The Musical Touch

Exploring Vibrotactile Augmentation of Music for Cochlear Implant Users

Paisa, Razvan

DOI (link to publication from Publisher):
10.54337/aau708324153

Publication date:
2024

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Paisa, R. (2024). *The Musical Touch: Exploring Vibrotactile Augmentation of Music for Cochlear Implant Users*. Aalborg University Open Publishing. https://doi.org/10.54337/aau708324153

THE MUSICAL TOUCH

EXPLORING VIBROTACTILE AUGMENTATION OF MUSIC FOR COCHLEAR IMPLANT USERS

BY
RAZVAN PAISA

PhD Thesis 2024

AALBORG UNIVERSITY
DENMARK

The Musical Touch

Exploring Vibrotactile Augmentation of Music for Cochlear Implant Users

PhD Thesis 2024

Razvan Paisa

Aalborg University Copenhagen
Department of Architecture, Design and Media Technology
A.C. Meyers Vænge 15
DK-2450 København SV

Submitted: January 2024

Main Supervisor: Professor Stefania Serafin
Aalborg University

Co-supervisor: Associate Professor Niels Christian Nilsson
Aalborg University

Assessment: Associate Professor Luca Simeone (chair)
Aalborg University, Denmark

Research Associate, PhD Hanna Järveläinen
Zurich University of Arts, Switzerland

Professor Deniz Başkent
University of Groningen, The Netherlands

PhD Series: Technical Faculty of IT and Design, Aalborg University

Department: Department of Architecture, Design
and Media Technology

ISSN: 2446-1628
ISBN: 978-87-94563-17-8

Published by:
Aalborg University Open Publishing
Kroghstræde 1-3
DK – 9220 Aalborg Øst
aauopen@aau.dk

Printed in Denmark by Stibo Complete, 2024

Abstract

Music listening is a quintessential aspect of human life; it has been a constant companion of humanity throughout the ages, for as long as we can find evidence of people. Nowadays, the scientific community agrees that music has numerous benefits for the individual well-being, mental and even physical health, by encouraging social relationships, emotional healing, memory recollection, as well as it's intrinsic value — music is good. A big obstacle in accessing music is hearing impairment that many individuals must live with. In the case of restoring severe hearing loss, the solution is usually found in a cochlear implant - a neuroprosthetic device that restores hearing by stimulating the inner ear directly. While these implants show fantastic results for speech related activities, indicated by the fact that most cochlear implant users live their daily life as normal hearing individuals, the music perception is severely degraded. As this segment of the population will only increase in the next 25 years, it is imperative to research new and innovative solutions that will work in cooperation with the inevitable advancement in hearing assistive devices.

One such opportunity can be found in using multiple sensory channels while ensuring the rules for multisensory integration are upheld. This idea is as old as the cochlear implant itself, as the creator of the first commercially available cochlear implant designed a tactile device called the "Tickle Talker", that was supposed to aid in the perception of sound by the hearing impaired user. This is the legacy this project continues, as advancements in the tactile technology afford designing significantly better devices, the research needs to explore and propose the best usage of those novel tools.

This dissertation is an interdisciplinary project conducted in collaboration with two institutions that have been immensely supportive of the work - *The Royal Danish Academy of Music* as well as the *Copenhagen Center for Hearing and Balance* from Rigshospitalet. Through an applied research project, we

explored the possibilities of using vibrotactile stimulation designed around the needs of cochlear implant users in order to improve their music hearing performance and experience.

The primary contribution of this doctoral research is in the eight papers presented in part II, as well as the discussion presented in part I. The first paper in this series presents a scoping review of vibrotactile devices applicable to music, underscoring recurring themes and gaps within existing literature, and emphasizing the lack of standardization in this field. The subsequent two papers delve into the study of musical perceptual features through the use of vibrotactile displays, while also addressing the constraints inherent in single-actuator devices. The remaining five articles explore different facets of music listening experiences for cochlear implant users in social settings, particularly in concert environments. Among these, some studies are dedicated to the development of tactile displays specifically designed for concert use, which I refer to as 'concert furniture' while others focus on assessing and training the auditory performance of CI users in live music scenarios. These studies collectively aim to provide a deeper understanding of how cochlear implant users interact with and perceive music in concert scenarios and how vibrotactile technology can enhance these experiences.

Resumé

Musiklytning er en væsentlig del af det menneskelige liv; den har været et konstant følgeskab for menneskeheden gennem tiderne, så længe vi kan finde beviser for menneskers eksistens. I dag er der enighedden videnskabelige enighed om, at musik har talrige fordele for individets velvære, mentale og endda fysiske sundhed, ved at fremme sociale relationer, følelsesmæssig heling, erindring, samt dens iboende værdi - musik er god. En stor hindring for adgang til musik er hørenedsættelse, som mange individer må leve med. I tilfælde af alvorlig høretab, findes løsningen ofte i et cochleaimplantat - en neuroprotese, der gendanner hørelsen ved at stimulere det indre øre direkte. Selvom disse implantater viser fantastiske resultater for taleaktiviteter, og de fleste brugere af cochleaimplantater lever deres daglige liv som normalthørende, er musikopfattelsen betydeligt forringet. Da denne del af befolkningen kun vil vokse i de næste 25 år, er det afgørende at forske i nye og innovative løsninger, der vil fungere i samarbejde med de uundgåelige fremskridt inden for høreassisterende enheder.

En sådan løsning kan præsenteres ved at bruge flere sensoriske kanaler, samtidig med at reglerne for multisensorisk integration opretholdes. Denne idé er lige så gammel som cochleaimplantatet selv, idet skaberen af det første kommercielt tilgængelige cochleaimplantat designede en taktil enhed kaldet Tickle Talker, som skulle hjælpe med opfattelsen af lyd for den hørehæmmede bruger. Dette er arven , som dette projekt fortsætter, da fremskridt inden for taktil teknologi gør det muligt at designe betydeligt bedre enheder, og forskningen skal udforske og foreslå den bedste anvendelse af disse værktøjer.

Denne afhandling er et tværfagligt projekt udført i samarbejde med to institutioner, der har været en enorm støtte for arbejdet - Det Kongelige Danske Musikkonservatorium såvel som Københavns Center for Hørelse og Balance fra Rigshospitalet. Gennem et anvendt forskningsprojekt undersøgte vi mulighederne for at bruge vibrotaktil stimulation designet omkring behovene

hos brugere af cochleaimplantater for at forbedre deres musiklyttepræstation og oplevelse.

Det primære bidrag fra denne ph.d.-forskning ligger i de otte artikler præsenteret i del II, samt diskussionen præsenteret i del I. Den første artikel i denne serie præsenterer en afgrænsende gennemgang af vibrotaktile enheder anvendelige til musik, og understreger tilbagevendende temaer og huller i den eksisterende litteratur, og fremhæver manglen på standardisering på dette felt. De efterfølgende to artikler dykker ned i studiet af musikalske opfattelsesegenskaber gennem brugen af vibrotaktile display, samtidig med at de adresserer "begrænsningerne ved enheder med enkelte aktuatorer". De resterende fem artikler udforsker forskellige facetter af musiklytteoplevelser for brugere af cochleaimplantater i sociale indstillinger, især i koncertmiljøer. Blandt disse er nogle studier dedikeret til udviklingen af taktile display specifikt designet til koncertbrug, som jeg refererer til som koncertmøbler, mens andre fokuserer på at vurdere og træne den hørelsesmæssige præstation af CI-brugere i live musikscenarier. Disse studier sigter kollektivt mod at give en dybere forståelse af, hvordan brugere af cochleaimplantater interagerer med og opfatter musik i koncertscenarier, og hvordan vibrotaktil teknologi kan forbedre disse oplevelser.

Contents

I Introduction 1

Thesis Details

Thesis title:	The Musical Touch : Exploring Vibrotactile Augmentation of Music for Cochlear Implant Users
Ph.D. Student:	Razvan Paisa
Supervisors:	Prof. Stefania Serafin, Aalborg University Asst. Prof. Niels Christian Nilsson, Aalborg University

The main body of this thesis consist of the following papers.

[A] Tactile displays for auditory augmentation - a scoping review and reflections on music applications for hearing impaired users

[B] The Relationship between Frequency and Hand Region Actuated

[C] A Comparison of Audio-to-Tactile Conversion Algorithms for Melody Recognition

[D] A Real-Time Cochlear Implant Simulator - Design and Evaluation

[E] Multisensory Integration Design in Music for Cochlear Implant Users

[F] Design and Evaluation of a Multisensory Concert for Cochlear Implant Users

[**??**] A Concert-Based Study on Melodic Contour Identification Among Varied Hearing Profiles

In addition to the main papers, the following publications have also been made during the PhD.

- Get a Variable Grip: A Comparison of Three Gripping Techniques for Controller-Based Virtual Reality, Pedersen, Astrid Krogh, et al. 2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). IEEE, 2023.

This thesis has been submitted for assessment in partial fulfilment of the PhD degree. The thesis is based on the submitted or published scientific papers listed above. Parts of the papers are used directly or indirectly in the extended summary of the thesis. As part of the assessment, co-author statements have been made available to the assessment committee and are also available at the Faculty.

Preface

Preface

Seven years ago to the day, I composed a techno track called 'Rants,' featuring a robotic voice that rambles about how *"music is good"* and that it *"encourages you to feel grief, sadness, anger, humbleness, achievement, pride, love, lust, etc..."*. Little did I know that in the following years, I would dedicate a significant portion of my time to making music accessible through technology for exactly this reason — to feel. I accidentally re-discovered this track while writing the abstract, realizing that the phrase *"music is good"* has been a recurring theme in my life for several years now. Music, and particularly sharing musical experiences, is something very close to my heart, and this project is deeply rooted in an honest and empathetic attempt to help people enjoy it. This worldview has served as a lighthouse, helping me stay focused on what matters most for cochlear implant users, who are the ultimate beneficiaries of my work.

Whether I succeeded or not remains to be seen, but I am confident that my contribution is not redundant and I feel that together with my collaborators we have started the discussion about concert music and cochlear implants — a discussion worth having, I hope you agree. Through this dissertation, I will walk you to to this discussion, but first I will introduce the project, my motivation and the stakeholders in the first section, followed by the *Background* where I will present a brief overview of the topics relevant to the understanding of the work. If the reader is well versed on the subjects like well-being and music, as well as vibrotactile displays for music listening and the underlying principles, then they can skip to section three, which describes the research questions and objectives, followed by a brief summary of the articles published as part of this doctoral research and a summary, concluding chapter one. Chapter two contains the eight articles that have been submitted or published as part of this project in their entire length. For the sake of consistency, the formatting of these articles has been converged to be consisted with the rest of the dissertation.

Acknowledgments

It is difficult to envision myself undertaking and completing a PhD program, yet here I am at the finish line. For this achievement, I am profoundly indebted to my supervisor, Stefania Serafin who somehow foresaw this moment, and has been the constant wind in my sails. For all she did for me — supervision, encouragement, guidance, or pouring an overly generous glass of Prosecco, I am forever grateful, but the thing that I am thankful for the most is for believing in me, a lot more than I did. I am truly privileged to have worked with her and she will forever be a role model. I also wish to express my gratitude towards Niels Christian Nilsson *(Nilšović)* for co-supervising my project, but also for being an extraordinary sparing partner — I have learned so much from him and he is an inspiration.

This was not a one-man project, far from it. There have been many people involved that I would like to thank for sharing their time and energy with me. Jesper Andersen is the one who jump-started the idea of using tactile displays for CI users, and for that alone I would like to thank him. Fortunately this evolved into a collaboration that I am most proud of, and I think this project wouldn't be half as good if it was not for the collaboration with Jesper. Likewise I would like to express a heartful thanks to Lone Marianne Percy-Smith for her advisory role, the numerous opportunities she presented as well as constantly advocating for the best of the users. I need more people in my life like Lone who consistently puts genuine effort for the good of others.

A special thanks to my co-authors and collaborators. Francesco Ganis has been a fantastic discussion partner and an even better drummer for all the live music needed throughout our studies. Doga Cavdir has challenged me to approach research from a different perspective, and I will always treasure her encouragement to focus on applied research, and *the people*. Peter Williams is another frequent collaborator who has been a great help into designing and executing most of the tactile displays used, and if that was not enough, he beautifully played the bass in all the live music based studies. Similarly I would like to thank Kirsten Louise and Antonia Barišić for their vocal performances in the concerts, and most importantly, for how effortlessly they adapted their singing style to our specifications. I am also grateful to Charlotte Thostrup and Mathilde Lumbye Orry from *SCR Kommunikation* for providing their invaluable input into integrating music in the life of CI users - they are truly selfless experts, and I have been lucky enough to collaborate with them.

I sincerely thank all my friends at ME-Lab, current and past. Without knowing you have contributed to my project in one way or another. I specif-

ically would like to thank to Lui Thomsen for being a supportive and caring friend that has help edit this document, as well as being an inspiration in perseverance and humbleness. Emil Rosenlund Høeg deserves a huge thanks for helping me get started with the PhD document and taking precious time from his family to answer my (probably silly) questions. Silvin Willemsen has been an unspoken benchmark of professionalism for me, and I am grateful to have work with him — I just hope I will (magically) become half of the programmer he is today. Ali Adjorlu has help me ground myself, always sharing his sincere opinions when requested, while being a real hub of fun. Similarly, i'd like to thanks Jon Ram Bruun Pedersen for showing me how diplomacy is applied without compromising ones values, as well as willingness to help selflessly, and Cumhur Erkut for being excited about the bleeding edge of sound processing, inspiring all of us. My life in the lab would not be the same without you around, and I hope I will never have to witness that.

The list of people who have supported me on this journey is probably longer, but the most exceptional role model I've had, who was also my initial one, has been my mother. I say this not only because she cared for me and motivated me to follow my passions, but also because she taught me how to navigate the challenges of the world, and for this, I am forever indebted to her.

I would like to end thanking to my amazing wife Lise for her unconditional love and support. You have been my biggest cheerleader and have helped me the most in actually completing this project. There is no way I could have finished without you, and I can't wait to finally have some time to relax and enjoy with the family. I would also like to thank Viggo, my 4 year old son for constantly smiling, and for understanding that *"tati has to work now"*; your infectious love of life is a privilege to witness. Thank you Viggo, thank you Lise, I love you.

Razvan

Part I

Introduction

CHAPTER 1

Introduction

Don't look for a great idea. Look for a ***good problem*** [147]...

1.1 Genesis

...this is what started my project. After 4 years of working as research assistant in the ME-LAB I have been part of several grant applications for a PhD position, getting very close to success a few times. What I was doing did not work — it was difficult to convince grant evaluators that replicating culturally relevant (but otherwise obscure) musical instruments using 21st century technology was a fantastic idea. Why couldn't they see how interesting it would be for the Danish Music Museum to have a modern *Ondes Martenot* or a *Nyckelharpa* for their visitors to play, and learn about? I could have built the coolest digital musical instruments.

My background revolves around programming, user experience, sonic interaction and *building stuff*. Throughout my bachelor in Medialogy I focused on human-computer interaction, programming and user experience evaluation. It was in 5th semester when I interacted with the museum world for the first time not as a visitor, but as a creator; together with my group we set to design an augmented reality interaction around the *Egtvedpige*. She was a mysterious character in the Danish history — supposedly a dancer for the gods, and her well preserved remains are found in a glass cabinet at the Danish Museum. We *brought her to life* through technology, and visitors seemed to really like it — this was way before the current VR craze; QR codes were a novelty back then.I graduated my bachelor with a project focusing on directional sound reproduction for car drivers, and that led me to studying Sound and Music Computing (SMC) as my masters degree.

Little did I knew that it would change my life, more through the people I met than the curriculum. During my masters I started designing and evaluating various sonic interaction systems; some were build from scratch - like a digital glass harmonica [119, 121], while some revolved on *hacking* existing technologies [66]. Most importantly,I have started to work under the supervision of prof. Stefania Serafin; little did I know then that it would be a long and extraordinary experience. After graduating SMC with a project about wavefield synthesis I got offered the research assistant position I was talking about earlier. It became great opportunity to further explore sonic interaction, by building and evaluating systems; a commonality among the projects I was part of is that everything had a tactile interface. I slowly realized (probably I was the last to know) that I enjoy working with physical artefacts; I like using the press drill, the soldering iron and the oscilloscope just as much as coding weird sounds. As a result I was part of projects revolving around re-creating old instruments [121, 170], or creating new musical human-computer interaction based on re-purposed hardware [40, 120, 125]. The experiences from these projects, along the people I worked with gave me confidence to apply for funding for my on project - *Re-creating extinct instruments using fabrication techniques and digital sound synthesis*. That did not go very well; I thought I have a great idea,I knew how to build instruments, but I did not have a good problem.

At the same time as my research endeavors into adapting medical haptic devices to work as musical instruments, tonnemester Jesper Andersen (now a frequent collaborator) started his quest into exploring music for the hearing impaired, from a musicological point of view that is. Being tasked with some mundane responsibilities regarding the organization of a concert for cochlear implant users, I got closer to the topic, and it was intriguing. I was reading about fairly familiar topics like signal processing for hearing aids, the perceptual system, musical feature extraction, etc. but from a complete new point of view. Even more, I found out about researchers designing tactile interfaces for audio applications — I was already doing that for my driving simulator in 6th semester, but never have I ever thought about the potential of using if for the hearing impaired. I was excited, curious, and oh so naive.

In the meanwhile I was reading *Mapping innovation: A playbook for navigating a disruptive age*, and it was here that Greg Satell was discussing about the issues with focusing on generating business ideas instead of focusing on existing problems, good ones that is [147]. What I understood from it was that a good problem is something that one does not know how to solve, or the solution is not immediately evident, but requires excitement, ambition and perseverance to uncover, and if those elements are ensured, success should follow. According to Satell, I was ready for a new direction, and augmenting the music listening experience of CI users with tactile sensations ticked all the boxes: it was new and exciting, it combined my skills and interests from previous projects, and most importantly it had a clear, incontestably good goal. I knew about most of the elements regarding the topic, but for once, I had no concrete idea about how to proceed — this was my *good problem*.

1.2 The project

The ME-LAB has an impressive record of using technology creatively to solve problems for specific need groups. As part of the *NordicSMC* project's open call, my proposal complemented the lab's direction by concentrating on developing solutions tailored for people with unique and specific requirements.

The main objective for my project was to to explore the possibility of augmenting music with vibrotactile stimulation so that persons with CI implants would find more enjoyment in music. We have focused the scope on social music listening scenarios, motivated by the cumulated social and mental benefits associated with it described in chapter 2.1. Another reason for this has to do with the *applied research* approach the ME-LAB is following in all project and a comment by the vice-president of the CI users' European association that was among the lines: "CI users don't want more devices to take care of at home, they already have one too many" (paraphrased from memory). We wanted to avoid designing devices that people would only use as part of research, for the sake of research. By focusing on social listening scenarios we could design and evaluate more than just tactile displays, but also scenarios and overall experience that extend past the technology involved. Throughout my project, I constantly created a parallel between the existing elevated platforms found in concert venues around the world, specifically designed for the needs of wheelchair users, or the presence of sign language interpreters and the concert furniture we designed. What if concert venues around the would would provide well designed vibrotactile stimulation for the CI users as well? *That would be pretty neat, wouldn't it?*

In other words the focus for my project was on designing and evaluating tightly integrated hardware/software systems and experiences that aimed to improve music perception for CI users. As a result, I developed novel methodologies for multisensory signal processing, especially for multi-actuator devices, that account for the physiological and perceptual characteristics of the touch and hearing. The articles presented in the second chapter describe my efforts in detail.

Before continuing to present the stakeholders, I want to briefly make some clarifications about different types of CI users my project is for. There are 2 grand categories of cochlear implanted people, segregated based on the time of implantation, relative to acquiring speech; they are called pre-lingual and post-lingual CI users [71]. This differentiation is made because learning to speak has such a big impact on the long term development of the brain with respect to sound perception, that the expected performances between the two groups are significantly different. I focused only on adult post-lingual CI users, at the recommendation of prof. Lone M. Percy-Smith from the *Copen-*

hagen Hearing and Balance Center (CHBC), who argued that this sub-group would benefit most from vibrotactile augmentation. Certainly, there are subgroups within the categories of CI users, such as early pre-lingually deafened, late-implanted, or post-lingually deaf children. However, given that the current state of vibrotactile augmentation research is still in its exploratory phase and primarily focused on broader aspects, these specific niches will be the subject of future research endeavors.

