all sound produced in the environment. If a user were in a loud environment, the system would become totally non-functioning with no way to inform the user. We decided not to pursue this solution, as this path would simply not address the root design obstacle we were facing.

Additionally, we explored developing a one-way device that streamed buffers over Bluetooth and was presented through a vibrotactile interface. However, this solution would not support the ability to create software and on-device interfaces, such as visualizers. We knew that our system, if it was to use an intermediary device to overcome the encryption obstacles, would have to support sending signals back to the device for processing.

Finally, we examined the properties of designing an AirPlay capable streaming receiver. During this exploration, we realized that a non-sandboxed software environment, such as a single-board Linux computer like a Raspberry Pi, could be setup to mimic the AirPlay speaker protocol and then forward back captured audio buffers in real time to the smartphone device. We proceeded optimistically with this path, with the goal of fully overcoming the limitations in accessing audio buffer data despite the streaming encryption techniques.

4.2 Capturing Streaming Service Audio Data

Developing an AirPlay receiver to decode encrypted audio buffers on a non-sandboxed device proved to be successful as we were able to capture and send the raw audio buffers of songs from both the Spotify and Apple Music platforms back to a smartphone device.

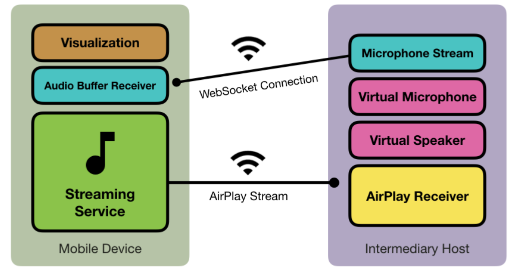

Figure 1: Outline of the software architecture to capture audio buffers from a streaming service.

Using the AirPlay protocol [2, 4, 5, 45], we implemented a Linux AirPlay receiver that output incoming streams to a virtual speaker. Setting the receiver's output speaker to a virtual speaker that captured the audio data from the output speaker while also feeding it into the devices default microphone allowed us to access the streaming audio data from the smartphone as a standard audio input. We created a Node.JS web server that used Socket.io, a WebSockets-based JavaScript library, to forward a streaming song's audio buffers from the provided virtual audio input back to the host smartphone. The process of capturing and forwarding audio buffer content was initiated and stopped whenever a device connected to the intermediary host through the WebSockets connection and the AirPlay receiver.

We connected the entire process on the Linux system into a single executable, and we embedded our audio observer and processor into a library that can be used within any iOS application[1] . To simplify the process of analyzing the raw audio buffers, the library allows the application either to retrieve the raw audio buffer data or a set of built-in audio transformation algorithms including the Yin [18] and Fast Fourier Transform (FFT) [11] algorithms.

5 PROTOTYPE INTERFACE EXAMPLES

Having introduced this intermediary to fix the technical obstacles we faced, we sought to prototype different interfaces to explore the advantages and limitations to the toolkit we developed. We designed two interfaces, one visual and one haptic, that both respond to the analysis of raw audio buffers provided by the intermediary.

5.1 Visual Modal Interface

Inspired by discussions with one of our community partners who experiences synesthesia and experiences colors when playing or experiencing music, we worked to create a visualization of the song by generating a new frame of the visualization from each unique audio buffer in the context of the audio buffers that made up the previous frames of the visualization.

Our initial system classified an audio buffer's context within a song modified work by Davis and Mohammad to synthesize a small musical representation of the trained images and to compare music

[1] https://github.com/InclusiveTechNU/bufferbeats

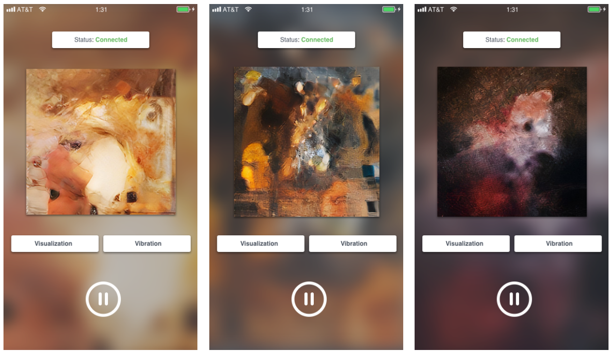

Figure 2: Example visualizations generated from the interface. Audio for each visualization was taken from the Moodo dataset, and was categorized by the authors as upbeat pop, classical with heavy brass, and dark and mysterious.

