Toward a Tool for Music Personalization for Cochlear Implant Users

Lloyd May[1,*], Aaron Hodges[1], So Yeon Park[1], Blair Kaneshiro[1] and Jonathan Berger[1]

[1]Stanford University, CCRMA, 660 Lomita Drive, Stanford, California, USA, 94305

Abstract

Listening experiences among people using assistive hearing technologies, such as hearing aids and/or Cochlear Implants (CIs), are diverse and impacted by various physiological and social factors. Given individual differences in such factors and in subjective preferences related to music enjoyment, there is a need for personalized, customizable solutions to improve music listening experiences for CI users. In this vision paper, we present methods and initial results of an exploratory study designed to gain insights into the empathetic processes involved in audio mixing and listening. Ten expert listeners (ELs) (each of whom has used a CI for over one year), and ten professional audio engineers (AEs) were recruited. AEs were then asked to mix and process audio under various conditions including using a CI-simulator as well as with written feedback from an EL or peer AE. All participants completed multiple interviews and surveys to better understand the nuances of this customization and listening process. Findings from this initial study will inform the creation of a web-based tool to empower CI users to customize their music listening experience.

Keywords

Accessibility, Aural Diversity, Music Personalization

Cochlear implants (CIs) are electronic, hearing-assistive devices that convert acoustic signals to electrical signals which then stimulate the cochlea. They are generally used by folks diagnosed with a high degree of hearing loss, and for whom standard hearing aids may not work as desired. CIs are optimized for speech perception, therefore experiences of complex auditory stimuli, such as music, differ greatly among CI users [1, 2]. Perception of certain musical features, such as tempo and rhythm, are comparable between CI users and those with traditional hearing, while the perception of melodic, harmonic, and timbral information differs more [3]. Current techniques to enhance the enjoyment of music listening among CI users include the adjustment of signal processing parameters on the device itself, algorithmic approaches to pre-process or "optimize" a piece of recorded music, and the creation of novel music specifically tailored to CI users [4]. While there are certainly merits to these approaches, they neither fully enable CI listeners to customize their experiences nor account for the diversity of listening experiences, strategies, and goals.

A potential site of exploration to develop a tool for CI users to customize their experience is the manipulation of multi-track or stem renderings [1] of a piece of recorded audio. The prevalence of multi-track and/or stem recordings being released, leaked, or algorithmically separated illustrates a strong desire from music makers, remixers, and general listeners to personalize and experiment with the musical material they are engaging with. This presents an opportunity to further develop a stem remixing and processing tool that centers D/deaf and Disabled joy.

We aim to create a tool that highlights the autonomy, curiosity, and mastery of one's own listening experience through multiple levels of controllable, stem-level processing and personalization. In the first phase of this project, ten professional audio engineers (AEs) and ten expert listeners with CIs (ELs) were recruited to complete a multi-step process of music mixing, revising, giving feedback, and evaluating mixes. Multi-track music clips from various genres (folk, hip-hop, electronic dance music, soft pop, and funk) and a speech-in-noise excerpt were given to the AEs to mix in four stages: (1) A baseline mix, as if mixing for commercial release, (2) Re-mix "Mix 1" through a CI simulation plug-in [2], (3) Re-mix "Mix 2" after receiving written feedback from an EL with the option for additional email correspondence, and (4) Re-mix "Mix 3" after receiving written feedback from a peer AE who is aware of all the details of the study.

Joint Proceedings of the ACM IUI Workshops 2023, March 2023, Sydney, Australia

*Corresponding author.

lloyd@ccrma.stanford.edu (L. May); athodges@stanford.edu (A. Hodges); syjpark@stanford.edu (S. Y. Park); blairbo@ccrma.stanford.edu (B. Kaneshiro); brg@ccrma.stanford.edu (J. Berger)

0000-0003-4692-8261 (L. May)

[1]Stem or multi-track representations refer to a piece of recorded music where each audio element, instrument, or group of instruments is available on their own track that can be individually manipulated before being summed into the final piece.

[2]While there is mixed evidence supporting the efficacy of disability simulation in empathy generation among non-disabled people, the effect of context and contact with people with disabilities has led to increased empathy generation [5].