1.2.1 Stakeholders

Another quintessential aspect of my project is that it is a collaborative effort. Without the many peers, partners and co-authors I would not be writing this document. I was always met with reciprocating enthusiasm, encouraged to explore the topic from my worldview, and trusted to apply the methods I mastered — an attitude shared by all stakeholders.

Applying the multi-level typology for stakeholders identification described by [33], the following list emerges:

1. **Macro-level** The Danish government, through its political direction with respect to hearing impairment and cochlear implantation. The government provides the opportunity for eligible individuals to receive a CI as well as a limited amount of post-surgery auditory-verbal rehabilitation through speech therapy. Furthermore, through regional or local governmental institutions, further rehabilitation programs can be accessed on a voluntary basis; some of these programs do include music listening training for CI users.

2. **Meso-level** This level encompasses several players:

    - NordicSMC is a Nordic University Hub supported by Nordforsk that unites top sound and music computing researchers from across the Nordic region. The network is distinctive in its comprehensive coverage of the sound and music field, spanning from the ″soft″ aspects like arts and humanities to the ″hard″ elements such as natural sciences, all underpinned by a strong technological expertise. NordicSMC does not have a fixed topic of research, therefore it affords great independence and creativity to the projects it supports. The involvement from NordicSMC was mostly through regular meeting and coordinating collaborative research.
    - The CHBC is the centralized contact between the audiologists, speech therapists, technicians, researchers and CI users. It treats children and adults with hearing loss and deafness and it is specialized in cochlear implantation and bone anchored hearing systems. Furthermore, it has a great focus on children age 0-18 that

can manifest as re-screening of infants, electrophysiological measurements, paediatric audiometric measurements, medical examination or genetic examination. Besides the clinical actives, CHBC acts as a leading research hub in hearing impairment by including 3 universities in their academic activities. A big part of the interaction between my research and the CI users was mediated through CHBC; furthermore I have received great advisory support from technicians, clinicians as well as doctors and researchers from the center, co-authoring several articles, for which I am extremely grateful.

- The Royal Danish Academy of Music (RDAM) is the leading music research institution in Denmark consisting of artistic practice (composition and musical performance), development activities (artistic research and pedagogical development), and research[1]. It has six excellent concert facilities that are used to (among others) to host over 200 free concerts every year, as part of their musical exercises and research activities. The development of my research revolved around several concerts organized for CI users that would have not been possible without the expertise and infrastructure found at RDAM.
- *Specialcenter Roskilde Kommunikation* (SRK) is one of the local institutions dealing with audio-verbal therapy as well as music therapy for CI users. During the COVID-19 pandemic lockdown, the speech and music therapist employed at SRK published a manual describing in detail their methodology regarding music listening as audio rehabilitation for CI users[2]. This manual has been provided to all communication centers across the country, and forms the foundation for their practices (where resources are available). My research has been an open dialogue with SRK, as they have first-hand experience teaching music for CI users, thus the training methodology requirements described in paper H have been outlined together with a music therapist from this center, but their influence on my project can be seen everywhere.
- Oticon Medial - limited time through an international project called *Audio-Haptic Research Group*. The group was composed of researchers from Denmark, Iceland, and the UK, that are working on the state of the art in tactile augmentation, with a focus on the hearing-impaired population. As Oticon Medical could no longer continue its backing due to a buyout by Cochlear Limited in Spring 2022, the group slowly stopped its official collaborative activities. Being part of this group allowed for easy coordination of international research, promoting collaboration over competition, as well

---

[1] https://www.dkdm.dk/

[2] https://scrkommunikation.roskilde.dk/da-dk/til-fagpersoner/mucik-pa-ny/

as providing an expert panel offering constructive feedback and brainstorming solutions to eventual challenges.

3. **Micro-level** The CI users, and to some extend the therapists that would interact with them as part of the auditory training and rehabilitation procedure (mostly in relation to the 3rd problem area described in section 3.3)

1.3 Research approach

The *Medialogy* and *Sound and Music Computing* programs are fundamentally specializations in different fields of Human-Computer Interaction (HCI). Most professors at *Aalborg University Copenhagen*, who have contributed to my education, are active researchers in their areas of expertise, adopting a multidisciplinary 'user-first' approach. This approach inevitably influenced my worldview, and by extension my research, therefore my project focused heavily on the user experience and ecological validity. I believe that lab based and ecological studies show different sides of the same truth, similar to how the quantitative-qualitative dichotomy works. Since the research field is still in its exploratory phase, in my opinion both approaches are crucial to advance the knowledge base fast and accurate, while minimizing blind spots. As a result, my research could be placed mostly in the *applied science* category, intertwined with the experimental nature of basic science, incorporating *lab work* and studies.

CHAPTER 2

Background

This chapter reviews the literature that became the foundation of my research. It aims to provide a comprehensive understanding of the topics discussed, enabling critical analysis of the articles published as part of my project. With an analytical approach, I intend to contextualize the theoretical underpinnings relevant to the project's domain and summarize findings from pertinent studies. The chapter begins by introducing the backbone of the project: well-being, with a particular focus on music listening. This is followed by an presentation of music listening experiences for individuals with cochlear implants, briefly explaining what a CI is, how it functions, and how it differs from conventional hearing aids. I will then outline the fundamental principle behind vibrotactile stimulation and it's association with music: multisensory integration. The ubsequent chapter will delve into vibrotactile music, especially in the context of audio-tactile augmentation for cochlear implant users.

2.1 Music and well-being

The fundamental goal of my research is to enhance the music listening experience for cochlear implant users. This is not because I possess a *miraculous potion* that I wish to sell at an inflated price, but because I have a deep love for music, and I understand that this sentiment is shared by many. Music is one of only two means of auditory communication, it is universal, and has been consistently present in humanity's life across civilizations [169]. Engaging in various musical activities, such as listening to music, singing, playing instruments both casually and formally, as well as creating music through exploration, composition, and improvisation, whether done alone or in a group, is a widespread practice among many people. While music is inherently a source of enjoyment, its impact extends well beyond mere entertainment. Researchers across various disciplines have been investigating the positive implications of music listening on individuals and communities, arguing for more music exposure. The exposure to music is not encouraged solely based on its intrinsic value, other factors also come into play, that will be discussed in section 2.1.2. This is particularly evident in the context of increasing healthcare expenses, a trend observed in both developed and developing, and the proposition of art-based interventions and solution made by the World Health Organization (WHO) [103].

I believe it is essential to begin this dissertation by defining the objective — some refer to it as *quality-of-life (QoL)*, others as *well-being*, and still others as *flourishing*. Throughout this chapter, I will introduce some work that forms a fundamental backbone of my research. They focus on the benefits of music listening, with a special emphasis on the older population and the cochlear implanted one, which represents the group of my thesis.

2.1.1 Well-being

Unlike health, which is generally accepted in terms of objectively measured metrics [104, 132], well-being encapsulates both the subjective and objective perspective. Sometimes well-being is used as a relative synonym to QoL, but some researchers and evaluation methods differentiate between the two. Nevertheless, the objective part of well-being is based on quantitative measurements of income, literacy rates, life expectancy, health, etc. [24]. On the other hand, subjective well-being looks at aspects related to one’s satisfaction with life itself, work or social relationships, as well as the amount of positive emotions and moods compared to negative ones. In other words, the subjective well-being is determined by the individual perception and outlook of the self [103].

Another approach towards defining the well-being comes from looking at a different sum of elements, referred as *PERMA*: *positive emotion, engagement, relationships, meaning, and accomplishment*, first described by Forgeard [41] and further described by Seligman [150]. Their claim is that *PERMA* constitutes the elements of well-being and it is just a different interpretation of the same state, defined in the paragraph above. An interesting finding was published by Goodman [65], who discovered that the elements of *PERMA* correlate moderately high (range .37 to .79, mean = .61) with each other, indicating that one aspect of well-being cannot exist without the others. This is an important element to keep in mind when discussing about improving well-being, regardless of its definition.

Now that well-being and its approximate synonym QoL have been outlined, Huppert and coauthors [76] present a framework that associates well-being with positive mental health. Their approach was to look at the clinical descriptors for depression and anxiety and find antonyms for those features, that will add up to what they called *Flourishing*. The chosen 10 features integrate both the emotional and functional elements: *positive emotions, engagement, positive relationships, meaning, competence, emotional stability, optimism, resilience, vitality and self-esteem*. If it might seem that the definitions keep getting more extensive and complicated, it is because it would be naive to believe otherwise — a great deal would be lost by measuring singular descriptors, although it's common for researchers to simplify their measurements due to external factors or pressure.

2.1.2 Well-being and music

One particular group of professionals work with health and well-being throughout their practices: music therapists. They generally focus on the therapeutic process, with its goals and mechanisms well defined in a clinical environment. If we are to look outside the clinical world, a review conducted in 2021 highlighted that social music listening promotes social connection and regulates mood in older adults [26]. Furthermore, simple music exposure can improve posture, movement and well-being of people with dementia, as well as improve cognitive health of older adults [26]. The same article goes even further and suggests that not only mental problems but physical ones like lung disease or stroke can be ameliorated by listening to music. It almost seems like the music itself is the *miraculous potion*. The authors do highlight that research in music and well-being is often prone to risk of bias related to sampling errors (low population, convenience sampling, etc.), short term evaluations as well as a potential limited independence throughout the assessment. While the authors of [26] do not explain how musical activities affect well-being, a further analysis of the literature available can shed some light on the mechanisms.

For example, an Australian study conducted by Chin and Rickard [21] investigates the impact of purposeful music engagement on mental health and well-being, emphasizing that the benefits of music use for cognitive and emotional regulation extend beyond mere affective styles. The study was conducted on 565 adults, with an average age of 24 years, and concludes that positive mental health outcomes are predominantly due to cognitive reappraisal strategies rather than suppressive emotional regulation strategies [21]. The findings suggest that intentional music use, independent of an individual's affect, consistently predicts positive mental health outcomes through cognitive reappraisal. Furthermore, the authors suggest that the beneficial effects are uniquely attributed to music use and not just a disposition to experience positive emotions. Conversely, music engagement coupled with expressive suppression is likely to predict poorer well-being. The study reaffirms that the method of music engagement is more critical for well-being than just exposure to music, highlighting music's role in enhancing life quality by facilitating adaptive emotional regulation, alleviating mental distress, and fostering flourishing, even considering individual differences in affective dispositions [21].

Another study conducted in Australia explores the profound and multifaceted impact of music on individuals, particularly focusing on older adults [70]. For many participants, the experience of music is deeply personal and serves as a social symbol, providing a means of understanding emotions, self, and spirituality. It's closely tied to their sense of identity and how they communicate feelings, ultimately using music as a medium to enhance well-being [70]. Being a qualitative study, it is full of touching descriptions of how the participants relate to music, in similar ways to me, and probably you. For example, two of them shared how music allowed for emotional sharing without words, intensifying joy and enabling a deeper connection with others, or how music creates *"a direct link back to your whole life"* [70]. Some participants also revealed music as a form of *self-therapy*, helping them maintain balance and feel *"whole"*, *"in tune"*, and *"competent"*. As one participant put it, *"music comes into your whole being and if you happen to be down... it uplifts you."*. Nevertheless, the study emphasizes that music is not just a therapeutic tool but a significant, symbolic medium for promoting wellness and offering a way for people to explore and express themselves [70]. Recognizing that music's role is complex, the study calls for further research to understand how different types of music influence various groups and how it can serve both positive and negative roles in people's lives, similarly to the authors of [109].

Similar results were obtained by a study surveying the Swedish elderly population. It reveals that older adults value music highly as a leisure activity, often employing various listening strategies to satisfy basic psychological needs, such as emotional functions (pleasure, mood regulation, relaxation) and issues of identity, belonging, and agency [90]. The researchers suggests that these strategies are associated with both affective and eudaimonic well-being, potentially leading to increased personal growth and positive affect. While personality and health status emerged as significant well-being predictors, the contribution of listening strategies, though smaller, is important due to the potential for actual well-being improvement. The study advocates for the inclusion of musical engagement in future research on leisure activities aimed at increasing well-being, emphasizing that the motives behind engaging in these activities are crucial [90]. Additionally, it suggests that specific listening strategies, notably *"mood regulation"* and *"identity and agency"*, are consistently positively associated with well-being. And similarly to the two Australians articles mentioned before, this study calls for caution in interpreting causality due to its cross-sectional design, but it underscores the significant role music plays in the lives of older adults and its potential to facilitate successful aging and overall well-being [90].

On the other side of the world, a group of Chinese researchers involved 66 participants aged 65 to 90 in a study examining the effects of music on QoL among older adults [91]. Most participants reported listening to music regularly, with Chinese music being the most popular choice. The study found that listening to music significantly improved QoL scores in the music group compared to controls, especially by week 4, with improvements noted in physical functioning, role limitations, bodily pain, general health, vitality, social functioning, emotional role, and mental health [91]. Although the study had limitations like a small sample size and potential Hawthorne effect [4] (due to the social norms in Chinese society), the findings suggest that music interventions may enhance well-being among older adults by inducing relaxation and distraction responses. The results align with previous research indicating that music can positively affect mental and physical health, supporting the hypothesis that music can significantly improve the quality of life [91, 156].

These are just a few studies that look into the correlation between music and well-being, and the consensus is that music does improve the QoL, but the mechanisms through which this is achieved are still unknown. Music is complex, and so is well-being, but if we look at it from the perspective presented by Goodman et al. in [65] that claims that one aspect of well-being cannot exist without the others, maybe it does not matter so much. Maybe it is OK to accept that the engineering, or clinical approach of *targeted intervention* does not apply to music and well-being.

2.1.3 Aging population

As probably noticed, most of the studies I discussed so far focus on the implications of music on the well-being of older people, and that is because the majority of post-lingual implanted individuals are part of this category [110]. Furthermore, this group is expected to represent a higher proportion of the population in the following years. Enhanced access to healthcare and education, particularly for women, has led to increased participation in the workforce, resulting in women having fewer children [159]. This shift is a positive stride for gender equality, life expectancy, and managing over-population, however, there is a potential downside. The decrease in the global fertility rate, along with a rise in life expectancy, is also shifting the age demographics towards an older population. According to the United Nation, the world population over 65 will account for 20% of the total by the year 2050, as opposite to the approx. 10% it is now [167]. The World Health Organization (WHO) [1] estimates that currently almost half a billion people suffer from disabling hearing loss, with a predicted 900 million by the year 2050. As the target demographics of this thesis averages to 60 years old, and it is expected that so many more people will be hearing disabled, it means that there is pressing need to ensure that these individuals have access to the best music experience possibly available. Optimal aging for older individuals is closely linked to their capacity to function at their best potential. Therefore, the key to well-being lies in staying physically, cognitively, and socially active for as long as possible [69].

---

[1] https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

2.1.4 Cochlear implants and well-being

When it comes to CI users, the discussion is usually revolving around QoL, rather than well-being. While the assumption of equality between the two terms is still maintained [68], it seems like the more systematic approach towards evaluating the impact of implants on the daily life is preferred, especially given the widespread adoption of the *Nijmegen Cochlear Implant Questionnaire* (NCIQ) [72] which measures health-related QoL. Another reason for this approach could be attributed to the necessity to obtaining quantifiable results that can be correlated with audiometric results — the de-facto measurement of success regarding the CI implantation and rehabilitation.

Regardless of the terminology used, there is mounting evidence that CIs do improve the QoL of recipients, even for the most difficult of the populations — the pre-lingual deafened and late implanted (PL-LI) that generally have poorer objective audiometric results. A recent study conducted by Rizkou et al. [139] suggests that there are significant benefits in *Self-reliance, Well-being and happiness, and Communication*, especially for those with a history of hearing aid use or speech therapy. The most intercorrelated subscales were communication, general functioning, and self-reliance, indicating perceived improvements in education, social interaction, and happiness post-implantation. While some variables like mode of communication and schooling status showed no correlation with subscales, others, like self-sufficiency and self-esteem, saw improvements. The study suggests including QoL measures alongside speech recognition scores in future research to better evaluate the outcomes of delayed CI implantation in the PL-LI population.

Another study that aims to discover the factors influencing the quality of life (QoL) in patients with CIs from Sweden and the USA found four significant predictors: environmental factors, perceived support, chronological age, and attitude towards patients' hearing difficulties [68]. Notably, increasing age was associated with better QoL, suggesting that younger individuals face more challenges affecting their psychological well-being due to greater social demands. Interestingly, the duration of deafness before implantation didn't significantly impact post-implantation QoL, contrary to some previous studies [97]. While the Psychological General Well-being Index(PGWBI) indicated lower scores for CI patients compared to the general population, qualitative studies emphasize considerable personal and social gains post-implantation [68]. This study underscores the importance of non-audiological factors like attitudes, support, and social participation in improving QoL for CI patients.

Several years later, a study from Norway used PGWBI to compare the well-being between CI users and a matched subgroup of the general population, revealing no significant difference between the two groups [137]. This aligns with some previous research but contradicts others, like the one described above [68] or [73], who reported lower psychological well-being among cochlear implant users. Notably, the only significant differences observed between the cochlear implant users and the general population were in the dimensions of general health and vitality, potentially due to pre-operative health selection of cochlear implant candidates [137]. Despite these minor differences, the overall psychological well-being was not substantially different. The study's design does not allow for causal conclusions about cochlear implants' impact on psychological well-being, as it was not a pre-and-post-implant comparative investigation. However, participants in an associated qualitative study reported improvements in psychological well-being due to the cochlear implants, suggesting potential positive changes post-implantation [137]. One thing that the authors suggest is that psychological well-being in hearing-impaired subjects is influenced not just by hearing loss but by its consequences on activity limitations and social participation, suggesting a multifaceted approach to understanding QoL improvements in CI patients [68, 137]. Factors like age, environmental attitudes, and support from others also contribute to psychological well-being, as found in other studies [21, 26, 90, 91]. The matching procedure ensured similarity in background variables between the groups, with the main difference being the experience of severe-to-profound hearing loss alleviated by cochlear implants in one group.

Lassaletta and their co-authors [88] published a cornerstone study investigating the music perception and enjoyment in post-lingual CI users and its influence on the QoL, by using a modified questionnaire introduced by [53] and the Glasgow Benefit Inventory (GBI), a generic QoL questionnaire [141]. The GBI scores indicated a positive overall effect of cochlear implantation, particularly in general improvement, with less impact noted in social support and physical wellness [88]. A significant relationship was found between the quality of musical sound through the implant and QoL, with those rating music perception more positively reporting higher scores [88]. A similar conclusion was proposed by Fuller et al. [45] when investigating the relationship between self-reported musical perception and QoL. This suggests that while the ability to enjoy music varies widely among individuals, improving the quality of musical sound may be crucial for music appreciation and, subsequently, for enhancing the QoL of CI users.

Neither the GBI or the NCIQ are very sensitive to music listening when evaluating QoL, therefore it's always slightly difficult to extract the "bigger picture" without relying on additional methods, but there seems to be new assessment tools to help with this. The *Cochlear Implant Quality of Life* [105] survey dedicates one category out of five to auditory entertainment evaluation, while focusing on the QoL. Another one is called *Music-Related Quality of Life* (MuRQoL) questionnaire [30], and it's specifically designed for CI users. Unlike previous questionnaires, which relied solely on expert opinions and basic validation methods, this assessment tool employed CI data to ensure that the final 18 items of the questionnaire were valid, reliable, and able to discriminate among various levels of music perception and engagement among CI users [30]. This approach also minimized item overlap and controlled for floor and ceiling effects. Nevertheless, MuRQoL has not witnessed a widespread adoption given it's infancy, but more researchers are starting to use the questionnaire, and translate it to their native languages.

2.2 Music and Cochlear Implants

Music, with its intricate and beautiful harmonies, presents a challenging signal for CI users to fully appreciate and discern; there are numerous studies that compare various aspect related to music listening between CI users and the general population, or other hearing profile groups [42, 45–47, 58, 154, 158]. Before delving into these studies (and many more), the journey of exploring cochlear implants and music starts with the basics — what is a cochlear implant?

2.2.1 Basics of cochlear implants

Cochlear implants are neuroprosthetic devices that partially restores auditory sensations for individuals with sever to profound hearing loss. The origins of the technology can be traced back to Alessandro Volta's experiments in the 18th century who connected each pole of an battery to his ears via metal probes. Upon activating a switch, he experienced an initial "shock in the head," followed by a noise that resembled "a kind of crackling, jerking, or bubbling as if some dough or thick material was boiling" [2].