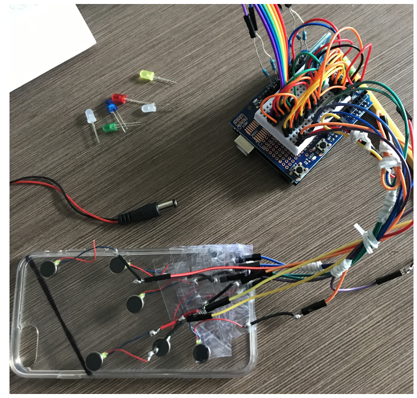

Figure 3: Vibration motors attached to a phone case which activate from audio buffers sent via Bluetooth to an Arduino.

being played to the most similar algorithmically generated song [17]. However, it soon became clear that the prototype was limited and that trying to synthesize music from paintings, while novel, did not generate particularly interesting visualizations because of the limitations of the generated songs.

Rethinking what context was actually required to create a contextual visualizations, we trained a regression model to synthesize a color representation of the audio buffer trained on Matevz et al.'s Moodo dataset of colors that participants associate with different audio clips [39]. The classifier received the new audio buffer as an input and output an HSV color space prediction for the color best representative of the current audio buffer.

It was important that the visualizations that were generated were not just of random pixel values but would actually connect and empathize with the user, emulating the same effect that listening to a song can have. To generate new images, we trained a Variational Autoencoder [28] with a 10-dimension latent space on 4,000 pictures of a diverse set of artwork from the WikiArt dataset [33]. Using the trained decoder model and a generated latent vector, new images were synthesized for each of the visualization's frames, and the difference in pixel values from the previous and current frame was changed by transitioning the color difference.

To calculate the latent vector that would be used to generate a frame in the visualization, the regressed color from the audio classifier component was compared to the most frequent color of each of the images in the WikiArt training set. The top 5 training images with the highest color similarity to the regressed color, as determined using the CIEDE2000 similarity algorithm [30], had their encoded latent vectors averaged, and the resulting vector was averaged with the latent vector of the visualization's current frame to create a new visualization frame latent vector.

5.2 Haptic Modal Interface

To explore the creation of a different modal interface for music streaming using the intermediary, we developed a mobile haptic interface for feeling synthesized vibrations from the music. Using 6 vibration motors connected to an Arduino and a Bluetooth module, we prototyped an interface that connected wirelessly to the music streaming intermediary and activated specific vibration motors depending on the intensity of the audio buffer received.

To determine the proper motors to activate from a buffer, a FFT was performed on each audio buffer, and the normalized magnitude range was split into 6 ranges, one for each motor. A specific audio buffer would activate whichever motor was linked to the range the normalized FFT magnitude value was contained in.

We initially tested this device on a larger breadboard where users could hold the vibration motors or see using LEDs which vibration motor would activate. However, the initial prototype's experience was bulky and not mobile. We switched to using a smaller breadboard and attached the motors to the back of a soft plastic iPhone case. As we continue improving our prototype, we hope to embed the entire system within the case of the phone. The motors were placed in different parts of the case's back, and the case was attached to the back of an iPhone. When music is streaming, due to different motors being activated depending on the intensity of each audio buffer, a user holding the iPhone will experience vibrations in different parts of the case at different times.

6 DISCUSSION

In this paper, we display a way to overcome the technical obstacles to develop a multimodal music streaming experience for deaf and hard of hearing music consumers. Through this design process, we surface tensions between the realistic necessity of anti-piracy measures in protecting the financial and creative rights of artists and the ways these measures stand to affect inclusivity in these streaming services. We argue that, in addition to protecting content creators' rights, streaming services should also support additional accessible interfaces for a wider audience to experience digital music. By thinking critically about our dominant description of music experiences as purely auditory, we can redesign music experiences for a more inclusive digital world.

To contextualize our work within the broader literature on Deaf culture, we acknowledge that there remains a debate over the importance of music within this community. Both of the people whom we worked with were not born Deaf or hard of hearing, and they do not self-identify as being a part of Deaf culture [15]. Therefore, this work does not seek to impose a view on the importance of music for those with hearing loss but only to support those who have hearing loss and who also enjoy, have enjoyed, or wish to enjoy digital music streaming experiences.

Our hope is that this toolkit will help others, whether they be researcher's prototyping a new tool or hobbyist makers trying to increase the accessibility of their devices, build novel designs for music accessibility, instead of focusing on solving obstacles to achieving basic audio processing for streaming services. This work is only one of many examples [48] showing that technologies such as machine learning, microcontrollers, and APIs can allow anyone with the knowledge to make new inclusive microworlds of creative expression, exploration, and thought. Instead of placing the role of those with disabilities as only early adopters and users of these tools, which means they are the first to experience the harm of them [10], digital fabrication has the potential to make those with disabilities the early pioneers and developers of these technologies.