The AEs were asked to mix using only default *Reaper* [3] audio effects and submit project files at the end of each round of mixing. During mixes 3 and 4, feedback was gathered through a survey, anonymized, and emailed to the appropriate AE. Participants additionally completed multiple surveys, as well as pre- and post-experiment interviews. Notably, ELs were asked to rate their enjoyment of a subset of these mixes using a re-purposing of the CI-MuSHRA (Multiple Stimuli with Hidden Reference and Anchor for CI users) framework [6].

A thematic analysis of EL feedback resulted in the following main themes: Subjective Preferences & Affective Perceptions, Sonic & Musical Characteristics, and Separation of Sonic Elements. In addition to the myriad of listening strategies and situational contexts that affect general music listening, CI users additionally noted the impact of their hearing assistive technologies (CIs and hearing aids) on not only the sonic qualities of the music but their mental evaluation process as well. This variety of themes illustrates the breadth of affective, technological, and aesthetic factors that impact CI users' music listening experience.

The results of this initial study will inform the creation of grouped audio effects, or meta-parameters, identified as enjoyable by ELs. These meta-parameters will be implemented in a web-based, publicly available tool developed using the Web Audio API and the Web Audio Module (WAM) plug-in standard. Once initial meta-parameters have been established, the research team will re-engage ELs using a participatory design framework to adjust meta-parameters and provide user experience feedback. Initially, the tool will only contain songs with publicly available 4-track stems.

Given the complexity of factors impacting the enjoyment of music listening among CI users, this project's preliminary results suggest that purely algorithmic approaches to address common complaints among CI users should be approached with sufficient context. Deep learning techniques have the capacity to be immensely helpful. Yet if applied incorrectly, such tools run the risk of creating and re-enforcing norms around music listening and appreciation among CI users that may significantly detract from user enjoyment. While well-defined research targets such as speech-in-noise scores and generic measures of audio quality may help address specific questions, consideration of affective responses such as listener *enjoyment* and desirability for repeated listening provide a more relevant means of guiding creative processes and encouraging more holistic design and implementation strategies. Without careful consideration of the diversity of hearing abilities, listening strategies, aesthetic preferences, and situational and cultural contexts, new technologies seeking to empower users to experience music in new ways may erase listener agency and stifle enjoyment.

The breadth of digital tools available that could be used to personalize, augment, and translate musical experiences for D/deaf and Hard of Hearing users shows great potential. The recent wave of accessibility research and customization options in the video game industry, from audio element volume control to tinnitus-sensitive audio censoring, illustrates the potential of customizable digital experiences to actively include and center the joy of D/deaf and Disabled users, as well as other users who might not identify with these groups but who also benefit from such settings. As new tools and interfaces for music creation and listening are crafted, there is an opportunity to shift accessibility concerns earlier into the design phase by engaging with D/deaf and Disabled communities from the beginning. This process can serve to shift the accessibility design framework beyond that of legal compliance to systems that actively value D/deaf and Disabled joy.

[3] A popular Digital Audio Workstation (DAW), available at: https://www.reaper.fm/

Acknowledgments

The research team would like to thank the Grammy Foundation for their financial support, Nette Worthey for their administrative support, and a special thank you to our participants for their time and thoughtful contributions.

References

[1] S. Spangmose, J. Hjortkjær, J. Marozeau, Perception of musical tension in cochlear implant listeners, Frontiers in neuroscience 13 (2019) 987.

[2] M. Maarefvand, J. Marozeau, P. J. Blamey, A cochlear implant user with exceptional musical hearing ability, International journal of audiology 52 (2013) 424–432.

[3] C. J. Limb, A. T. Roy, Technological, biological, and acoustical constraints to music perception in cochlear implant users, Hearing research 308 (2014) 13–26.

[4] W. Nogueira, A. Nagathil, R. Martin, Making music more accessible for cochlear implant listeners: recent developments, IEEE Signal Processing Magazine 36 (2018) 115–127.

[5] S. Burgstahler, T. Doe, et al., Disability-related simulations: If, when, and how to use them in professional development, Review of Disability Studies: An International Journal 1 (2004).

[6] A. T. Roy, P. Jiradejvong, C. Carver, C. J. Limb, Assessment of sound quality perception in cochlear implant users during music listening, Otology & Neurotology 33 (2012) 319–327.