First notable implant surgery was conducted in 1957 in Paris by André Djourno and Charles Eyriès [171], and one year later by Roger Maspétiol, that worked with Djuourno as well. Despite initial success, the progression was slowed by interpersonal and ethical conflicts and abandoned soon after [149]. This early attempt demonstrated the potential for direct electrical stimulation of the auditory system, though the device failed after a short period of only a few weeks. This pioneering work was not well-known initially, but it eventually inspired further efforts in the United States and globally.

Subsequent developments in CI technology were significant, with multiple initiatives worldwide in the late 1960s and 1970s, focusing on electrical stimulation of the auditory system using electrodes inserted into the cochlea. Dr. William F. House and Dr. F. Blair Simmons were among the early innovators, each contributing to the evolution of CIs through their work [171]. However, during the early 1980s, skepticism was rife about CIs, with many experts doubting they could do more than provide a basic awareness of environmental sounds and speech cadence. Nevertheless the progress has been monumental since the late 1980s and early 1990s, with modern CI systems now providing high levels of speech reception [171]. This advancement has transformed skepticism into recognition of the potential of CIs, evidenced by the increasing number of successful implantations and the demand for more challenging tests of sentence intelligibility due to high patient performance.

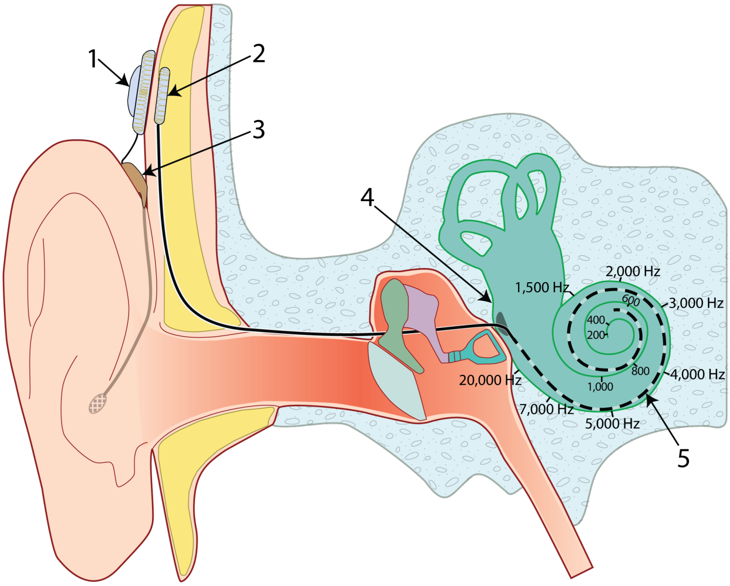


Fig. 2.1: The architecture of a common CI and tonotopy; 1 - External emitting antenna, 2 - Internal receiving antenna, 3 - Control unit with processor and microphone(s), 4 - Cochlea, 5 - Cochlear implant electrodes

Current systems achieve 50% - 60% accuracy when tested on monosyllabic words after 24 months of implantation [171], and 70% - 80% correct sentence recognition in quiet environments [180]. As of 2022, over 1 million people have received CIs, a testament to the system's efficacy and the long journey from initial skepticism to widespread acceptance and success.

Nowadays, most CIs are rather similar in terms of hardware structure and operation; they capture the surrounding sound that is subsequently processed and transmitted to the implant, which in turns stimulates different regions of the cochlea to activate the auditory nerve fibers [175]. There are three main hardware components of a modern CI: 1) a microphone array, 2) battery powered processor; those two are enclosed together and usually sit behind the ear of the user, and 3) the electrode array that is surgically implanted. The number of electrodes in an array is close to the limit of today's technology in terms of number of wires that can fit into the space while still being robust enough to last a lifetime. The electrode array receives the signal form the processor via an radio receiving antenna that sits permanently under the skin and tethers to a matching external transmitting one; the entire system can be seen in Figure 2.1

Similarly, the majority of speech-processing strategies in CIs employ what is known as a vocoder-centric approach or its derivatives [95]. In a typical CI system, the microphone array captures the acoustic scene, and perform some pre-processing to ”clean” the signal of unwanted noise. The assumption is that there is a target signal that needs to be extracted, and the rest is noise that should be eliminated. The pre-processed signal is then divided by a series of band-pass filters. Each filter is tuned to a specific central frequency, with the bandwidths varying systematically across the filters, reflecting the organization of the cochlear basilar membrane. The signals from each filter undergo half-wave rectification and low-pass filtering to isolate the temporal envelope. Subsequently, this envelope modulates a constant-rate pulse train that is delivered to specific electrodes implanted along the cochlea. Electrodes corresponding to lower frequency filters stimulate apical regions of the cochlea to activate nerve fibers attuned to lower frequencies, whereas electrodes for higher frequency filters target more basal regions for high-frequency sounds. Thus, the auditory system is stimulated electrically. This is just the beginning in the effort of restoring hearing. After the surgical procedure is completed, a long rehabilitation process beings which requires users to re-learn how to interpret the new sensation of hearing. Various elements can impact the ultimate auditory outcomes for users of cochlear implants, such as the characteristics of the device, the surgery, the length and timing of the hearing loss pre-implantation (before or after language acquisition), the effectiveness of postoperative recovery, the approaches used in therapy and rehabilitation, and many more [74].

The assessment and eligibility criteria for a person to get a cochlear implant have evolved over time, driven by technological advances and improved patient outcomes, but many aspects are considered: age of the candidate, presence of residual hearing, health of the cochlea and auditory nerve, and, of course various audiometric test batteries [23]. Most common tests rely on speech intelligibility for assessment, and focus on various aspects of spoken language like specifically designed sentences, words, or group of words. One such test is the Minimum Speech Test Battery (MSTB) [145]. Nevertheless, many professionals have identified limitations in current testing systems and called for expansion, suggesting that candidacy for a CI should be considered on an individual basis rather than strictly based on a speech recognition score. This approach is particularly important as early referral and provision of a CI as soon as possible could positively impact the results patients receive with CIs [182].

Eligibility is only half of the path to a CI, as access to the implantation significantly differs worldwide due to various country-specific factors. While developed countries like Australia, Sweden, and the UK have high pediatric CI utilization rates due to universal newborn hearing screening and effective healthcare systems (90% or higher of eligible infants), adult utilization remains remarkably low across the world, including in the United States and European countries (5% - 7% of eligible adults) [157]. This low rate in adults is attributed to inadequate screening for hearing loss, lack of awareness among primary care physicians about CI, and stringent candidacy criteria [157]. In some regions, discrepancies in information provided and influence from the Deaf community also play a role in lower utilization rates [157].

Additionally, financial models and political issues in healthcare systems influence CI testing, with some countries offering more support than others. Despite the potential benefits of CI, many eligible individuals globally remain without access due to these multifaceted challenges.

2.2.2 Music elements and cochlear implants

Cochlear implants have been remarkably successful at restoring speech to the deaf population. Despite advances in CI technology, music perception for CI users is still far from normal hearing individuals, with high-level musical understanding achieved by regular implanted individuals. The primary challenge lies in the complexity of the music that is not transduced appropriately. The issue is the degradation of the musical signal fidelity during the conversion of acoustic sound into electrical impulses, which affects the spectral, temporal, and timbral complexity, as well as the dynamic range of music [12, 27, 80, 129]. Particularly, CI users face severe limitations in pitch representation, crucial for understanding melody and harmony in music. The devices, typically optimized for speech, are limited by the number of electrodes, which cannot match the fine frequency-specific information transmitted by thousands of inner hair cells in a normally functioning ear. This leads to a distorted representation of pitch, and sometimes a gap in the frequency spectrum presented, especially low frequencies, which are essential for the richness and fullness of music [84, 174]. This was believed to be due to the fact that the apex of the cochlea is responsible for handling the very low end of the spectrum, and the general depth insertion of a CI is not enough to reach those — Figure 2.1 shows the approximate distribution of frequencies. The consequence is that CI users often perceive pitches as higher than the real ones, leading to a mismatch between the natural and electrode spectral maps maps, which further distorts the music perception [142]. However, recent studies suggest that there is no benefit in terms of low frequency perception with a deeper insertion of the electrode array [5, 14]. The lack of precise frequency identification leads to a difficulty in discerning sound consonance, typically perceived by individuals with normal hearing as simple ratios between frequencies [158]. Consequently, when considering the intricate elements of music associated with pitch perception, such as melodic motion and contour as well as emotional nuances, these aspects become particularly challenging for CI users to detect, though training in music can enhance their ability to recognize melodic contours [48, 50].

On top of the lack of low frequency presentation as well as the pitch mismatch, the broad stimulation area from individual electrodes results in imprecise electrical fields reaching more nervous fibers than targeted, further diminishing pitch perception [93]. This leads to altered intervals between harmonics. CI users primarily depend on temporal fluctuations in the electrical current to identify the fundamental frequency of sound, effective only up to 300 Hz [158]. Enhancing place-coding, especially for higher frequencies, could improve timbre perception in CI users . However, the limitations in CI encoding lead to altered timbre perception and poorer performance in timbre discrimination tasks, although CI users can utilize specific temporal cues to distinguish between broad instrument categories due to differences in their sound's temporal envelope [158].

Another important aspect of music processing that CI users face challenges with is auditory stream segregation, essential for isolating specific instruments or melodies in complex musical settings [49, 96]. The inability to effectively use fine structure and temporal cues limits the recognition of different timbres and the segregation of simultaneous sound streams, making it difficult for CI users to follow individual instruments or voices in polyphonic music. This difficulty aligns with previous findings indicating CI users' challenges with instrument identification amidst ensembles and speech recognition with competing talkers. The core issue is the weak fundamental frequency (F0) coding and coarse spectral resolution in CIs, which hinders the ability to segregate sounds, even when there are significant F0 differences [49]. Interestingly, some CI users with more musical experience showed a better ability to distinguish polyphonic melodies relying on timbre and F0 differences [49]. This suggests that musical experience might enhance attention to the limited pitch and timbre cues available through CIs. However, due to the limited spectral and temporal resolution of current CIs, even these experienced individuals face significant challenges compared to normal hearing individuals. Until CIs can provide fine structure cues essential for complex pitch perception and timbre discrimination, music experience and targeted training appear to be the most beneficial strategies for improving music perception for CI users.

This diminished ability to discriminate monophonic or polyphonic pitches, results in an impaired perception of harmony and chord structure [100]. In a study by Caldwell et al. [17], ten CI users and twelve normal-hearing (NH) listeners rated 36 piano pieces, which varied in chord types, from very unpleasant to very pleasant. NH listeners preferred consonant triads and found dissonant chords least pleasant. Conversely, CI users rated all chord types similarly, indicating a lack of distinction between consonant and dissonant chords, despite enjoying the music. Knobloch et al. [85] further revealed that CI users could somewhat discern consonance, rating major chords as more consonant than other types. However, they couldn't differentiate between an authentic cadence and a modified one with an altered final tonic.

When it comes to the amplitude of sounds, the dynamic range in CIs is significantly compressed compared to normal hearing, which profoundly affects the perception of music's expressive elements [151, 179]. Furthermore, the dynamic range and resolution of perceived loudness differ significantly between acoustic and electric hearing [60]. Normal-hearing individuals can detect a wide range of sound intensities, up to 120 dB (acoustic), with the capability to distinguish between as many as 200 steps from the quietest to the loudest sounds [60, 116]. In contrast, CI users experience a much narrower intensity range, typically between 10 to 30 dB, and the number of discernible intensity steps is substantially reduced to about 20 [93, 181]. As a result, CI users tend to perceive polyphony as a combined signal — a cacophony of pitches, especially when the frequencies are relatively near, therefore they prefer music with simple melodic lines like pop or country [101, 158]. An additional variable for CIs is the mapping functions aiming to mimic the natural loudness progression of incoming sounds within this limited dynamic range. While the impact of such constricted dynamic range and resolution on the perception of music sound quality among CI users is not fully understood [60], there are studies that show that implanted individuals preferred enhancements to certain elements of complex music (like drums, bass, and vocals) implying that they might seek to simplify the music to better suit their limited dynamic range [16].

One silver lining in the music perception of CI users is that rhythm is generally unaffected, and CI and normal hearing perform approximately equivalent on rhythm perception tasks [158]. However, in complex music where rhythm is conveyed by various instruments, the patterns perceived by a CI user may become overlapping and indistinct. Consequently, it's not surprising that CI users tend to favor simpler, more rhythmically defined music [158].

Another dimension of music listening worth discussing is the access to musical meaning, as this might be more valuable for enjoyment than correct pitch or amplitude perception. A cornerstone study investigated music meaning through event-related potentials (ERPs), specifically the N400 effect [89], which reflects the brain's response to meaning in stimuli [13]. The authors report that the N400 effect indicates that despite reduced input information, CI users can access preformed meaningful representations when listening to music [13]. They activate similar associations to normal hearing individuals, but don't always follow these associations in their behavioral judgments. The study also notes that CI users can access semantic auditory concepts built up during the time of normal hearing, implying cortical plasticity that enables them to understand musical meaning despite degraded hearing [13].

2.2.3 Music listening with cochlear implants

So far I have been describing CIs and the listening experience they provide form a seemingly grim perspective, focusing on the limitations they inherit, but I don't want to insinuate that life with cochlear implants is devoid of music. Despite the objective limitations, many CI users do not reject music, and it's not unheard of some of them to actively search and engage in musical activities. There are even professional musicians that rely on CIs for their hearing, doing wonders to counter the inaccurate (but unfortunately widespread) assumption that there is no music after implantation. Music is still there, but requires substantial intrinsic motivation to re-discover. A study by Gfeller et al. investigated the listening habits and reported that only 23% of CI users indicate that listening to music is less satisfying, but 43% reported that the sound of music is constantly improving, and it's better than no music at all, while still mentioning that it's less pleasant than before, and 23% noted that music sounds at least as good or even better than before hearing loss [53]. Throughout the following sub-chapter, I want to shed some light on how CI users engage with music and their preferences.

The musical experience of CI users is varied and influenced by a multitude of factors, including pre-implantation music involvement, the technological aspects of the implant, cognitive abilities, and individual preferences [53]. For many CI recipients, music plays a significant role in their lives, but the extent and nature of this involvement vary greatly due to individual differences and the complex nature of music perception with a cochlear implant. Prior to implantation, individuals typically engage with music to various degrees, but post-implantation, there is often a reported decline in the time spent listening to music. This decline might be attributed to the differences in sound quality and music perception experienced through the implant. For those who gradually lost their hearing and previously relied on hearing aids, cochlear implants offered a more enriching musical experience, despite not providing a perfect representation of music [53]. Nevertheless the same study concludes that this category of users report a more detailed perception of pitch and timbre compared to what they experienced with hearing aids [53]. Interestingly, the study found no significant correlation between the duration of deafness or implant use and the enjoyment of musical instruments. This suggests that other factors play a role in musical appreciation post-implantation. Cognitive abilities, for instance, were found to correlate with the appreciation of music and the ability to discern different musical instruments [53]. Exposure to music post-implantation could benefit especially post-lingual deafened individuals, that can relate to past musical experiences. Motivated CI users who resumed musical activities soon after implantation adapted to using the device in musical contexts [108]. According to Migrov et al., a significant portion of the participants (30%) stated they would consider undergoing implantation solely for the ability to listen to music, indicating a strong appreciation for music through the CI [108]. Furthermore, listening habits and music appreciation do not seem to be significantly associated with factors like age of implantation, gender, duration of deafness, type of CI device, or speech perception abilities [133]. However, a trend was noted where fewer older individuals continued to listen to music [53, 108]. A notable finding is that all pre-lingual deafened patients in the study conducted by Migrov et al. continued to listen to music, while 32% of post-lingual deafened patients did not [108]. Overall it can be concluded that that beyond the technological aspects of the implant, individual cognitive factors and motives significantly influence the music listening experience [53, 101].

The listening environment also plays a critical role in the enjoyment and perception of music for CI users. Quiet environments, with good acoustic properties and low reverberation are strongly preferred, just as is the case for a comfortable volume. A study investigating the impact of reverberation reported that that minimal reverberation (RT60 = 0.2s) received the highest mean ratings for musical enjoyment, while those with the highest reverberation (RT60 = 10s) received the lowest [20]. It seems that a more reverberating space increases the number of auditory signals the CI must process, exacerbating the spectral smearing effect inherent in CI technology and consequently reducing spectral peak recognition. This reduction might impair music enjoyment in a manner similar to the effects observed in speech recognition under CI conditions.

Additionally, general music enjoyment seems to be positively correlated with the amount of time spent listening to music post-implantation, suggesting that more exposure can enhance the music listening experience for CI users [53]. Besides this, being familiar with the music listened, a simple musical structure, a high quality recording as well as a strong, clear beat also correlates to a positive listening experience [15, 53, 101, 158].

Despite the challenges with CIs, no particular genre of rhythmical music stands as universally preferred or more enjoyable for CI users [53]. Classical music, by being more intricate, is less likable compared to pop and country, likely due to its complex structures or the absence of lyrics and steady beat that could aid in understanding the music [101]. While there is some decline in enjoyment across all genres compared to pre-deafness levels, the drop is relatively modest. This indicates that while the quality and perception of music might change with a cochlear implant, the recipients' diverse musical tastes and preferences continue to play a significant role in their listening habits and enjoyment. What is important to mention is that there is evidence of a lack of correlation between audiological outcomes and CI user satisfaction with music listening [133].

Implants are designed primarily for speech transmission, and they may convey better the simple melodies and rhythms of pop and country music [101]. One way to reduce complexity both objectively and subjectively is to alter the relative amplitude between instrumentation and vocal parts. Attenuating background music and emphasizing clear vocals and drums might make music more enjoyable for CI subjects [15]. This benefit seems to be more prominent for male singers than female ones, possibly due to the difficulty in segregating male vocals from the background music, as indicated by the authors of [52].

These are just some of the studies reporting factors that contribute to musical enjoyment under CI conditions, but the list can be extended to include instrument combinations, rhythmic and acoustic complexity, duration of musical tones, and instrument type, playing style, etc. Understanding the ideal combination of factors will require further research to maximize musical enjoyment for CI users. Nevertheless, there is still a lack of consensus between the different articles discussed above, highlighting the important for more collaboration on further, larger scales research.

The good part is that CI users do not uniformly dislike music, but their enjoyment is significantly influenced by the implant's transmission quality, the timing of hearing impairment, their cognitive processing abilities, and the listening environment. Selecting less complex music, with an emphasis on drums, bass, and vocals, along with a quiet listening environment with low reverberance, might enhance musical enjoyment for CI users. This only highlights the complex, multidimensional nature of understanding music with a cochlear implant.

2.2.4 Training and rehabilitation

As mentioned before, the surgical implantation procedure is just the beginning of the journey towards hearing restoration. Most of the effort is expected as part of the rehabilitation process. With good reason, though, as the biggest predictor of a successful recovery is the amount of time spent on structured, goal-oriented, disciplined training. For example, a study investigating the effects of 6 months of musical ear training on CI users focused on music perception improvements and their relationship to speech perception and emotional prosody recognition [127]. It showed significant improvements in overall music perception compared to the control group, with notable advances in timbre, melodic contour, and rhythm discrimination [127]. The post-training scores in timbre and rhythm discrimination for the treatment group were comparable to those of normal hearing individuals. The training's most substantial impact was on the ability to identify musical instruments. This improvement in timbre discrimination suggests that CI users can learn to identify musical instruments by their timbre, indicating that implants transmit enough spectral information for this task [127]. This is encouraging as it may contribute positively to the aesthetic enjoyment of music listening for CI users. However, the study also highlights that melodic discrimination remains a challenge for CI users, even after training [127]. The results show a broad variability among participants, suggesting that individual preconditions such as the duration of hearing loss and residual hearing, might influence training outcomes [127].

When looking at the effects of training on melodic contour identification (MCI) and familiar melody recognition (FMR), Galvin et al. [48] found that training improved CI users' performance in both tasks, with significant improvement in MCI performance and notable, but lesser improvement in FMR. In their study, the training used an adaptive computer-generated algorithm, increasing difficulty as users progressed, and required users to attend to increasingly difficult contrasts. This approach, demonstrated generalizability to untrained music listening tasks. However, significant inter-subject variability was noted, a common occurrence in CI studies as mentioned in the study discussed in the previous paragraph [127].