ACKNOWLEDGMENTS

We are indebted to the individuals who spent their time guiding our design processes through sharing stories, experiences, and thoughts. We also thank Emily Wang, Maitraye Das, Adam Goodkind, and Lily Barghi for valuable discussions.

REFERENCES

[1] [n.d.]. FairPlay. https://developer.apple.com/streaming/fps/
[2] [n.d.]. Shairport Sync. https://github.com/mikebrady/shairport-sync
[3] [n.d.]. Widevine. https://www.widevine.com/
[4] 2011. AirTunes 2 Protocol. https://git.zx2c4.com/Airtunes2/about
[5] 2017. open-airplay. https://github.com/jamesdlow/open-airplay
[6] 2019. Visualizer. https://developer.android.com/reference/android/media/audiofx/Visualizer
[7] 2020. Spotify Company Info. https://newsroom.spotify.com/company-info/
[8] Jack Banks. 1997. Video in the machine: the incorporation of music video into the recording industry. *Popular music* 16, 3 (1997), 293–309.
[9] Tony Bergstrom, Karrie Karahalios, and John C Hart. 2007. Isochords: visualizing structure in music. In *Proceedings of Graphics Interface 2007*. 297–304.
[10] Jeffrey P Bigham and Patrick Carrington. 2018. Learning from the Front: People with Disabilities as Early Adopters of AI.
[11] Peter Carr and Dilip Madan. 1999. Option valuation using the fast Fourier transform. *Journal of computational finance* 2, 4 (1999), 61–73.
[12] Charlie C Chen, Steven Leon, and Makoto Nakayama. 2018. Are You Hooked on Paid Music Streaming?: An Investigation into the Millennial Generation. *International Journal of E-Business Research (IJEBR)* 14, 1 (2018), 1–20.
[13] Frances Choi. 2016. Equal Access Requires Full Captioning of Music and Song Lyrics for the Deaf and Hard of Hearing. *Loy. LA Ent. L. Rev.* 37 (2016), 271.
[14] Peter Ciuha, Bojan Klemenc, and Franc Solina. 2010. Visualization of concurrent tones in music with colours. In *Proceedings of the 18th ACM international conference on Multimedia*. 1677–1680.
[15] Alice-Ann Darrow. 1985. Music for the Deaf. *Music Educators Journal* 71, 6 (1985), 33–35.
[16] Alice-Ann Darrow. 1993. The role of music in deaf culture: Implications for music educators. *Journal of Research in Music Education* 41, 2 (1993), 93–110.
[17] Hannah Davis and Saif M Mohammad. 2014. Generating music from literature. *arXiv preprint arXiv:1403.2124* (2014).
[18] Alain De Cheveigné and Hideki Kawahara. 2002. YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America* 111, 4 (2002), 1917–1930.
[19] D. P. Ellis. 2006. Extracting Information From Music Audio. , 32-37 pages.
[20] David Fourney. 2012. Can computer representations of music enhance enjoyment for individuals who are hard of hearing?. In *International Conference on Computers for Handicapped Persons*. Springer, 535–542.
[21] Anastasia Gumulia, BartBomiej Puzon, and Naoko Kosugi. 2011. Music visualization: predicting the perceived speed of a composition–misual project–. In *Proceedings of the 19th ACM international conference on Multimedia*. 949–952.
[22] Rumi Hiraga and Nobuko Kato. 2006. Understanding Emotion Through Multimedia: Comparison Between Hearing-impaired People and People with Hearing Abilities. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility* (Portland, Oregon, USA) *(Assets '06)*. ACM, New York, NY, USA, 141–148. https://doi.org/10.1145/1168987.1169012
[23] Eric J Isaacson. 2005. What You See Is What You Get: on Visualizing Music.. In *ISMIR*. 389–395.
[24] ROBERT Jack, ANDREW McPherson, and TONY Stockman. 2015. Designing tactile musical devices with and for deaf users: a case study. In *Proceedings of the International Conference on the Multimodal Experience of Music, Sheffield, UK*.
[25] Maria Karam and Deborah Fels. 2008. Designing a model human cochlea: issues and challenges in crossmodal audio-haptic displays. (01 2008), 8. https://doi.org/10.4108/ICST.AMBISYS2008.2837
[26] M. Karam, G. Nespoli, F. Russo, and D. I. Fels. 2009. Modelling Perceptual Elements of Music in a Vibrotactile Display for Deaf Users: A Field Study. In *2009 Second International Conferences on Advances in Computer-Human Interactions*. 249–254. https://doi.org/10.1109/ACHI.2009.64
[27] Jeeeun Kim, Swamy Ananthanarayan, and Tom Yeh. 2015. Seen music: ambient music data visualization for children with hearing impairments. In *Proceedings of the 14th International Conference on Interaction Design and Children*. ACM, 426–429.