Gfeller et al. [57] explored synthetic training using connected melodies and real-world stimuli. This training aimed to promote top-down processing and encouraged trainees to use contextual cues. The study found that formal training significantly improved melody recognition and appraisal for real-world melodies but only showed modest learning for computer-generated ones. This suggests that real-world music, with its unique timbral blends and rhythmic patterns, provided a more suitable stimuli for training.

Regarding timbre training, Gfeller et al. [56] found that systematic training improved CI recipients' recognition and appraisal of musical instruments. The training involved direct instruction on various instruments played in different styles and emphasized characteristic features. The results indicated significant improvement after 12 weeks of training. However, Driscoll et al.1 [29] further explored the timeline and methods of timbre training, finding that significant auditory learning could occur with as little as 3 weeks of instruction. They compared different types of training inputs and found that feedback and direct instruction were more efficient than repeated exposure without feedback [29].

Musical training for CI users also influences music perception and enjoyment [20]. While it doesn't necessarily correlate with enjoyment, those with more musical experience might recognize songs better, but also tend to be more critical of the sound quality through the CI [20, 133].

A similar conclusion was suggested by a study aimed to determine if pitch-timbre training or group music therapy could improve music perception and/or speech perception in CI users, and to identify which method is most effective [47]. The findings indicated that while pitch/timbre training significantly improved music perception, specifically MCI performance, it did not significantly enhance speech perception. On the other hand, group music therapy showed a small but significant improvement in vocal emotion identification, a cross-domain effect, but just like pitch-timbre training, it did not significantly impact word or sentence identification [47]. Neither training method significantly improved speech perception in noise or overall quality of life as measured by the NCIQ, although the music therapy group reported subjective improvements in perception across sessions [47].

The pitch-timbre training group's improvement in MCI performance aligns with previous studies and suggests that specific, targeted training can enhance certain aspects of music perception in CI users [54, 55]. In contrast, the music therapy group, which engaged in a dynamic, multimodal, and social training environment, showed improved vocal emotion identification [47]. This improvement may be attributed to the specific training in emotion identification and the dynamic nature of the training that involved singing, playing instruments, and listening in a social context. The subjective reports of improved perception and enjoyment from the music therapy group further suggest that this type of training might impact broader cognitive and emotional areas, potentially enhancing overall engagement and motivation for continued training [47].

A specific study by Lerousseau et al. [92] extends beyond the immediate benefits of music perception, advocating for the integration of music education into the rehabilitation process for CI users. It suggests that this inclusion is not merely beneficial but transformative, capitalizing on the intricate relationship between music and language processing. The reason for their suggestions are based on the fact that musicians exhibit enhanced pitch discrimination, which is crucial not only for music but also for language, aiding in speech tone and vowel discrimination [92]. This heightened sensitivity to auditory cues can significantly benefit CI users, for whom pitch discrimination is often challenging. Furthermore, the superior timing abilities of musicians, essential for consonant recognition in speech, suggest that rhythmic training could greatly enhance the phonemic awareness of CI users [92, 94]. Importantly, training in timbre perception can enrich the auditory experience of CI users, making music and speech more enjoyable and distinct. Additionally, the ability to extract and understand the hierarchical rhythms in auditory streams is not only crucial for music comprehension but also enhances speech understanding and social interaction [92]. Moreover, the enhanced working memory observed in musicians indicates potential benefits for language processing in CI users. With the positive effects of musical training on speech in noise perception and understanding prosody, integrating such training into the rehabilitation process could substantially improve the communication abilities and overall QoL for CI users.

Advocating for music training as an integral part of the rehabilitation process for CI users is not just about enhancing their musicality, but about holistically improving their auditory, cognitive, and social skills, paving the way for a more effective and enriching auditory rehabilitation journey.

2.2.5 Future perspectives

Cochlear implantation is acomplex process, and it's safe to say that the research community is still exploring creative ways to understand the implications of various factors on the hearing experience. Looking into the future, this is good news, as the understanding of the technology and the complex human perceptual system will only increase. We know that the doctors, the engineers and the therapists will converge on the best solution, but what would that be it's hard to tell at the moment.

There are several aspects that are being explored simultaneously in order to present a system with best possible hearing restoration capabilities. First and foremost it's the continuous, incremental development of the audio processing systems — a constant journey towards the best balance between audio quality, battery life, physical size and cost. While the topic is still mostly researched from a speech perception perspective, whether it's noise reduction, speaker estimators, de-reverberation, etc., we can observe an increase in focus on music processing as well. An important aspect of this research is focused on auditory scene-aware processing and listening modes, and with the outstanding evolution of AI systems, we can expect that processors will be self-adjusting with a high degree of accuracy. I am personally excited about the prospect of a multi-dimensional spectrum of listening modes, where listening environment and acoustic scene are used as input to configure the audio processing parameters in real time.

While processors will get better and smarter, their physical size is not expected to change much (except totally implantable CIs), but one way to expand on their capturing capabilities is to integrate them into the larger hearing system. Many CI users are using bimodal or bilateral systems, and the communication between those will provide new hearing opportunities. Of course, there will be a learning curve for the research community in order to understand the intricacies of multi-microphone directional audio presented on two hearing impaired ears, but at least in terms of hardware and software, the technology is almost ready. When that happens, there will be a considerable jump in sound quality and especially sound localization for CI users through the presence of binaural cues.

With more processing power, better algorithms and multi-directional microphones available for beamforming, the focus will probably expand towards a bespoke, patient-specific hearing experience. So far the research is limited and provides mixed results, as the main approach towards customization is to deactivate subsets of electrodes in regions with bad neural functions [19]. The challenge is to correctly identify the electrodes that need to be de-activated, and there is some skepticism in the research community that this approach will work at all [11]. Nevertheless, the aim for a custom experience past the electrode mapping sessions will remain, and could provide new perspectives facilitated by the advancement in hardware, software and in the academical knowledge.

This approach of incremental advancement of the technology and the experience it provides has produced many QoL benefits for CI users, but the truth is that the progress is slower than expected by the users themselves, as the baseline performance of an average single-side implanted individual has remained largely the same for 30 years [180]. This is not due to a lack of effort, as many prototypes have been evaluated, but few ideas resulted in significant improvement in listening performance. To overcome this roadblock the progress needs to be equally rapid for the transmission system, not only for the capturing and encoding as described so far. Unfortunately the performance obtained with the fundamental design of a small number of electrodes implanted in the scala tympani is plateauing.

One approach with proven results is to preserve the residual hearing in the implanted ear, and it seems that music perception is a strong candidate for witnessing improvements through hybrid stimulation, even tough the acoustic hearing is mostly observed under 300Hz [51, 56, 58, 64]. Some more radical ideas have been looking at re-designing the stimulating interface completely. One such suggestion is to directly implant an array of electrodes in the auditory nerve, bypassing the cochlea completely [107]. This study shows objective benefits of this approach, but so far only animal tests have been conducted. Another direction is suggested by Pinyon et al., [130] who propose the delivery of drugs to the auditory nerve, in order to stimulate it's growth into the cochlear electrodes. An alternative approach all-together is looking into discarding electrical stimulation and substitute it with an optical one [19]. This would imply genetically manipulation of the spiral ganglion so that they become photo-responsive, as reported by Dieter et al. [25]. However, challenges include energy-efficient delivery and the safety concerns associated with very early steps of the technology. Its efficacy may be limited by variable neural survival rate in human CI users, but that is unknown since optogenetic stimulation was mostly tested in recently deafened animals [19]. Further improvement in audio restoration may come from a combination of these techniques, but since most of the novel ideas have not been validated with human patients, we can expect a longer timeline for successful restoration through alternative simulations.

2.3 Multisensory Integration

In everyday life, our senses continuously gather information from the world around us, creating vividly rich experience. Imagine being at a live concert: the band is playing, lights are flashing, the bass is shaking the floor and you feel everything all at once. These sensory inputs don't work in isolation. Instead, our neurons perform an intricate dialogue, seamlessly weaving together these diverse cues to provide a coherent and efficient awareness of our environment. This interaction among the senses and the combination of their independent information stream is called *multisensory integration* [160]. This mechanism, critical for navigating and interacting with our complex world, showcases the brain's remarkable ability to resolve multiple streams of sensory information simultaneously. This chapter will briefly describe the mechanisms of multisensory integration in order to highlight it's potential with respect to CI users' musical listening experience.

Multisensory integration was first described in a cornerstone publication by Stein and Meredith as early as 1990 [160], but the interaction between different sensory modalities have been explored much earlier. The authors state that individual neurons can be receptors of multimodal input and the response to a stimulus from one (e.g., visual) could significantly change when accompanied by a stimulus from another modality (e.g., auditory) [160, 161]. Similarly, cells that didn't respond to individual stimuli from one modality were profoundly affected when these stimuli were combined with others, indicating their multisensory nature. This proved for the first time that the brain uses multiple sensory modalities to facilitate attention and orientation behaviors. More recent studies recognize within-modality and crossmodal integration as fundamentally multisensory operations, arguing for the necessity to re-evaluate studies that researched each type of stimulation independently [59]. One perspective proposed by Ernst and Banks [32] describes how the brain is functioning as a *"Bayesian"* estimator of the environment, and it aims to minimize the uncertainty of sensory estimates by integrating multiple, independent measurements. Multisensory integration is critical in this process as it allows for the combination of complementary cues from different sensory modalities, providing a more accurate representation of the environment. For instance, during conditions like dusk, where visual cues might be unreliable, the brain can combine those with auditory ones to form a clearer perception of reality [3]. Additionally, cues from different modalities often predict each other, allowing the brain to anticipate and adjust to sensory information more effectively [124]. This Bayesian approach views multisensory integration as a means to reduce uncertainty and improve perceptual accuracy [3].

There are several principles that determine how and when different sensory inputs combine to form a unified perception. One such principle is spatial correspondence, suggesting that sensory inputs from different modalities are more likely to be integrated when they converge in the same spatial location [146]. However, the extent to which this principle applies can vary based on the task at hand and the nature of the stimuli [18]. This indicates that while spatial overlap is a significant factor, its influence on integration is not absolute and can be modulated by other contextual factors.

Temporal proximity is another vital principle, suggesting that sensory signals received close in time are more likely to be integrated. However, the importance of timing is not uniform across all types of stimuli and tasks [163]. It is known that stimulation from different modalities reaches the brain at different rates, suggesting that the perceptual system might be rather poorly equipped to deal with simultaneous events in the real world. One such discrepancy exists between auditory (10-30 ms) and visual inputs (55-125 ms); yet, audio-visual integration is still prevalent and robust [160]. The potency of these interactions relies less on the absolute latencies of sensory stimuli and more on the overlap of their influences on the cell [160]. Stimuli typically trigger an extended wave of excitation or inhibition, allowing for a broad "*temporal window*" of multisensory interactions, sometimes exceeding 1500 ms [160]. This prolonged period of multisensory interaction has significant implications for the detection and response to external stimuli. The noted variability in integration timing suggests that the brain's approach to integrating temporally aligned inputs is nuanced and adapted to the specific nature of the sensory information and the current task. Nevertheless, a recent study on audio-tactile speech concluded that the timing aspect is critical, and best results were found when the tactile signal led the auditory one by 100 ms or less [138]. This indicates that more research is needed to fully understand the timing aspect of audio-tactile integration, and some of it will be discussed in chapter 2.4.

The principle of inverse effectiveness is particularly intriguing, as it describes how the strength of multisensory integration inversely relates to the effectiveness of individual sensory reception [162]. In contexts where single sensory cues are weak or ambiguous, multisensory integration tends to have a more significant impact, enhancing the perception of the event or object. Conversely, when one sensory modality provides strong and clear information, the additional benefit of integrating another sensory input diminishes [3]. An example of this principle would be when being traffic and an a police car in mission approaches from the opposing lane; the perception of the police car is based on sight and sound. When it is far away, the computation involved in the multisensory integration results in a supraaditive perception, as the response is stronger than any of the two senses individually. As the police car approaches, the computation changes to additive (equal contribution from both sight and auditory stimuli), and sub-additive, favoring vision when the car is very close [161]. This principle underscores the adaptive nature of multisensory integration, optimizing the brain's response based on the reliability of available sensory information.

Multisensory integration is not a static or uniform process but is highly adaptable and context-dependent. Its principles, such as spatial correspondence, temporal proximity, and inverse effectiveness, interact with a multitude of factors, including the nature of the stimuli, the current task, and the individual's previous experiences and expectations. This dynamic interplay allows the brain to flexibly and efficiently process complex multisensory environments, ensuring a coherent and accurate perception of the world.

2.3.1 Audio-tactile integration

Now that the basic mechanisms and principles of multisensory integration have been introduced, it's important to narrow the focus towards the audio-tactile integration, as this is a fundamental aspect of my project. Research within auditory-tactile interactions has shown that tactile stimulus can influence auditory stimulus and vice-versa [117, 118, 140]. It can therefore be observed that auditory and haptic stimuli are capable of modifying or altering the perception of each other when presented in unison [176]. Before looking at how the interplay between the two senses has been applied, there will be a brief presentation of the involved systems.

The auditory system

The auditory system converts sound waves into neural signals, integrating them with other sensory input to guide behavior. This transformation begins in the external and middle ear, which collect and amplify the sound, transmitting it to the inner ear's cochlea. In there, the auditory signals is decode into simpler sinusoidal components preserving information about frequency, amplitude and phase through a series of bio-mechanical processes [134]. These simple signals are transduced by the hair cells and auditory nerve into electrical activity. The result of this acoustical decomposition process is the structured depiction of sound frequencies spanning the cochlea's length, known as tonotopy [134]. This crucial characteristic remains consistent across the central auditory pathways. The limits of this sensing structure are approximately 20 Hz to 20 kHz, for a healthy (and very young) individual [164]. These numbers are not absolute thresholds that describe the entire population, as most adults loose the ability to hear the higher frequencies proportional to their age and eventual hearing damage. Furthermore, these limits are also amplitude dependant, as extremely loud signals can be perceived even if their frequency exceeds 20 kHz . Nevertheless, they are an elegant generalization of the human capabilities. The connection between frequency and amplitude extends in a non-linear fashion across the entire perceivable spectrum, with highest sensitivity ranging between aprox. 1000 Hz to 5000 kHz [164]. In terms of amplitude, the human ear is capable of detecting sounds over 140 dB, but constant exposure to sounds over 100 dB is know to be harmful — the louder the sound, the shorter the time required for permanent damage to occur. Since the unit of measure for sound is the *Bel(B)* that describes a ration of two values of a root-power quantity on a logarithmic scale, it's important to fix the lowest of the two, to set a reference point that other sounds can be compared to it. In terms of human listening, 0 dB represents the absolute lowest pressure level required for a pure tone between 2-4 kHz to be perceived by a person with excellent hearing, under anechoic listening conditions. This values is $2 * 10^{-5}$ Pa, but the number is not so relevant, what's important is the reference level of 0 dB.

The human ear's sensitivity extends to the quality of the sounds. Pure sounds contain only one frequency, like those produced by a tuning fork, or a sine wave generator. However, in most cases, sounds are intricate combinations of various frequencies, leading to the perceptual characteristic known as timbre. When comparing tones of equal loudness and pitch produced by different instruments, like a violin and a flute, one can distinctly hear the timbral differences. When compared to pitch and amplitude, timbre perception is considerably more complex, combining a multitude of perceptual attributes [102].

The somatosensory system

The second component in audio-tactile integration is the somatosensory system with the fundamental purpose to provide the brain wihth information about the mechanical state of the body. The main source of information are ligaments, tendons, muscles and the skin. Throughout my project I focused exclusively on the skin, however it's worth noting that most of the body's soft tissues are mechanosensitive, but the exact contribution of each sensitive channel to the sensation of touch is not established [122]. The skin comes in three categories with different attributes and functions: mucosal, glabrous(non-hairy) and hairy. The first one covers the internal surfaces of the body and are generally kept moist. In this category is found the tongue, with it's documented impressive sensory capabilities. It can detect object size, shape, very small curvatures or hardness, offering a high enough resolution to be used in video-haptic sensory substitution [122]. The glabrous skin houses different types of specialized receptors but for music perception, two of them prove to be useful: the Meissner corpuscles, also known as Rapid Adapting (RA) receptors which have a very high innervation density and have a limited frequency range of 10Hz-100Hz, with a peak sensitivity around 40Hz and the Pacinian receptors [87]. The latter are larger than the RA ones, have a low spatial resolution, a frequency response between 40 Hz and 1000 Hz, and are most sensitive around 250 Hz. These types of receptors are not distributed evenly among the skin surface, with hands and the mouth area hosting a larger percentage of them, as frequently showed in Penfield's homunculus [126]. The hairy skin does not have such a complex organization, but instead each hair is associated with sensory fibers that innervate an organ called the hair follicle [122].

One particular propriety of the somatosensory system is the ability to recognize stimuli presented in different locations on the body. As per Goldstein, spatial resolution exhibits variations depending on the stimulus location, ranging from an average of 10mm in the hand and lips to more than 40mm in the back and calf [63]. Additionally, the accurate recognition of the location of two vibratory stimuli hinges on two temporal factors: the duration of the stimuli and the inter-stimulus onset asynchrony (ISOA), which indicates when each actuator is activated and deactivated [136]. It's important to note that acuity diminishes as the number of simultaneously presented stimuli increases [62]. Therefore, it is evident that frequency, amplitude, body location, the number of stimuli, and ISOA are all crucial parameters to consider awhen designing tactile displays.

As mentioned before, understanding timbre depends on the spectral content of audio signals. Despite the narrow tactile perception band affecting the ability to recognize subtle spectral variations, individuals are still capable of distinguishing the timbre of various musical instruments (such as piano, cello, trombone) solely through vibrotactile stimuli [143]. The sense of touch, capable of recognizing signal waveforms, uses mechanoreceptors as tactile filters in this process [67]. This ability facilitates rendering the texture of sound, or timbre, as vibrotactile texture. In a study investigating vibrotactile discrimination of musical timbre, different signal aspects like waveform, temporal envelope, fundamental frequency, harmonics, duration, and ISOA were varied, showing that both normal hearing and hearing-impaired individuals could differentiate these timbral representations [143]. Another method for representing timbre in vibrotactile signals involves quantifying noisiness in the audio signals and reproduce it as an interpolation between a 500Hz sine tone and white noise. This process creates a vibrotactile representation of timbre, where the amplitude of the noise component is analogous to spectral brightness [77]. However, participant evaluations in this study focused more on the overall quality of the vibrotactile stimuli than on specific timbre discrimination, indicating the need for further research.

Tactile representation of loudness is relatively straightforward, as it can be directly linked to the intensity of actuators [128]. However, there are psychophysical factors that complicate the interaction between loudness and tactile rendering of music. Studies indicate that loudness perception is frequency-independent within the 20 to 40 Hz range [10], but other research has shown that low frequencies can influence loudness perception [86]. This variability adds complexity to determining the appropriate bandwidth for rendering musical information in a tactile format. Furthermore, as noted by Verillo [168], psychophysical phenomena like summation and suppression affect vibrotactile loudness perception. Summation refers to an increased perceptual loudness when two stimuli are presented within the same psychophysical channel (Pacinian or Non-Pacinian) [136]. In contrast, suppression denotes a decrease in perceived loudness of a second vibrotactile stimulus when two tones are presented in independent psychophysical channels [136]. This adds another layer of complexity to tactile rendering of loudness, particularly in multi-actuator tactile displays that deliver several vibrotactile stimuli simultaneously.

Audio-tactile interaction

There are many apparent similarities between the two sensing systems, like the response to stimulation from periodic signals, and the non-linear frequency response, going so far that deaf percussionist Evelyn Glennie states that "Hearing is basically a specialized form of touch". Nevertheless, listening for pitch is almost always dependent on the frequency of the audio content, while the timber and amplitude rarely have an impact on pitch perception [10]. In contrast, the perception of frequency from a vibrotactile stimulus is more complicated due to the multi-channel nature of the skin [10]. On top of that, the perception of frequency is amplitude and time dependent, and it varies significantly depending on the position on the body where the stimulation takes place. However, there is one important similarity between auditory and tactile pitch perception: within certain frequencies, the discrimination fits a critical band model [98]. Specifically, certain frequency ranges are perceived as distinct sensations, indicating that with enough exposure, tactile pitch perception can be interpreted similarly to the auditory one, as evidenced by previous work involving hearing impaired people [13, 113].

A well documented interaction between the tactile and auditory systems is related to the perception of loudness. A study by Schürmann et al. [148] delves into audio-tactile interactions among individuals with normal hearing reporting that when participants held a vibrating tube reproducing a congruent version of the auditory stimuli, they consistently selected lower auditory intensities to achieve equal loudness.

A similar study conducted by Gillmeister and Eimer explored the impact of unrelated tactile events on auditory detection and perceived loudness of sound [61]. One of their results revealed that weak sounds become more detectable when synchronized with tactile stimuli [61]. Furthermore, the study reported that auditory stimuli paired with synchronous tactile vibrations were perceived as louder compared to when presented in isolation or with asynchronous tactile events. This tactile enhancement effect was more prominent for lower auditory intensities but unaffected by the spatial alignment of auditory and tactile stimuli. These findings extend previous research on auditory-tactile interactions described above in [148], and underscore the principles of inverse effectiveness and temporal proximity required for multisensory integration. In a related investigation utilizing short 250 Hz sinusoidal tactile stimuli, Wilson et al. [172] reported findings consistent with Gillmeister and Eimer [61]. Previous research has indicated that the enhancement of auditory perception induced by somatosensory stimuli is more pronounced when the auditory and tactile stimuli share similar or adjacent frequencies, particularly in the lower frequency range centered around 250 Hz [173]. Interestingly, the relative phase between these stimuli appears to have minimal impact within this frequency range [172].

In a recent study utilizing electroencephalogram (EEG) to investigate the impact of prolonged tactile stimulation that fluctuates in-phase or anti-phase with concurrent auditory noise, the results revealed that when tactile stimulation was in-phase with auditory stimulation, it significantly increased cortical responses to the fluctuations in the auditory noise [44]. This enhancement was not observed when comparing anti-phase tactile stimulation to purely auditory stimulation, indicating that the effect was primarily driven by the synchrony between auditory and tactile stimuli. This finding aligns with previous neuroimaging studies and behavioral evidence,further indicating that synchronous tactile inputs can augment auditory responses [61, 148, 173]. What the study by Fu et al. brings as novelty is the obveration that anti-phase tactile stimulation induced a significant phase shift in cortical responses to the auditory noise, suggesting that the relative phase differences in the physical stimuli were partially retained in cortical processing [44]. In contrast, in-phase tactile stimulation resulted in the alignment of responses to peaks in both auditory and tactile stimuli [44]. This finding indicates that the relative phase of ongoing tactile input can influence cortical responses to auditory noise, even though participants were largely unaware of this phase, even when asked to pay attention to it [44].

While the spatial alignment of auditory and tactile stimuli typically governs multisensory integration, there is evidence that auditory-tactile enhancement operates independently of spatial congruency, suggesting a unique characteristic of this type of interaction [61, 112, 178]. Specifically, Murray et al. conducted research to comprehend how the brain processes information during audio-tactile stimulation and to determine whether this processing is affected by the spatial origins of these sensory inputs – whether they originate from the same or different locations in space [112]. The authors found that regardless of whether the stimuli were spatially aligned or misaligned, the brain exhibited consistent patterns of interaction and demonstrated enhanced processing in response to combined sensory inputs [112]. This suggests that our brain has mechanisms for combining audio-tactile information even when the sensory inputs are spatially separated — a unique propriety observed only on the interaction between these two senses.

When it comes to rhythm perception, Bernard et al. claimed that the mechanism is shared between the two system [9]. Their study demonstrated that an interaction between audio and haptic perception, previously known for pitches above 100 Hz, extended to the perception of rhythm and its temporal changes. In their experiments involving a surface-haptic device able to synthesize arbitrary audio-haptic textures (e. g., smooth, porous, rough etc)., participants were able to perceive rhythm haptically after a minimal exploration distance on a simulated surface [9]. Furthermore they discovered that if the frequency of the stimulus changed more rapidly, the distance needed to be explored with the finger to perceive this change haptically was less than it would be for slower frequency changes, following a specific predictable pattern described by the power law with an exponent of 0.5 [9]. This pattern mirrors the behavior observed in auditory perception when the tempo of a sound increases or decreases. Adding audio feedback congruent with the haptic sensation led to a notable interaction between the two modalities. When auditory tempo variations matched the haptic signal's pattern, participants detected the rhythm with 12% less exploration distance [9]. These findings suggest a bimodal integration of audio and haptic stimuli, implying that the perception of energy envelopes in both audio and haptic signals may involve shared perceptual mechanisms.

A particular application for audio-tactile integration is sensory augmentation or substitution, sometimes researched in the context of sensory deprivation. This is usually manifested in one of the following categories: vibrotactile aids for the vision impaired population, audio-tactile augmentation of speech or music for the hearing impaired, and musical haptics. The music applications will be discussed in the following Chapter 2.4. While researching vibrotactile stimulation for the visual impaired can provide valuable insight about applications of multisensory integration, in the interest of conciseness I will focus only on the vibrotactile augmentation of sound.

Riecke and their team researched the mechanism of cortical speech-envelope tracking in the context of audio-tactile speech integration, hypothesizing that supra-additivity in cortical activity would indicate multisensory integration [138]. The study found that tactile speech-shaped stimulation could enhance the cortical encoding of degraded auditory speech, and as mentioned earlier in this chapter, that audio-tactile integration was observed when tactile input led auditory speech by 100 milliseconds or less. This finding suggests that the timing of tactile stimuli relative to auditory input is critical for effective integration, with tactile cues potentially preparing the cortex for upcoming auditory speech peaks [138]. Furthermore the study revealed that audio-tactile speech-envelope integration involved significant contributions from cortical activity in the delta, theta, and alpha frequency bands. The delta band, in particular, showed the strongest contribution [138]. Surprisingly, despite the observed enhancement in cortical speech-envelope tracking due to tactile input, there was no evidence of a corresponding improvement in speech intelligibility. This discrepancy between neural results and behavioral outcomes may be due to factors such as participants' familiarity with the stimuli and the complementary nature of the audio-tactile information [138]. The study indicated that audio-tactile speech-envelope integration primarily affected neural generators in the auditory cortex rather than the somatosensory cortex [138]. This suggests that the observed effects were more related to enhancing auditory processing through tactile input, possibly because the meaning of the stimulation was carried by the sound.

A similar conclusion was reached by another group of researchers that investigated the effects of vibrotactile augmentation on perceiving speech-in-noise [22]. They found that when auditory signals were degraded, adding complementary tactile stimulation following the fundamental frequency of the voice significantly improved speech understanding in noise. "This enhancement was both automatic and consistent among participants, leading to a notable group benefit of 6 dB [22]. This is especially significant considering that a 3 dB increase equates to a doubling of sound intensity, while a 10 dB increase corresponds to a perceived doubling of loudness.

M. Fletcher and his research team have made significant efforts in documenting how the introduction of tactile stimuli, shaped to mimic the temporal patterns of auditory speech, can enhance the comprehensibility of auditory speech presented concurrently in noisy environments under various scenarios. This phenomenon has been documented in individuals with normal hearing [38] as well as in cochlear implant recipients, following their participation in audio-tactile speech training programs [36, 37, 37]. One of their studies creatively extend the focus towards multi-talker scenarios where spatial audio perception is critical [39]. The researchers found that audio-tactile stimulation significantly improved speech recognition in multi-talker noise scenarios where sounds were spatially separated. This improvement was observed in CI users with a single implant, who make up the majority of the CI community. Both ipsilateral (noise on the same side as the implant) and contralateral (noise on the opposite side of the implant) noise conditions showed improvements, with mean Speech Reception Threshold (SRT) improvements of 2.8 dB and 2.6 dB, respectively. Notably, this improvement was achieved after just 30 minutes of training and was consistent across users of different cochlear implant systems.

2.4 Tactile displays and music

So far I have briefly described the building blocks of my project that are part of the discussion about vibrotactile displays, for music listening by CI users. Throughout this chapter I finally approach the main topic of my research, and discuss the implications of augmenting music with tactile stimulation.

The goal is to provide a basic understanding of the concepts necessary to interpret the articles in the second part II, and to contextualize the work both in terms of academic research but also regarding artistic or commercial relevance. However, I will start by briefly introducing the technology behind vibrotactile augmentation.

2.4.1 Tactile displays

At the hart of any sensory augmentation or substitution system is the interface used to convey the message, in my case the tactile display. These systems are heavily responsible for the experience users have, and they vary greatly depending on the application. For example, and augmented violin designed to improve performing capabilities of novice players [123] has a completely different set of requirements than a *Tactile Phoneme Sleeve* [135] created to research vibrotactile speech perception. While those examples have radically different designs, they do follow a similar architecture, adapted for their bespoke requirements. A common architecture of a tactile display can be seen in Figure 2.2 that describes the signal chain from auditory source to tactile stimulation (the figure is presented in paper A). What is interesting is that every step is optional, and it is perfectly possible and valid to create tactile displays only including only the source and the output elements, if the application requires so. Nevertheless, my review of the tactile displays presented in paper A, reports that most of the existing projects employ more than one element (with various degree of sophistication), with the majority having at least capture, pre-processing and mapping as part of the signal chain.

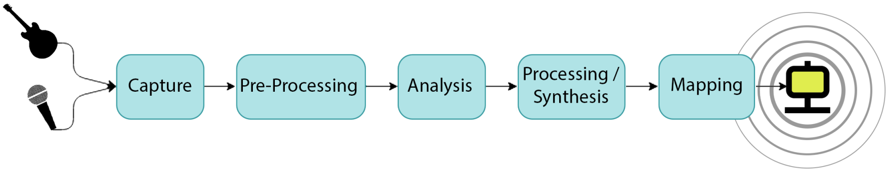


Fig. 2.2: Archetype of a tactile display - Figure taken from paper A

Actuator types

When it comes to the actuator — the final element in the chain presented in figure 2.2, there is a large palette of possibilities to choose from, with great effect on the user experience. The choice of actuators depends on the role of the vibrotactile stimuli in the interface, it's size and energy limitations, as well as the tactile stimuli expected to be reproduced [62]. To complicate things even more, it is possible to combine different types of actuators as part of the same display; not only that, but it is generally a good idea to do so, as they can be mixed according to their properties. In the interest of clarity, I will briefly present the types of actuators, and their properties, with respect to vibrotactile augmentation.

- Voice coil actuators are probably the most popular type used for musical reproduction — most loudspeakers are build on this design. The principle is simple, there's a permanent magnet interacting with an electromagnet that is attached to a moving component. For loudspeakers this is the cone, for tactile actuators it's a mass. What makes the voice coil actuators easy to work with is the fact that most of the audio technology is built upon them, and there are many standards for voltage and current, and the subsystems are ubiquitous (amplifiers, converters, processors, etc.). Another benefit of voice coils actuators is that there are easy to scale up, if enough power is provided. The downside is that they cannot be made to be small, and the intensity of the stimulation is generally correlated to the size of the actuator. On this principle, several variations are built:
    - Bass shakers are voice coil actuators marketed towards audio listening applications, be it films, games, playing or listening to music. They are generally large, heavy and require abundant power to operate, but for that they can provide intense whole body stimulation for more that one person, and good frequency response, especially in the lower frequency bands.
    - Contact speakers are the smaller cousin of bass shakers, with intended applications limited to reproduction of audio materials. Due to their rather large mass, and good contact surface, these actuators can be re-purposed to be used for local tactile stimulation, but due to their small size, they struggle to produce low frequency at convincing amplitudes.
    - Subwoofers are not generally used for vibrotactile stimulation, but with enough power and size, they do produce convincing tactile sensations. Unfortunately, this comes as a byproduct of audio stimulation in the first place, so there's limited application for those given the extreme loudness required for tactile stimulation.
- Linear Resonating actuators are somewhat similar to the voice coil ones, as they also contain an electromagnet controlled by alternative current, which interacts with a permanent magnet. Unlike the voice coil ones, the linear resonating actuators have a limited bandwidth, as they have to be driven at the resonant frequency of an in-built spring, to produce perceptible stimulation. Another differentiating factor is that they can be built much smaller and cheaper than voice coil actuators, contributing to their popularity.
- Eccentric Rotating Mass actuators are another type of rotating actuators, but instead of relying on AC, they are driven by a direct current signal,

which couples the amplitude to the frequency. This type of actuators can be build fairly small, and have a good size-to-intensity ration, but only operate in a very limited frequency range, usually around 200Hz. These are very popular and cheap actuators, and are found in most mobile phones (but not iPhones, which features dual-mode one).

- Dual-Mode Actuators can be seen as a combination of two linear resonating ones in the same enclosure, tuned to resonate at different frequencies. The combination of these can produce complex signals that are generally perceived as stronger and more accurate than a similarly sized and powered linear one. While no official source from Apple Inc. has declared this, their *"Taptic Engine"*, found is most modern iPhones is a dual-mode actuator.

- Solenoids are electro-mechanical devices that have binary states - open or closed, and depending on their design one of these states is the default one, usually ensured by a spring that resets a pin to it's original position. Attaching a mass to the pin of a solenoid can create a convincing tactile sensations, but due the the construction of the actuators they have limited operating frequency band, and are generally loud. They have been popular in the past, but with the advancement in voice coil, and resonating actuators, they have not been used frequently in current vibrotactile displays.

- Piezo actuators vary from all other as they vibrate when current is applied to them as a result of changes of the shape of their piezoelectric material. They are also part of the resonating category, but due to the lack of moving parts, these actuators can be custom built into any shape. Unfortunately, their stimulation intensity is relatively modest, and are rarely used in vibrotactile displays as a result.

- Electro-hydraulic shakers are industrial systems that rely on changes in hydraulic pressure to vibrate. They can be built to order and are very versatile, but extremely expensive (when compared to the other types presented above). Hydraulic actuators are mostly used in vibration analysis and acoustic industries, and in rare occasions in commercial applications, like the *6D Cinema* concept, where they are deployed on rows of seats.

What it can be seen is that most tactile actuators are variations on the moving voice coil design, and it is for good reason. Audio technology evolved on the assumption that the output will be a voice coil loudspeaker, powered by alternative current and in order to make use of all this existing infrastructure and knowledge, it's more efficient to adapt designs than create new ones all together. Nevertheless, Dual-Mode actuators are relatively new, and do show that new ideas are worth exploring, as they provide quantifiable benefits over older designs.

Number of actuators and mappings

Unlike sound that can stimulate only two fixed locations — the inner ears, vibrotactile displays can be attached anywhere on the body, and the choice of location will have an enormous impact on the user experience. Furthermore, vibrotactile displays can be use to stimulate multiple parts of the body at the same time, with different types of actuators and signals. This flexibility opens the doors to unlimited opportunities, but also raises many question regarding optimal choice of locus, stimuli and actuator type. Nevertheless, it's not necessary to have more than one actuator, as single actuator displays have been used successfully in a plethora of projects.

Another aspect that is fundamental when working with tactile displays is the mapping scheme employed. When talking about *mapping* in the context of audio-tactile stimulation, there are actually two types that need to be designed: audio-to-tactile mapping scheme, as well as body mapping — the later applying to multi-actuator systems. Starting with the simplest of the two (at least in terms of hardware), a single actuator would rely heavily on the audio-to-tactile mapping scheme if the designer intends to encode complex information in vibrotactile stimulus. The most simple one is actually no mapping — with certain types of actuators (mostly voice coils), one can use the sound material as source for vibrotactile stimulation. This is not a bad mapping either, but it limits control over the tactile display. A common variation of this mapping relies on a low-pass filter to remove spectral content lying outside the range perceivable by the skin, as presented by Frid and Lindetorp [43] in their *Sound Forest* installation found in the Swedish Museum of Performing Arts. This method, while computationally lightweight, offers a practical solution for systems with limited processing power.

A slightly more complex audio-to-tactile mapping used with single actuator displays (but not excluding multi-actuator devices) implies using a sinusoidal signal with temporal characteristics matching the audio one. This mapping has been used by Merchel and his team that designed a chair to investigate the impact of vibrotactile stimulation on concert reproduction [106]. In addition to this mapping scheme, they employed a tactile stimulation method where the vibrating frequency matched the pitch of the audio, as demonstrated in [38, 153, 166], or was slightly offset while still preserving the relative relationship. In the context of audio-tactile music experiencing, these are the most common mapping schemes for single-actuator devices; more complicated mappings have been trialed, usually in combination with bespoke tactile displays, but no particular one has been widely adopted.

The discussion about mapping gets increasingly more complex when multi-actuator tactile displays are used. In terms of audio-to-tactile mapping schemes, there seems to be less emphasis on complexity, as direct audio (with minimal pre-processing) is commonly used, as discussed in Paper A. When it comes to the body mappings, one particular scheme stands out: pitch-to-position. This design splits the audio signal into multiple streams that are each reproduced on distinct actuators ate different locations on the body. The concept of utilizing a pitch-to-position mapping, which is the fundamental idea behind cochlear implants (CI), has been used for an extended period. This persistence can likely be attributed to the fact that the original creator of CI and one of the first audio-tactile sensory augmentation device (the *"Tickle Talker"*), employed the same mapping approach in both inventions. This shared approach might have served as a source of inspiration for subsequent researchers, primarily focused on making incremental enhancements to these systems rather than pursuing revolutionary approaches, as exemplified in the studies conducted by Nanayakkara et al. [113] and Karam et al. [81]. A variation of this mapping scheme that became popular in recent years with the advancement of vibrotactile music research is instrument-to-position which implies using multi-track music source and present each instrument on a different body part as described by Karam et al. [82]. Furthermore, the authors of this study propose a system that has been explored and confirmed by many researchers, who suggest to employ a positive correlation between the frequency of stimuli and the vertical arrangement of actuators on the human body. They named this mapping scheme "T*he Model Human Cochlea"*, as it draws inspiration from the tonotopic distribution of frequencies in the ear. This has been a frequently used mapping, as users seem to understand intuitively that low frequencies come from lower areas, as documented in paper F. In my multi-actuator tactile displays, I have used this spectral distribution as well, mapping the low frequencies (or instruments playing the lower register) lower on the body, relative to the higher instruments.

2.4.2 Audio-tactile interaction and music

Tactile stimulation can influence and even enhance musical experiences, as discussed in chapter 2.3, but the intricate cross-modulation between the two systems has not been discussed yet. Throughout the next paragraphs I will try to briefly approach the impact of vibrotactile augmentation on music perception, as wells as the music experience. The goal is to present what are the know tactile parameters that manipulate the experience of listening to music through audio-tactile stimulating devices. Since most existing vibrotactile music enhancement devices aim to preserve the intensity [34], the timing or both between the two modalities, the discussion that follows will assume audio-tactile intensity and time alignment congruence.

In a landmark study, my peers from across the city investigated the impact of audio-tactile congruence on music enhancement, and shed light on the specific vibrotactile parameters that influence this enhancement [1]. The study reveals that audio-tactile congruence significantly enhances vibrotactile music, with a degree depending on specific vibrotactile parameters aligning with the music [1]. Time alignment and intensity congruence notably influence participants' ratings, while frequency congruence does not significantly alter ratings, except in certain conditions involving musicians, who exhibited better vibrotactile frequency discrimination [152]. The most substantial effect on participant enjoyment was observed when the tactile stimulus was synchronized in time with the auditory stimulus [1]. Misalignment in timing consistently resulted in lower ratings, indicating the critical role of timing alignment in multisensory enjoyment [1].

Musicians, in particular, demonstrated a greater sensitivity to the intensity congruence of stimuli, suggesting that their enjoyment was more significantly impacted by intensity alignment compared to non-musicians. This finding opens up a discussion about *"The Musician Effect"* (which claims that musicians perform better than non-musicians at audio-perceptive tasks post cochlear implantation) with respect to tactile stimulation, not only to auditory one, as investigated by Başkent and her team [6, 7, 46]. However, the [1] study utilized a single melody based on a pop song, potentially limiting the generalizability of the results to other musical styles and genres [1]. Additionally, the potential influence of rhythm and intensity as interrelated factors was not fully explored, and the musicality of the tactile stimulus itself was not controlled, which could have contributed to the music enhancement observed.