[28] Hanock Kwak and Byoung-Tak Zhang. 2016. Generating images part by part with composite generative adversarial networks. *arXiv preprint arXiv:1607.05387* (2016).
[29] Scott Lunsford. 2005. Seeking a Rhetoric of the Rhetoric of Dis/abilities. *Rhetoric Review* 24, 3 (2005), 330–333.
[30] M Ronnier Luo, Guihua Cui, and Bryan Rigg. 2001. The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur* 26, 5 (2001), 340–350.
[31] Thomas Mejtoft. 2015. Proceedings of the 11th Student Conference in Human–Information-Thing: Interaction Technology 2015. In *Student Conference in Human–Information-Thing Interaction Technology, Umeå, May 2015*. Umeå universitet.
[32] Sebastian Merchel and M. Ercan Altınsoy. 2014. The Influence of Vibrations on Musical Experience. *Journal of the Audio Engineering Society* 62 (04 2014), 220–234. https://doi.org/10.17743/jaes.2014.0016
[33] Saif Mohammad and Svetlana Kiritchenko. 2018. Wikiart emotions: An annotated dataset of emotions evoked by art. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.
[34] Jorge Mori and Deborah I Fels. 2009. Seeing the music can animated lyrics provide access to the emotional content in music for people who are deaf or hard of hearing?. In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*. IEEE, 951–956.
[35] Fabio Morreale and Maria Eriksson. 2020. " My Library Has Just Been Obliterated": Producing New Norms of Use Via Software Update. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
[36] Meinard Mueller, Bryan A Pardo, Gautham J Mysore, and Vesa Valimaki. 2018. Recent Advances in Music Signal Processing [From the Guest Editors]. *IEEE Signal Processing Magazine* 36, 1 (2018), 17–19.
[37] Suranga Chandima Nanayakkara, Lonce Wyse, Sim Heng Ong, and Elizabeth A Taylor. 2013. Enhancing musical experience for the hearing-impaired using visual and haptic displays. *Human–Computer Interaction* 28, 2 (2013), 115–160.
[38] Benjamin I Outram. 2016. Synesthesia audio-visual interactive-sound and music visualization in virtual reality with orbital observation and navigation. In *2016 IEEE International Workshop on Mixed Reality Art (MRA)*. IEEE, 7–8.
[39] Matevž Pesek and Primož Godec. 2014. INTRODUCING A DATASET OF EMOTIONAL AND COLOR RESPONSES TO MUSIC. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*. Taipei.
[40] Benjamin Petry, Thavishi Illandara, Don Samitha Elvitigala, and Suranga Nanayakkara. 2018. Supporting Rhythm Activities of Deaf Children Using Music-Sensory-Substitution Systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 486, 10 pages. https://doi.org/10.1145/3173574.3174060
[41] Benjamin Petry, Thavishi Illandara, and Suranga Nanayakkara. 2016. MuSS-bits: sensor-display blocks for deaf people to explore musical sounds. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction*. ACM, 72–80.
[42] Stefan Pham, Franck Russel Kuipou, Stefan Arbanowski, and Stephan Steglich. 2018. On the current state of interoperable content protection for internet video streaming. In *2018 IEEE Seventh International Conference on Communications and Electronics (ICCE)*. IEEE, 13–17.
[43] Douglas C Sicker, Paul Ohm, and Shannon Gunaji. 2006. The analog hole and the price of music: An empirical study. *J. on Telecomm. & High Tech. L.* 5 (2006), 573.
[44] Urvish Trivedi, Redwan Alqasemi, and Rajiv Dubey. 2019. Wearable musical haptic sleeves for people with hearing impairment. In *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, 146–151.
[45] Clément Vasseur. 2012. Unofficial AirPlay Protocol Specification. https://nto.github.io/AirPlay.html

[46] Quoc V Vy, Jorge A Mori, David W Fourney, and Deborah I Fels. 2008. EnACT: A software tool for creating animated text captions. In *International Conference on Computers for Handicapped Persons*. Springer, 609–616.
[47] Jonathan Weinel. 2019. Cyberdream VR: Visualizing Rave Music and Vaporwave in Virtual Reality. In *Proceedings of the 14th International Audio Mostly Conference: A Journey in Sound on ZZZ*. 277–281.
[48] Marcelo Worsley, David Barel, Lydia Davison, Thomas Large, and Timothy Mwiti. 2018. Multimodal interfaces for inclusive learning. In *International Conference on Artificial Intelligence in Education*. Springer, 389–393.