A similar conclusion was reached by Huang and his team that researched the impact of vibrotactile augmentation on music perception, particularly in CI users [75]. The study observed that the multisensory stimulation significantly improved pitch perception in CI users, regardless of their musical training experience. The tactile stimulation in the study was low-passed at 500 Hz [75]. The advantage of this low-frequency tactile stimulation might not apply to a wider frequency range or more complex musical sounds. However, it's crucial to ensure that it falls within the skin's sensitive range. Musical training was found to have a significant impact on the enhancement provided by tactile stimulation in melody recognition tasks [75]. Pre-CI musical training led to better post-CI melody recognition in musician CI users compared to non-musicians, possibly due to enhanced auditory-somatosensory integration and neural responses to pitch information. Musicians were also able to recognize melodies through low-frequency tactile stimulation alone, suggesting that musical training enhances tactile processing and auditory-tactile integration [75]. Although musicians had greater absolute melody recognition scores than non-musicians, pre-CI music training did not result in more relative enhancement from vibrotactile stimulation integration [75]. The relative enhancement, normalized by the electrical stimulation baseline, was similar between musicians and non-musicians in both rhythmic and non-rhythmic melody conditions [75]. This indicates that the benefit provided by adding tactile stimulation in music processing is independent of musical training.

A study by Fletcher et al. investigated how an multi-actuator tactile display worn on the arm can enhance pitch discrimination in normal-hearing subjects listening to cochlear implant (CI) simulated audio [35]. The average pitch discrimination threshold with haptic stimulation was just 1.4% without noise, markedly better than the target of 6% (1 semitone), and even the worst-performing participant achieved a threshold of 3.5% [35]. This performance aligns with the best-performing CI users' pitch discrimination abilities [28, 80]. Some participants attained pitch discrimination thresholds as low as 0.8%, comparable to normal-hearing listeners for similar auditory stimuli [78]. Interestingly, no difference was found between the audio-haptic and haptic-alone conditions, suggesting that the quality of pitch information from auditory stimulation did not degrade performance, potentially due to the principle of inverse effectiveness [146]. Due to the design of their device, the pitch discrimination performance remained robust against background noise, with no noticeable effect of noise on pitch discrimination thresholds, even at -7.5 dB SNR. However, it is difficult to generalize their results to other tactile displays that do not encode pitch into tactile stimulation position, as it seems that the participants in the presented study relied mostly on the position of tactile stimulation for identifying auditory pitch. While the authors do not report anything about music listening experience, it's fair to assume that the increased pitch discrimination ability did not affect it, as an increase in musical listening performance does not automatically result in a better listening experience [47].

Continuing the discussion about audio-tactile congruence and pitch perception, the focus shifts to a study that measured the impact of vibrotactile augmentation for both simple and complex waveforms, but only at 160 Hz. [177]. The main effect of waveform type on pitch perception was not significant, sine wave stimuli showed a more distinct curve between groups compared to complex waveforms. This suggests that the complexity of the waveform influences the perception of pitch in extra-auditory vibrotactile feedback exercises, with the effect being less pronounced for more complex waveforms. This finding doesn't reduce the potential application of complex waveforms in vibrotactile feedback but highlights the need for a balance between simple and complex signals in real-world applications. The study also observed differences in Just Noticeable Difference (JND) values between auditory and tactile systems. The tactile system's JND for a 150 Hz sinusoidal stimulus, with constant amplitude, was found to be broad, approximately ± 18% (27 Hz), equating to 28.8 Hz at 160 Hz [131]. In contrast, an auditory-only JND experiment would typically expect a variation of 3 Hz for sine wave and 1 Hz for complex waveforms below 500 Hz [8]. The study's results indicated that JND values for both the audio-only and the audio-tactile groups were relatively equal across all waveform types, with only minor improvements when vibrotactile information was added. For instance, in the audio-tactile group, JND for sine waveforms was 1.83 Hz, while for saw and square waveforms, it was 1.89 Hz and 1.78 Hz, respectively. It's worth nothing that since the study is only conducted at a single frequency, and one that is rather optimal for the tactile system, it is hard to generalize these results for the entire spectrum perceivable on the skin. Furthermore, the study also evaluated only one amplitude, further limiting the possibility to generalize, as there is strong evidence that stimulation amplitude impacts the perception of tactile frequency [111].

Looking at the overall experience, Merchel and Altinsoy explored audio-induced vibration generation approaches and their effect on the perceived quality of concert reproduction using loudspeakers sound and vibrating seats. The tactile display they used significantly enhanced the music experience, with all evaluated vibration-generation approaches scoring better than experiences without vibrations [106, 122]. They also explores some different audio-to-tactile mapping schemes concluding that a simple low-pass filter yielded good quality ratings [106, 122]. Additional processing like compression in the frequency range, such as octave shifting, can reduce unwanted sounds while preserving quality. Even amplitude-modulated sinusoidal signals were found to be effective, allowing for simpler and computationally cheaper vibration systems [106, 122], requiring only to extract the temporal envelope of the original signal followed by eventual dynamic compression [106, 122]. Furthermore, their participants showed tolerance to a wide range of music and seat vibration combinations, and reporting that the type of music affected the perceived quality of vibrations, with rock music benefiting more from added vibrations than classical compositions [106, 122].

2.4.3 Musical Haptics

One emerging research filed closely related to my project is called *Musical Haptics*, and it lies at the intersection between music performance and haptic interaction. It is interdisciplinary by nature and it explores the roles of touch and proprioception in music-related contexts. This field integrates knowledge from haptic engineering, human-computer interaction (HCI), applied psychology, musical acoustics, aesthetics, and music performance. The primary objectives of research in musical haptics are twofold: firstly, to gain insights into how haptic interactions contribute to the experience of music and the performance of musical instruments; and secondly, to develop innovative musical instruments or devices that provide significant haptic feedback [122]. These goals lead the researchers into investigating digital musical instruments (DMIs), frequently augmenting them with haptic systems in order to evaluate their performative qualities, or the role of the additional stimulation on various parameters (e.g. the quality of the instrument, as per [144]). While the fundamental research blocks in the field of musical haptics coincide with my projects' (e.g. tactile displays, mapping schemes, tactile perception, etc.), the large difference in target groups and the special requirements associated with that results in little overlap in research direction. Nevertheless, I am sure that in the near future the artistic approach of musical haptics researchers will find interest in the challenges presented by cochlear implant music listening, resulting in creative solutions for this particular target group.

2.4.4 Tactile music installations for hearing impaired individuals

Based on the principles of multisensory integration presented in chapter 2.3, as well as the technology and techniques listed above, several sensory augmentation systems have been created to aid the hearing impaired population with their music listening experience. The following section will present and discuss some of those projects; it's worth noting that there are not that many systems designed exclusively for the hearing impaired, as most devices are experimental and focus more on researching perception, or are created for augmenting the music experience for normal hearing individuals. While there is considerable overlap between the knowledge necessary to successfully create installations for everyone, I will only focus on those that have accounted for the specific needs of hearing impaired, and by extension CI users.

One of the earliest initiatives to develop music augmentation systems for hearing-impaired users was undertaken by Nanayakkara and his colleagues, who created the *Haptic Chair* [113–115]. This chair was developed through a participatory research study in collaboration with the hearing-impaired community. Its design concept was inspired by feedback from deaf musicians, who suggested that enhancing the body's perception of sound vibrations, similar to what occurs in natural environments, could improve music enjoyment compared to relying solely on visual cues or amplified audio. Initially designed for sound experience through touch, later versions of the chair incorporated voice-coil motors into its back, extending the range of frequencies covered [113]. This evolution in design recognizes the diverse spectrum of hearing impairments, acknowledging that many individuals with partial deafness can still perceive certain sounds through normal air conduction.

One of the earliest initiatives to develop music augmentation systems for hearing-impaired users was undertaken by Nanayakkara and his colleagues, who created the *Haptic Chair* [113–115]. This chair was developed through a participatory research study in collaboration with the hearing-impaired community. Its design concept was inspired by feedback from deaf musicians, who suggested that enhancing the body's perception of sound vibrations, similar to what occurs in natural environments, could improve music enjoyment compared to relying solely on visual cues or amplified audio. Initially designed for sound experience through touch, later versions of the chair incorporated voice-coil motors into its back, extending the range of frequencies covered [113]. This evolution in design recognizes the diverse spectrum of hearing impairments, acknowledging that many individuals with partial deafness can still perceive certain sounds through normal air conduction.

A few years later, a collaborative project between Queen Mary University and the Deaf arts charity group *Incloodu* led to the development of an installation featuring a sofa and armchair [77]. This installation incorporated a *subPac*[2] device placed under the seating area and voice coil actuators in the backrests and armrests. The armrests were designed to replicate a noisy component associated with timber. The distribution of spatial auditory information was arranged from low to high frequencies along the backrest, reflecting a cochlear metaphor as described in [82]. The furniture was designed by a severely deaf architect with expertise in creating accessible furniture. Their analysis highlighted that the style of music played a crucial role in the user experience, with highly rhythmic music generally eliciting more positive responses than music predominantly characterized by harmonic movements.

Yet another variation on the vibrotactile chair comes from the efforts of Marozeau and his team. At the Museum of Art and History in Geneva, a year-long exhibition allowed visitors to experience the *Tactile Chairs*. These are part of the continued development of a project named *Augmented Music*, that created several prototypes of seated vibrotactile installations [155]. The first prototype of the chair is designed to distribute low and middle vibratory frequencies across various parts of the user's body. The second prototype diverges from transmitting vibrations through the chair's base and instead employs wooden rods along the chair's sides. This design is influenced by the structure of the basilar membrane in the inner ear [99].

Diverging from the seated position, there are very few studies exploring vibrotactile augmentation of music for the hearing impaired. The *VibGrip* is such a device - a multi-actuator hand held tactile display using piezoelectricity to produce individual vibrotactile stimulation for each finger [79]. Unfortunately the authors did not evaluate the system yet. Similarly, a *Haptic Sleeve* was designed to be used on the legs of hearing impaired people, mapping musical notes to positions around the tibia [166]. Just as with the previous tactile display, the *Haptic Sleeve* was not evaluated in the intended context, making it hard to interpret the efficacy of such a system.

---

[2] https://subpac.com/

CHAPTER 3

Research Questions

The main goal of the project is to advance the state of the art in vibrotactile augmentation of music targeted at cochlear implant users. As previously described, pitch recognition, and its impact in melody perception, along poor timbre discrimination and source separation are some the most affected music properties for CI users, but these limitations do not correlate perfectly to musical experience or enjoyment. Music is more than these descriptors, therefore the project will be approached from different perspectives.

3.1 Problem area 1

PA1 summary: Perception and discrimination of musical features through vibrotactile displays

The central questions in this problem area, with implications for the entire project and research field, revolves around the perceptual characteristic of the tactile sense. Initial literature overview highlighted a need for studying some of the most fundamental aspects of the perception of vibrotactile stimulation. This problem area is more directed towards basic research paradigms.

3.1.1 Perception and discrimination of musical features through vibrotactile displays

Problem Area 1, to understand tactile musical perception, has been a central area of investigation ever since the start of the project, as a good knowledge of the current state of the art is fundamental for any research on the topic of this project. The initial goal was to identify an algorithm to translate any musical material for usage with vibrotactile displays. I started from the assumption that such an algorithm exists and it was descried in existing literature - which was incorrect. Vibrotactile stimulation for musical applications is a very complex topic, that is not understood well enough. The existing literature deals with simple cases and naive approaches for music-to-tactile translation. Furthermore there seems to be no consensus about an optimal algorithm — a discovery I found during the background research process — as existing studies vary on the type of music presented, number and type of actuators available, body area actuated, and more. With this in mind, the PA1 was narrowed down to these specific research questions:

Q. 1 *Which optimal audio-to-tactile conversion algorithm should be used with vibrotactile displays designed for music listening by CI users?*

Q. 2 *Is there a relationship between stimulation frequency and area of the hand of maximal perception when using single-actuator vibrotactile displays?*

3.2 Problem area 2

PA2 summary: Vibrotactile augmented concerts

The second problem area of my project revolved around concert experiences, and incorporating vibrotactile displays in these social music listening environments. Together with a multidisciplinary team we aimed to improve concert experiences for CI users — this problem is relying on the knowledge obtained through researching the first area, and is leans more towards applied research methodologies.

3.2.1 Vibrotactile augmented musical experiences

The benefit of music exposure is well documented for normal hearing users, and most of them are applicable for CI users as well, as explained in Chapter 2. Attending concerts supplements the well-being with social interaction and engagement. However, the concert experience presents a unique set of challenges and opportunities for CI users that differ from the normal hearing attendees. For example, from a human-computer interaction perspective, there is an opportunity to design assistive technologies like vibrotactile displays that can enhance the concert experience for CI users, addressing specific difficulties such as pitch identification, instrument segregation, and overall music perception. On the other hand, from a clinical research perspective, understanding how these displays impact the overall well-being and musical engagement of CI users is important, as the hearing rehabilitation process needs to be undisturbed by eventual ”crutches”. Additionally, approaching the issue from a musicology and phenomenological perspective is essential as well, considering the subjective and embodied nature of musical experiences, specifically in a social context. As such, there is a dual challenge: designing effective vibrotactile displays and creating inclusive concert environments that attract and cater to the unique needs of CI users, without excluding normal hearing individuals that might accompany the hearing impaired ones. The following research questions emerge:

Q. 1 *How to design vibrotactile displays that enhance the concert experience for CI users?*

Q. 2 *How to create concerts that increase the attendance from CI users?*

3.3 Problem area 3

PA2 summary: Music listening training using vibrotactile displays

The final problem area investigated through my project focused on training CI users to listen to music. The goal was to observe whether exposure to guided music training would result in objective improvement in the music listening abilities of CI users. This problem is a drawing insight from problem area 2 but is fundamentally a basic-research method in an applied context.

3.3.1 Music listening training using vibrotactile displays

As mentioned in Chapter 2.2, exposure to music and training are solid predictors of better hearing performance. So far, researchers have attempted to train for specific aspects of music listening, with bespoke exercises like the one seen in [47, 50]. These approaches have shown some improvement regarding trained aspects (e.g., MCI). However, to our knowledge, no study has been conducted with multisensory stimulation, with existing music and not only particular exercises, leading to the following research question:

Q. 1 *How does audio-tactile music training at home impact the listening performance of CI users?*

CHAPTER 4

Summary of Included Papers

The main contribution of the thesis has been documented in eight articles presented in Part II of this dissertation. Each paper describes an empirical study or research synthesis addressing one of the research questions found in Chapter 4. The papers are not presented chronologically, but instead are grouped based on tools used and research question(s) addressed. This chapter provides a brief overview of the motivation, methodology and findings of each of the included articles.

Paper A: Tactile displays for auditory augmentation - a scoping review and reflections on music applications for hearing impaired users

This review article focuses on investigating and organizing the tactile displays that have been used for augmentation of sound or music, usually attempting to address aspects of hearing impairment. The focus is mostly on the technology and how such devices have been evaluated.

Motivation

The review detailed in the article spawned from many questions that arose during the PhD proposition and planning phase, as the information about vibrotactile devices was scattered or incomplete. Through this article I organized all the tactile devices used for sound or music related applications detailed in literature and performed some statistical analysis on the main features. Specifically I looked for patterns and commonalities in terms of types and number of actuators, signals used, mapping schemes, as well as into what body areas are commonly paired with specific actuators, and eventual evaluation methodology.

Methodology

The main method used to identify the relevant articles was derived from the *PRISMA-ScR* methodology and structure [165], and was applied on literature indexed by the *Scopus*® database. This meant that there was a sequential exclusion system, with various criteria identified at each step resulting in a truncation from 3555 potential articles, to 63 scientific publications included for further analysis.

The eligible articles were studies and disseminated into a table containing tech descriptors as well as evaluation practices that served as a foundation for a statistical analysis of the literature.

Findings

Many interesting aspects have been uncovered through the analysis, and some of those are included in the article. Probably the most important one is that the vibrotactile augmentation field is still in an early phase, characterized by an exploratory approach and preliminary results, with most of the literature published after 2009. This phase is further emphasized by a large variety in designs and approaches, with no apparent convergence. Another important finding revolves around the low validity and reliability in the evaluation practices, calling for more longitudinal studies as well as ecologically valid ones, as many studies are conducted exclusively in laboratory with designs that are rarely designed with a user-centric approach.

Besides the documented findings, the process of creating this article helped me on understanding the landscape of vibrotactile augmentation, highlighting the strengths as well as the shortcomings of the academic community, thus helping me focus in my research activities.

Paper B: The Relationship between Frequency and Hand Region Actuated

This article documents a study conducted on a single actuators handheld vibrotactile display called *the VAM*, and how different frequencies impact the perceived actuated area, as part of a larger effort focusing on building novel devices for musical augmentation.

Motivation

The spark that started the study is a series of reports from participants in the article C that claimed they felt some stimuli in the fingertips, while others in the palm of their hand, when using the same tactile display. Therefore the study aimed to understand how constant amplitude vibrotactile stimuli with distinct frequencies interact with different areas of the hand. The goal is to demonstrate that specific frequencies elicit stronger sensations in certain regions of the hand, resulting in potentially cheaper and simpler tactile displays.

Methodology

Two studies were conducted with similar tactile displays; one had the actuator mounted closer to the bottom side (thus closer to the thumb), while the other had the actuator closer to the top side (palm and distal palmar areas). Participants (65) were exposed to 11 distinct stimulation frequencies several times and were asked to report 3 out of 28 possible areas that the perceived the vibration most intensely. The data gathered was used to create a mixed effect regression model that could explain the effects.

Findings

The study confirmed the hypothesis that different stimulation frequencies are perceived in distinct areas of the hand, but no apparent mapping has been identified. This only emphasized the complexity of the skin's frequency response, influenced by various factors such as amplitude, duration, location of stimuli, as well as the multi-channel nature of the sensing system. Moreover, the study revealed that the perceived frequencies and areas depend on the design of the tactile display, further increasing the input variables number required to understand and control when designing tactile displays. These findings contribute to the understanding of how vibrotactile stimuli can be tailored for specific areas of the hand, with implications for designing tactile displays with widespread applications.

Reflections

This study was unique in that it was not originally planned; however, the feedback from users of the VAM consistently piqued our interest — they often mentioned feeling sensations *in their fingers*. To the best of my knowledge, this is the first study to investigate spatial perception of tactile stimulation using a single actuator. In retrospect, I would have preferred to control for tactile pressure at each spot, as grip force might correlate with the distribution of perceptual sensations. It would also have been ideal to measure the frequency response at every possible position to ensure that no resonant frequencies were present only in certain areas. Nonetheless, I have applied a compensation filter based on an average of measurements from three different positions, but having a comprehensive map of frequency responses would have provided significant insight. Regardless, I am happy that I started the discussion about single-actuator multi-area stimulation, and I hope my study is the proof of concept necessary for further research.

Paper C: A Comparison of Audio-to-Tactile Conversion Algorithms for Melody Recognition

This article presents a study that compares audio-to-tactile conversion algorithms with respect to their ability to convey melodic information through vibrotactile feedback alone. Two of the algorithms are from existing literature, while one is propose by us; all have been evaluated on the same VAM vibrotactile display as paper B.

Motivation

The study was initiated as one of the first steps in developing vibrotactile displays for music augmentation specifically created for CI users. One important part contributing to the success of new displays is the mapping between full spectrum music to stimulation suitable for the proprieties of the skin; unfortunately there's no consensus on the best approach, and this study was set to evaluate several options, as well as the general usability of the tactile display.

Methodology

A total of 34 participants from evaluated the conversion algorithms, through a task of identifying which tactile stimuli is coherent with the auditory one presented through headphones — a three alternative forced choice design. Each trial consisted of listening to the same auditory melody three times, with different haptic stimuli for each. The auditory stimulation consisted of 72 unique melodies written in a popular western music style and were played on (virtual) bass, subtractive synthesizer, and trumpet.

Findings

The results of the experiment reveal that, while there is no significant difference between any of the three proposed processing techniques and no processing at all in terms of melody matching, there was a tendency for users to perform better when exposed to haptic stimuli featuring shorter notes. Furthermore, a similar result was found with respect to differences between instruments, as they performed similarly on all four conditions. Our results contradict some of the existing literature emphasizing the importance of the audio-to-tactile conversion algorithm, at least in terms of it's capability of convening melodic information.

When it comes to the tactile display, most people were satisfied with the comfort and intensity of the stimulation, but overall it was underlined that individual customisation is desired.

Reflections

As the first study of my doctoral research, it provided a good insight into what is ahead — uncertainty and opportunities. Fundamental to my planned research was the existence of an optimal audio-to-tactile conversion algorithm, that should have been (easily) identified though this study. Being the first study, I inevitably had some blind-spots that might have set me on a sub-optimal path. Firstly, the instrument matching task was to difficult for most participants, thus it would have been good to start with a simple melody recognition one and scale up the complexity after gathering fundamental data. Furthermore, it would have been beneficial to present all melodies with the same rhythmical structure, in order to minimize the confounding effect.

Paper D: A Real-Time Cochlear Implant Simulator - Design and Evaluation

This article describes the process of implementing and evaluating a real-time cochlear implant simulator that can be used live or with pre-recorded material. The simulator is intended as a tool for normal hearing users to explore and understand the sonic palette of a CI users, without resorting to electrdograms sonification.

Motivation

People across many disciplines working on cochlear implants in some capacity are not implanted themselves, therefore it is sometimes hard to understand how the CI users experience daily sounds. To help with this discrepancy, we created the simulator, featuring a flexible architecture with many user-configurable parameters like number of active electrodes, implantation depth, as well as different carriers for a vocoder such as sine, noise, and pulse-spreading harmonic complex (PSHC). This can be valuable tool for investigating the sonic experiences of CI users by exploring parameters that could enhance their music listening experiences.

Methodology

The simulator was built as an a compilation of extension of existing ones presented in literature, selecting the most relevant features from each, and implementing them as a real time system — a novelty in the field of CI research. The goal was to apply as much of the signal processing existing in the real CI processors as possible, and using a variable-band vocoder to synthesize the final result. The system was evaluated using a *melodic contour identification* task by 19 normal hearing participants. In order to provide comparable results with existing literature, the simulators' parameters and stimulation were selected to match another system described by [83].

Findings

The evaluation revealed that the simulator can mimic the decrease in melodic contour identification abilities observed in cochlear implant users, particularly in distinguishing smaller semitone intervals, suggesting that it could be considered a relevant tool in CI research. However, the system was evaluate with a single configuration, on normal hearing people that had only several minutes to accommodate to the *artificial hearing*, unlike CI users who undergo many mapping sessions in order to optimise the signal processing and reproduction. Therefore, our simulator did not evaluate an accurate comparable perception, and it should not be used as a replacement for a user-centered design approach. Nevertheless the CI simulator offers a platform for further investigations aimed at enhancing music perception for individuals with cochlear implants.

Reflections

The most difficult aspect of this study was balancing the comparability of our research with previous one done on bimodal CI users that featured a very simplistic CI simulator, and utilizing the more advanced features of or flexible simulator. This was a constant point of discussion, but due to the limited time associated with a semester project, we had to compromise and use similar settings as the original study. Now that we have validated our system, it would be interesting to explore the CI simulator's flexibility together with bimodal CI users and investigate to what extent the two listening experiences match.

Paper E: Multisensory Integration Design in Music for Cochlear Implant Users

This article describes a participatory design process that focuses on involving CI users in the creation of novel tactile displays and experiences. This work is part of the same larger effort as paper B and C focusing on creating efficient multisensory systems for enhancing the music experience of CI users.

Motivation

The motivation behind this paper stems from the limited personal understanding of CI users' engagement and interaction with music. Given the diverse challenges faced by CI users in perceiving and enjoying music, it is paramount to invite them early into the design process and approach their challenges in a collective effort. Therefore, we have prepared four distinct systems with multiple actuators and loudspeakers to present to the participants and explore the possibilities they afford.

Methodology

At the core of this study lies the user-centered design approach through a participatory process, giving the users agency and control over the designs. Four CI users with different hearing profiles participated in the workshop, that consisted of a short interview, as well as interaction with the installation setups. Each setup focused on at least one aspect associated with music listening by CI users: spatial sound perception, multisensory music listening (live and recorded) with vibrotactile stimulation, as well as extra-timpanic air-coducted music exposure. The workshop allowed users to adjust several aspects of each experience, allowing them to tune the experience to their individual preference; their visit concluded with an interview.

Findings

The findings reveal significant variations in CI users' preferences, highlighting the need for personalized and customisable approaches in designing assistive musical technologies. The installations, ranging from vibrotactile devices to multisensory setups, showcase the complexity of CI users' experiences. Participants express preferences for certain tactile stimuli, identify challenges in music perception, and offer insights into the potential of integrating visual feedback. The study emphasizes the importance of an integrated and participatory research process, guiding designers toward creating more inclusive and effective musical interaction experiences for CI users.

Reflections

This collaboration with D. Cavdir and the series of workshops that resulted out of that was a pivotal point in my research, as it shifted focus from single use, to multi-use, multi-actuator tactile displays. Probably more important, it provided a fantastic opportunity to discuss music listening with CI users in an honest and intimate way. By providing a modular playground of actuators and signals, we have gained first hand insight into what works, what does not, and where should we focus our efforts. Looking back, it would have been great to continue this approach of inviting CI users into our design space and process on a regular basis.

Paper F: Design and Evaluation of a Multisensory Concert for Cochlear Implant Users

A continuation of the efforts started in paper E are described in this article. Specifically we have created a system of tactile displays based on the workshop results and organized a concert to evaluate their performance as multisensory augmentation devices. The concert was organized in partnership with the Royal Danish Academy for Music and was part of a larger event dedicated to music research for the hearing impaired population called *CoolHear*.

Motivation

The main goal of this study is to synthesize the knowledge obtained through the participatory workshop into one or more vibrotactile displays that could be used in a concert scenario by CI users. Our efforts resulted in a multi-actuator concert furniture that provided multichannel vibrotactile feedback to the participants, in order to enhance their perception and musical appreciation.

Methodology

The concert furniture was based on the requirements identified in the participatory workshops documented in paper E — a customisable system that accommodates for individual preference, and aids to the concert experience and perception by providing separate stream of stimulation for the bass and vocals. Furthermore, we used a combination of observations, interviews, and surveys to collect data about the concert furniture and analyze the user experience.

Findings

The concert revealed that tactile augmentation provides a pleasant and engaging experience for CI users, but also further highlighted that individual preference as well as musical genres/play style have a significant impact on the user satisfaction. Furthermore we have observed that multi-user vibrotactile displays need to be carefully calibrated not to cause discomfort for any of the participants. Lastly, the study highlights the benefits of visual feedback, such as gestural performance, to supplement the auditory and tactile channels.

Reflections

This was the first concert organized during my PhD research. It was part of a larger event where other CI oriented research was presented and demoed to the audience. It was a great satisfaction for all parties involved that it was a success — CI users loved it and so did the audiology technicians and the clinical staff. Nevertheless, there were several shortcomings that surfaced. Firs I exceeded the threshold for tactile comfort of some users. For the encore, I aimed to bring the tactile levels to the exact same values, but apparently I overshoot by a few millimeters on the mixer fader, resulting in participant noticing and mentioning it after. Secondly, my naive approach of conducting interview with CI users in a large concert hall was sub-optimal

— I overlooked the importance of a quiet place when discussing with hearing impaired people, neither of us could understand each-other well. With better planning and coordination between the academic team and the clinical one this could have been easily avoided. On a brighter note, the participants kept requesting more songs, to the point the band had to perform unplanned numbers — this was a pleasant surprise, and I am grateful for working with such a professional group of musicians that adapted new songs to the playing style discussed, while on the spotlight.

Paper G: A Concert-Based Study on Melodic Contour Identification Among Varied Hearing Profiles

In this article we documented a study on melodic contour identification organized in an ecologically valid setting — a concert. The melodies were performed live by a professional band consisting of players of: drums, bass, piano and accordion. Doing so we have created a library of melodies following four contours, that other researchers can use in their MCI research[1].

Motivation

The main goal of the study was to investigate how different hearing profiles affect the perception of melodic contours in a concert scenario, using live music as stimulation. Furthermore we wanted to compare the results with existing MCI studies that were usually conducted in a laboratory and to investigate how different combination of instruments affect the MCI scores, as instruments are rarely played solo in concerts.

Methodology

The study was organised as part of a free concert created for hearing impaired users, resulting in 43 people taking part, with 5 different hearing profiles: normal hearing, double hearing aids, bimodal users, single-side cochlear implanted and bilateral implanted. A total of 48 melodies were performed on 4 combination of instruments: drums + piano, drums + piano + bass, drums + accordion, and drums + accordion + bass. These combinations allowed to investigate the impact of bass as a masker, as well as the impact of spectro-temporal differences between the two lead instruments. The data were gathered using an online system accessed through the participants' smartphone, and were analysed using Bayesian logistic mixed-effects models.

Findings

The study revealed that the melodic contour identification performance for piano melodies was not affected by the electric bass masker, in contrast to

---

[1] https://github.com/razvysme/MCI-In-Concert

the accordion ones as well as previous studies. Furthermore, it was observed that accordion melodies were more difficult to follow than piano ones, especially with the bass-as-masker present. Interestingly, there was no significant difference between hearing aid users and other hearing impaired groups, questioning the assumption they would have a better performance than CI users.

Reflections

This was an important study for my research, as it *painted* the picture of hearing impairment and hearing performance, in terms of MCI. We had a good turnaround of 52 for a concert/study that had no agreement with any hearing impaired participants to show up. It was truly a pleasant surprise to see so many people interested in concerts for CI users. Armed with the lessons learned from the *CoolHear* event, we created a step-by-step protocol of collecting demographic and performance data. Our approach to assign animal names to participants in order to preserve anonymity was well received, and it create a positive mood in the audience, as it became a topic of discussion among attendees. Furthermore, it was surprising to see the high level of IT skills of the population with respect to the data gathering method — everyone managed to use the online voting system. Not everything was perfect though and looking back it would have been great to collect data about the CI processors individuals used, but from previous discussions with audiology technicians it was apparent that most don't know the brand name/model. Nevertheless, we should have tried. Furthermore, we could have been more transparent in the poster and event page that there will be a data gathering segment, as some participants felt *tricked* into coming to a concert, only to have to answer questions about short melodies. This ambiguity did not help on top of the fact that the MCI test lasted a bit over 20 minutes. Nevertheless, during the post-experiment interview all participants contacted were understanding and supportive of the goal of the study, but claimed that a clearer communication would have been appreciated.

Paper H: Effects of Music Training on Concert Experiences and Melodic Contour Identification for Hearing Impaired Individuals

This article is a continuation of the study of melodic contour identification in concert scenarios, and the focus is on a bespoke training system. The participants were offered a podcast-style guided listening training program for a period of 3 months. Half of the participants had access to the VAM presented in paper **B** and **C**, utilizing a different audio-to-tactile conversion algorithm.

Motivation

The research sought to determine whether a structured music training program could enhance the hearing performance and concert experience for HI individuals. Furthermore we wanted to explore the potential benefits of incorporating vibrotactile feedback in music training for CI users.

Methodology

This study involved two concerts with a six-month gap, during which participants underwent a home-training program focused on music listening. The training program, delivered via an online platform, comprised podcast-style lessons on four popular songs, emphasizing elements like instrumentation and melody. Participants used Bluetooth speakers for training, and half received a vibrotactile display for multisensory stimulation. The study used a parallel converging mixed-methods approach, gathering quantitative data through a MCI test during the concerts and qualitative data from post-concert interviews. The MCI test assessed participants' ability to identify three basic melodic contours(ascending, descending, undulating) in various instrumental combinations: drums + piano, drums + piano + bass, drums + accordion, and drums + accordion + bass.

Findings

The study revealed that the training program positively influenced participants' music perception and concert experiences. Although no participant completed the entire program, many reported enhanced music enjoyment and attention to details in songs. However, the length and cognitive load of the training sessions posed challenges as there was an noticeable difference between perceived and actual training effort. The vibrotactile display was not convincing enough to use since all participants reverted to auditory stimulation exclusively. MCI test results showed general improvement post-training, but no clear statistical difference between trained and untrained groups, possibly due to the small sample size. The study's ecological design presented challenges in controlling variables like participant interest and seating position. The research highlights the need for personalized auditory training approaches for CI users and indicates the potential benefits of such training in enhancing music appreciation in live scenarios.

Reflections

Continuing where the study in paper *G* left off, we had high expectations of a similarly sized audience and training participants. Unlike the first study, when everyone could participant, recruiting participants for the study raised several problems. Out of the people who agreed to evaluate the training programme, some got sick while others could not attend the given date. We even moved the concert date once in order to accommodate for the majority, but inevitable we had a lower turnaround. When it comes to the training program, it was well received and participants praised the pedagogical approach. Nevertheless, it was not an interactive system, therefore participants received no feedback regarding their performance. This is an aspect that could be improved, as it has been documented that feedback during training results in better hearing performance outcomes [50]. Similarly, we observed that no one completed the test, even though through the interviews some claimed that they even did some lessons multiple times (a statement that was not true). For the next iteration, we will focus on a more structured training scheme, with some interactive elements that would help users gauge their progress.

CHAPTER 5

Conclusion

During my doctoral research the main goals were to explore the use of vibrotactile stimulation in order to enhance the listening experience of cochlear implant users. My efforts have been focused on exploring how tactile displays can be used congruently with music. I strongly believe in the capabilities of CI users to re-gain their *"musical ear"*, at least to some degree, through engagement with music. There are several ways I can see the vibrotactile stimulation benefit the CI users in their music listening practice. What I don't see possible is to fully enjoy music relying exclusively on the tactile sense, due to it's psycho-physiological limitations. The key element in my research is the interplay between the two senses, with the tactile one supporting the music reaching the ear.

The overarching goal for my research is to promote music listening for CI users, in order to improve their quality of live, or well-being. Music has been a trusty companion for humans for as long as humanity itself, and to loose this friendship is something that saddens me. Nietzsche famously claimed that *"without music, life would be a mistake"*, but for CI users, the reality is frequently absent of music. It would be a mistake from us, the researchers, to accept this lack of music.

On a more practical note, this project was conducted wit a novel user-centered design approach, from a multidisciplinary perspective. This aligns well with the large spectra of needs observed in CI users, with respect to their music listening experiences but also their daily life. By adopting a multidisciplinary approach, my project looked into seemingly disconnected aspects of tactile stimulation, reported as problem areas in chapter 3, exploring the multifaceted world of vibrotactile augmentation of music. The following sections will provide a short summary of my and findings, organized around the aforementioned problem areas.

5.1 The VAM and Problem Area 1

The VAM discussed in papers **B** and **C** was used as a high-fidelity portable single actuator tactile display. It was designed as a research platform and it shows some promise. Although it may not match the precision of reference vibration exciters, its portability and affordability are notable strengths, which allowed us to fabricate enough to distribute for the home training programme described in paper **H**. Initial feedback on the VAM has been predominantly positive, particularly in solo listening scenarios. Users have appreciated it's ability to convey (some) musical information through tactile stimulation. Despite its design being as straightforward as possible, the VAM was not successful as a home-use tactile display for music, as all users who were given one as part of the training presented in paper **H** abandoned it at some point. To improve its usability, it could be re-designed as a wireless peripheral, but that would increase the user user's maintenance effort by requiring charging, and the technological complexity exponentially.

Even with a single actuator display like VAM, it is possible to create a spatial distribution of sensations based on the stimulation's frequency; as discussed in paper **B**. Unfortunately, there is no known mapping to between the frequency and amplitude of the tactile display and characteristics of the skin (e.g., the density of Pacinian receptors) that would provide an easy solution for implementing tactile spatialization without using multiple actuators. Furthermore, integrating this into musical applications without disrupting the original audio stimulus is complex and challenging.

A core issue with VAM, and indeed with audio-tactile technology in general, is the absence of a known optimal audio-to-tactile conversion algorithm for music. This gap means that translating audio signals into tactile sensations remains an open question, and comparing results between studies is problematic. Furthermore, there is no consensus on reporting and interpreting hardware descriptors of tactile displays, increasing the uncertainty when comparing tactile displays form literature. Frequency response is usually discussed, but even that is somehow limited as it's dependent on the displacement as well as the load of the capacity of the transducer.

Another significant constraint of the VAM is its limited ability to present complex signals like polyphonic ones or music or compositions with multiple instruments, and ensure that these would be distinguishable by the user. The analogy of mixdown taking from music recording is not directly translatable to tactile signals. This is because the tactile perception needs to operate in a more restricted frequency and dynamic range, and it's usually monophonic. Moreover, public familiarity with tactile music is often linked to experiences with loud music, making it challenging to introduce new *tactile mixdown* techniques effectively. This translates into a requirement of extensive longitudinal training that prototypes and evaluates new methods of representing complex stimuli through the tactile medium — a task that no research group attempted yet.

In conclusion, while the VAM presents exciting possibilities in the field of tactile audio technology, it also encounters substantial challenges. These include its limited applicability in home environments, the absence of a definitive audio-to-tactile conversion algorithm, and difficulties in accurately conveying complex musical compositions due to its single-actuator nature. This means that its design needs adjustment, should it be considered a viable tactile display for music listening. Nevertheless, it has been a useful research tool, and it could be used in further studies investigating single-actuator tactile displays.

5.2 Concert furniture and Problem Area 2

The vibrotactile concert furniture described in papers **E**, **G** builds upon the insights gained from the VAM. This furniture aims to enhance the concert experience for CI users by presenting congruent, multi-actuator tactile stimulation. Initial results suggest that the setup was appreciated, and we observed that once people were introduced to the concept by the concert host, they engaged with it.

One of the key learnings from the VAM project is that a single actuator can produce varied sensations depending on the area it stimulates and the pressure applied. This information has significant implications for the design of vibrotactile concert furniture, as the focus had to be on flexibility, and accommodation for various user needs. This is not only from an ergonomic point of view, but also regarding the tactile experience. Our goals was to provide multiple vibrotactile experiences with a limited number of actuators, by affording the user to alter the contact position and pressure on the skin.

Another influence the VAM had on the concert furniture design process was reflected from the understanding that there is no optimal conversion algorithm for translating audio signals into tactile feedback. As a result, we decided to design a simple yet effective conversion algorithm as a starting point, based on expert knowledge, with the goal to improve it with further research and user feedback.

An important aspect of designing vibrotactile displays discovered by using the furniture is the delicate balance with amplitude. Understanding the threshold of discomfort for vibrotactile actuation is essential, particularly considering the infinite combinations of frequency presented, areas actuated, and the contact area size afforded by the concert furniture. There is relatively little known about these thresholds, as most research is done with small scale actuators. This requires further testing and consideration of individual differences in sensory perception and tolerance, extending beyond simple discomfort levels with single actuators and/or signals. I can imagine an equivalent of the *Fletcher-Munson curves* for touch, but extended for multiple areas of the body.

The concert furniture was not designed in isolation, but part of a concert experience that I organized together with my peers. In total, we organized three events, iteratively improving various aspects of the concert. Through our research we discovered that, in order to provide a pleasant experience to CI users, we should not only augment the experience with tactile stimulation, but also take into consideration the musical preferences of the audience. This meant focusing on performing music familiar to the audience (preferrably pre-implant), with a simplified arrangement and a relatively low volume. One anecdotal conclusion we reached through dialog with several attendees as well as audio-verbal therapists therapists is that CI users do not want *new music* to be created accommodating for their hearing characteristic — they do not consider themselves deaf, therefore do not need special treatment. This preference underscores the importance of curating music that resonates with the audience's existing musical tastes and, probably more important, its memories.

Our experiences also highlighted the success of sing-alongs with the audience as introduction to the concerts. These interactive segments significantly enhanced audience engagement and enjoyment. This might be mostly true for a Danish audience, where sing along activities are widely practised in social gatherings [31]. However, it is important to note that this success might be influenced by cultural factors, whose extent of connotations is challenging to gauge and might vary significantly across different audiences. Therefore, understanding the cultural background and preferences of the audience is just as important as the music selection, in order to design an inclusive and enjoyable concert experience for CI users.

Another critical aspect observed was the sensitivity of CI users to the mixing of instruments in the concert. Achieving a balance that pleases everyone was difficult, as individual preferences can vary widely, and it's somewhat related to the relative area where they sit. Moreover, we observed that if the volume levels are kept relative low, CI users are comfortable with concerts that are longer than expected since it is known that CI users display a higher cognitive load when listening to music. This finding was iteratively understood, as we cautiously increased the concert duration in response to requests from attendees for longer concerts. This gradual adaptation allowed us to find an a good concert length that balances audience comfort with their desire for a more prolonged musical experience.

In conclusion, the development of vibrotactile concert furniture and the related musical experience is an evolving process, deeply influenced by audience feedback and preferences. The goal was to create an environment that is not only presenting a vibrotactile augmentation, but also musically rewarding for CI users. This requires a good understanding of their preferences for familiar music, cultural etiquette, instrument mixing, and concert duration.

5.3 Music training and Problem Area 3

The music training program described in paper **H** was designed using the VAM, with the extended knowledge learned from the concert furniture, particularly regarding the conversion algorithm used for translating audio signals into tactile feedback.

Our key finding was that the training improved the MCI performance of hearing impaired participants in the study, regardless of their hearing profile, but given the small sample size we would be careful suggesting generalization. What is more important is that, while MCI performance and music listening experience are largely uncorrelated, the participants in our study reported an improved perception of music listening performance, that is shown to have an effect on the enjoyment [45].

One aspect that was particularly appreciated in our training program was the pedagogical approach that focused on individual instruments in isolation before presenting the complete mixdown. This feature allows users to follow individual components of the music, enhancing their understanding and appreciation of complex musical compositions.

A novel feature of our training system was the accurate logging of training times, for each participant. We were able to capture and analyze the discrepancy between perceived training time and actual training time, revealing a notable mismatch. This suggests that the effort involved in the training might be higher than anticipated. This finding is important for calibrating further training lessons to better align with users' perceptions and capacities.

Another significant observation was regarding the optimal duration of training sessions. While CI users were able to attend concerts for extended periods, indicating a surprising tolerance for music in a concert setting, the same did not necessarily apply to training at home. It was observed that 15-minute training sessions might be too long for some users, suggesting a need for shorter, more focused training modules. This discrepancy highlights the difference in user engagement and tolerance levels in different environments and emphasizes the need for flexible and adaptable training durations.

In summary, the development of the music training device represents a different step forward in our efforts to enhance the musical experience for CI users. While there was a positive impact on the measured listening performance by all listeners, not all aspects of the training program were in line with our expectations, especially regarding the cognitive load associated with individual music training at home.

5.4 Vibrotactile augmentation of music for CI users

Wrapping it up, the field of vibrotactile augmentation for music is in an exploratory stage, marked by a lack of established methods and a reliance on frequent, low-scale prototype evaluations. One of the key challenges in this phase is the absence of clinical accuracy in these evaluations, which is crucial for developing reliable and effective solutions for CI users. The research, while promising, is still in its infancy, with much ground to cover before reaching a stage of maturity and standardization.

This does not mean that the technology is under-performing, quite the opposite. Tactile displays, in general, have been well-received, particularly when their design accounts for the users' requirement as well as the listening situation. Nevertheless, more research is necessary to create truly successful tactile displays. This is due to the already niche nature of the field and the specialized requirements of CI users. As a result, there is a relatively low interest from commercial institutions in investing in such technology, which could potentially slow down innovation and wider adoption.

Given the current landscape, I believe it to be more beneficial to initially focus on augmenting musical experiences in public spaces, such as concert halls or cinemas. This approach can serve as a stepping stone to gradually introduce the technology for private use. However, there is a risk that vibrotactile augmentation for CI users would be rendered obsolete by a breakthrough in cochlear stimulation method. While current indications suggest that no major advancements in CI technology are imminent in the short term, the field remains dynamic. The opposite is equally plausible, where tactile stimulation becomes an integral part of the hearing experience.

To summarize, the realm of vibrotactile augmentation for music for CI users is a field with significant potential but also faces numerous challenges. It requires a careful balancing act between technological innovation, user-centered design, commercial viability, and an understanding of the evolving landscape of auditory assistive technologies . As researchers in this field it's important to maintain a focus on the practical needs and experiences of CI users, while staying attuned to the current and future advancements in cochlear implantation and related technologies.

5.5 Future directions and emerging questions

So, what's next? The research done so far answered a few questions, and fortunately raised more, and I would not have it any other way. In the following lines I would like to formalize further questions related to vibrotactile augmentation of music, that have been, to some extent, raised during my project.

Tactile displays

Previous studies on tactile displays, including the one presented as part of this dissertation, rely on tactile displays, but the interpretation of the human body is different across research groups. For example, some focus on high density of actuators in a confided space like the research group led by M. Fletcher [34], other focus on whole body stimulation [115], or a combination of both. To my knowledge, there is no study available into the relation between number of actuators, the areas of the body stimulated and music listening experience, therefore I cannot help by wonder *What is the optimal number of actuators for audio-tactile experiences?*. Similarly, there are no comparative studies between existing mapping of music signals using multi-actuator displays, further complicating the previous question. A possible intermediary step towards answering these types of questions could be an *accurate mapping of the frequency sensitivity distribution over the entire human body, at various amplitude intervals, and contact areas*. This hypothetical three dimensional space would provide a window into the human sensing capabilities, and allow creative and controlled use of actuator combinations.

Audio-tactile integration for music

Another aspect, which is just beginning to be explored, concerns the salient features of music and their relationship to audio-tactile experiences. This is strongly related to the audio-to-tactile conversion algorithm discussion I was presenting earlier in my research, but it extents it towards investigating multisensory music perception and appreciation. I believe that there is a great opportunity to answer questions like: *"What musical features are most important to be presented as tactile stimulation, in multisensory music listening scenarios by CI users?"*, and probably more importantly *"How do different musical features transmitted as tactile stimulation contribute to the music experience?"*.

While such questions can be researched in laboratories, my next one revolves around vibrotactile concerts for CI users. More specifically, the evaluation of such events. As part of my work we have organized 3 concerts, with over 100 cumulative participants attending them, but we could not evaluate the concert experience in an objective way. This is primarily because our emphasis was on MCI and training. However, the question of *"How to evaluate vibrotactile concerts for CI users?"* was a re-occurring point of discussion, and it will likely guide the direction of my future research. A good understanding of the mechanisms underlying a concert experience for the hearing impaired would allow us, the researchers, to correctly identify *pressure points*, as well as attract interest from the funding organizations.

Music training for CI users

The last series of questions relate to training, and luckily there are other research groups who are feeling this is an important aspect of CI rehabilitation process. Nevertheless, through our hands-on experience by organizing concerts and creating the training program, we encountered many uncertainties that would be great further research questions. Most of them were practical ones like *"How does the set of musical features encompassed in the training set-list impact the success of the training for CI users?"*. Similar questions related to frequency and length of the training are also worth exploring.

We can consider the questions about tactile displays answered to a satisfactory level, but even without that being the case it would be beneficial to evaluate the impact of adding vibrotactile stimulation to the training, and more important what would be the best approach: should the same instrument be presented multisensory? If not, which one should be tactile? If yes, does it apply to complex stimuli as well? etc. Tightly related to all the questions about training is the aspect of evaluation of its success. This question can be approached from multiple perspectives, highlighting the complexity of the field we are working with. On one hand, it can be looked at from objective measurements in terms of music listening performance, but on the other hand, the success can be evaluated from the well-being point of view.

5.6 Conclusion

Introducing vibrotactile stimulation for music in the lives of CI users — be it for private or public usage is a beautiful challenge. This issue provided a distinctive opportunity for me to engage with applied research in practical settings, utilizing a pragmatic method to handle problems as they surfaced. Is it safe to assume that in 10 to 20 years, CI users will rely on vibrotactile stimulation for their musical experiences? It is difficult to predict, but I believe that even if the CIs as we know now them would not exist exist anymore, as they would have been surpassed by incredible assistive hearing devices, the knowledge presented throughout my doctoral research would have aided, even in a minuscule quantity to a better experience for the people I worked for. I am proud to have contributed to this field, and I am immensely excited to witness it grow and mature, alongside my professional journey.

CHAPTER 6

References

[1] S. C. Aker, H. Innes-Brown, K. F. Faulkner, M. Vatti, and J. Marozeau, "Effect of audio-tactile congruence on vibrotactile music enhancement," *The Journal of the Acoustical Society of America*, vol. 152, no. 6, pp. 3396–3409, 2022.

[2] I. Asimov, "Asimov's biographical encyclopedia of science and technology: the lives and achievements of 1195 great scientists from ancient times to the present, chronologically arranged," *(No Title)*, 1972.

[3] N. V. Atteveldt, M. M. Murray, G. Thut, and C. E. Schroeder, "Multisensory integration: Flexible use of general operations," 2014.

[4] H. Bastian, "The hawthorne effect: An old scientists' tale lingering 'in the gun-smoke of academic snipers.'," *Scientific American*, 2013.

[5] U. Baumann and A. Nobbe, "The cochlear implant electrode-pitch function," *Hearing Research*, vol. 213, 2006.

[6] D. Başkent, C. D. Fuller, J. J. Galvin, L. Schepel, E. Gaudrain, and R. H. Free, "Musician effect on perception of spectro-temporally degraded speech, vocal emotion, and music in young adolescents," *The Journal of the Acoustical Society of America*, vol. 143, 2018.

[7] D. Başkent and E. Gaudrain, "Musician advantage for speech-on-speech perception," *The Journal of the Acoustical Society of America*, vol. 139, 2016.

[8] J. Benesty, M. M. Sondhi, Y. Huang *et al.*, *Springer handbook of speech processing*. Springer, 2008, vol. 1.

[9] C. Bernard, J. Monnoyer, M. Wiertlewski, and S. Ystad, "Rhythm perception is shared between audio and haptics," *Scientific Reports*, vol. 12, 2022.

[10] D. M. Birnbaum and M. M. Wanderley, “A systematic approach to musical vibrotactile feedback,” 2007.

[11] T. Brochier, F. Guérit, J. M. Deeks, C. Garcia, M. Bance, and R. P. Carlyon, “Evaluating and comparing behavioural and electrophysiological estimates of neural health in cochlear implant users,” *JARO - Journal of the Association for Research in Otolaryngology*, vol. 22, 2021.

[12] S. J. Brockmeier, D. Fitzgerald, O. Searle, H. Fitzgerald, M. Grasmeder, S. Hilbig, K. Vermiere, M. Peterreins, S. Heydner, and W. Arnold, “The music perception test: A novel battery for testing music perception of cochlear implant users,” *Cochlear Implants International*, vol. 12, 2011.

[13] L. Bruns, D. Mürbe, and A. Hahne, “Understanding music with cochlear implants,” *Scientific Reports*, vol. 6, 2016.

[14] C. A. Buchman, M. T. Dillon, E. R. King, M. C. Adunka, O. F. Adunka, and H. C. Pillsbury, “Influence of cochlear implant insertion depth on performance,” *Otology & Neurotology*, vol. 35, 2014.

[15] W. Buyens, B. V. Dijk, M. Moonen, and J. Wouters, “Music mixing preferences of cochlear implant recipients: A pilot study,” *International Journal of Audiology*, vol. 53, 2014.

[16] ——, “Evaluation of a stereo music preprocessing scheme for cochlear implant users,” *Journal of the American Academy of Audiology*, vol. 29, 2018.

[17] M. T. Caldwell, P. Jiradejvong, and C. J. Limb, “Impaired perception of sensory consonance and dissonance in cochlear implant users,” *Otology and Neurotology*, vol. 37, 2016.

[18] C. Cappe, A. Thelen, V. Romei, G. Thut, and M. M. Murray, “Looming signals reveal synergistic principles of multisensory integration,” *Journal of Neuroscience*, vol. 32, no. 4, pp. 1171–1182, 2012.

[19] R. P. Carlyon and T. Goehring, “Cochlear implant research and development in the twenty-first century: A critical update,” 2021.

[20] M. V. Certo, G. D. Kohlberg, D. A. Chari, D. M. Mancuso, and A. K. Lalwani, “Reverberation time influences musical enjoyment with cochlear implants,” *Otology and Neurotology*, vol. 36, 2015.

[21] T. Chin and N. S. Rickard, “Beyond positive and negative trait affect: Flourishing through music engagement,” *Psychology of Well-Being*, vol. 4, 2014.

[22] K. Cieśla, T. Wolak, A. Lorens, B. Heimler, H. Skarżyński, and A. Amedi, “Immediate improvement of speech-in-noise perception through multisensory stimulation via an auditory to tactile sensory substitution,” *Restorative Neurology and Neuroscience*, vol. 37, pp. 155–166, 2019.

[23] N. L. Cohen, “Cochlear implant candidacy and surgical considerations,” vol. 9, 2004.

[24] E. Diener and K. Ryan, “Subjective well-being: A general overview,” *South African Journal of Psychology*, vol. 39, 2009.

[25] A. Dieter, D. Keppeler, and T. Moser, “Towards the optical cochlear implant: optogenetic approaches for hearing restoration,” *EMBO Molecular Medicine*, vol. 12, 2020.

[26] G. A. Dingle, L. S. Sharman, Z. Bauer, E. Beckman, M. Broughton, E. Bunzli, R. Davidson, G. Draper, S. Fairley, C. Farrell, L. M. Flynn, S. Gomersall, M. Hong, J. Larwood, C. Lee, J. Lee, L. Nitschinsk, N. Peluso, S. E. Reedman, D. Vidas, Z. C. Walter, and O. R. L. Wright, "How do music activities affect health and well-being? a scoping review of studies examining psychosocial mechanisms," 2021.

[27] P. J. Donnelly, B. Z. Guo, and C. J. Limb, "Perceptual fusion of polyphonic pitch in cochlear implant users," *The Journal of the Acoustical Society of America*, vol. 126, no. 5, pp. EL128–EL133, 2009.

[28] W. R. Drennan, J. J. Oleson, K. Gfeller, J. Crosson, V. D. Driscoll, J. H. Won, E. S. Anderson, and J. T. Rubinstein, "Clinical evaluation of music perception, appraisal and experience in cochlear implant users," *International Journal of Audiology*, vol. 54, 2015.

[29] V. D. Driscoll, J. Oleson, D. Jiang, and K. Gfeller, "Effects of training on recognition of musical instruments presented through cochlear implant simulations," *Journal of the American Academy of Audiology*, vol. 20, 2009.

[30] G. Dritsakis, R. M. van Besouw, P. Kitterick, and C. A. Verschuur, "A music-related quality of life measure to guide music rehabilitation for adult cochlear implant users," *American Journal of Audiology*, vol. 26, 2017.

[31] O. Ekholm, K. Juel, and L. O. Bonde, "Associations between daily musicking and health: Results from a nationwide survey in denmark," *Scandinavian journal of public health*, vol. 44, no. 7, pp. 726–732, 2016.

[32] M. O. Ernst and M. S. Banks, "Humans integrate visual and haptic information in a statistically optimal fashion," *Nature*, vol. 415, no. 6870, pp. 429–433, 2002.

[33] J. Fares, "A multi-level typology for stakeholder influence: A systematic literature review using the structural approach," *European Management Journal*, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S026323732300107X

[34] M. D. Fletcher, "Can haptic stimulation enhance music perception in hearing-impaired listeners?" 2021.

[35] M. D. Fletcher, N. Thini, and S. W. Perry, "Enhanced pitch discrimination for cochlear implant users with a new haptic neuroprosthetic," *Scientific Reports*, vol. 10, 2020.

[36] M. Fletcher, "Using haptic stimulation to enhance auditory perception in hearing-impaired listeners," *Expert Review of Medical Devices*, vol. 18, pp. 63–74, 2021.

[37] M. Fletcher, A. Hadeedi, T. Goehring, and S. Mills, "Electro-haptic enhancement of speech-in-noise performance in cochlear implant users," *Scientific Reports*, vol. 9, 2019.

[38] M. Fletcher, S. Mills, and T. Goehring, "Vibro-tactile enhancement of speech intelligibility in multi-talker noise for simulated cochlear implant listening," *Trends in Hearing*, vol. 22, 2018.

[39] M. Fletcher, H. Song, and S. Perry, "Electro-haptic stimulation enhances speech recognition in spatially separated noise for cochlear implant users," *Scientific Reports*, vol. 10, 2020.

[40] F. Fontana, R. Paisa, R. Ranon, and S. Serafin, "Multisensory plucked instrument modeling in unity3d: From keytar to accurate string prototyping," *Applied Sciences*, vol. 10, no. 4, p. 1452, 2020.

[41] M. J. Forgeard, E. Jayawickreme, M. L. Kern, and M. E. Seligman, "Doing the right thing: Measuring wellbeing for public policy," *International journal of wellbeing*, vol. 1, no. 1, 2011.

[42] S. L. Fowler, H. Calhoun, and A. D. Warner-Czyz, "Music perception and speech-in-noise skills of typical hearing and cochlear implant listeners," *American Journal of Audiology*, vol. 30, 2021.

[43] E. Frid and H. Lindetorp, "Haptic music - exploring whole-body vibrations and tactile sound for a multisensory music installation," *Proceedings of the Sound and Music Computing Conferences*, vol. 2020-June, pp. 68–75, 2020.

[44] X. Fu and L. Riecke, "Effects of continuous tactile stimulation on auditory-evoked cortical responses depend on the audio-tactile phase," *NeuroImage*, vol. 274, p. 120140, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811923002914

[45] C. Fuller, R. Free, B. Maat, and D. Başkent, "Self-reported music perception is related to quality of life and self-reported hearing abilities in cochlear implant users," *Cochlear Implants International*, vol. 23, 2022.

[46] C. D. Fuller, J. J. Galvin, B. Maat, R. H. Free, and D. Başkent, "The musician effect: Does it persist under degraded pitch conditions of cochlear implant simulations?" *Frontiers in Neuroscience*, 2014.

[47] C. D. Fuller, I. John J. Galvin, B. Maat, D. Başkent, and R. H. Free, "Comparison of two music training approaches on music and speech perception in cochlear implant users," *Trends in Hearing*, vol. 22, no. 9999, p. 2331216518765379, 2018, pMID: 29621947. [Online]. Available: https://doi.org/10.1177/2331216518765379

[48] J. J. Galvin, Q. J. Fu, and G. Nogaki, "Melodic contour identification by cochlear implant listeners," *Ear and Hearing*, vol. 28, 2007.

[49] J. J. Galvin, Q. J. Fu, and R. V. Shannon, "Melodic contour identification and music perception by cochlear implant users," vol. 1169, 2009.

[50] J. Galvin, E. Eskridge, S. Oba, and Q. J. Fu, "Melodic contour identification training in cochlear implant users with and without a competing instrument," *Seminars in Hearing*, vol. 33, 2012.

[51] B. J. Gantz, C. Dunn, E. Walker, T. V. Voorst, S. Gogel, and M. Hansen, "Outcomes of adolescents with a short electrode cochlear implant with preserved residual hearing," vol. 37, 2016.

[52] K. Gfeller, A. Buzzell, V. Driscoll, B. Kinnaird, and J. Oleson, "The impact of voice range and instrumental background accompaniment on recognition of song lyrics," *7Th Asia Pacific Symposium on Cochlear Implants and Related Sciences - Apsci*, 2009.

[53] K. Gfeller, A. Christ, J. F. Knutson, S. Witt, K. T. Murray, and R. S. Tyler, "Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients," *Journal of the American Academy of Audiology*, vol. 11, 2000.

[54] K. Gfeller, D. Jiang, J. J. Oleson, V. Driscoll, C. Olszewski, J. F. Knutson, C. Turner, and B. Gantz, "The effects of musical and linguistic components in recognition of real-world musical excerpts by cochlear implant recipients and normal-hearing adults," *Journal of Music Therapy*, vol. 49, 2012.

[55] K. Gfeller, C. Turner, J. Oleson, X. Zhang, B. Gantz, R. Froman, and C. Olszewski, "Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise," *Ear and Hearing*, vol. 28, 2007.

[56] K. Gfeller, S. Witt, M. Adamek, M. Mehr, J. Rogers, J. Stordahl, and S. Ringgenberg, "Effects of training on timbre recognition and appraisal by postlingually deafened cochlear implant recipients," *Journal of the American Academy of Audiology*, vol. 13, 2002.

[57] K. Gfeller, S. Witt, J. Stordahl, M. Mehr, and G. Woodworth, "The effects of training on melody recognition and appraisal by adult cochlear implant recipients," *JARA*, vol. XXXIII, 2000.

[58] K. E. Gfeller, C. Olszewski, C. Turner, B. Gantz, and J. Oleson, "Music perception with cochlear implants and residual hearing," vol. 11, 2006.

[59] A. A. Ghazanfar and C. E. Schroeder, "Is neocortex essentially multisensory?" *Trends in cognitive sciences*, vol. 10, no. 6, pp. 278–285, 2006.

[60] M. Gilbert, P. Jiradejvong, and C. Limb, "Effect of compression on musical sound quality in cochlear implant users," *Ear and hearing*, vol. 40, 2019.

[61] H. Gillmeister and M. Eimer, "Tactile enhancement of auditory detection and perceived loudness," *Brain Research*, vol. 1160, 2007.

[62] M. Giordano and M. M. Wanderley, "Perceptual and technological issues in the design of vibrotactile-augmented interfaces for music technology and media," in *Haptic and Audio Interaction Design: 8th International Workshop, HAID 2013, Daejeon, Korea, April 18-19, 2013, Revised Selected Papers 8*. Springer, 2013, pp. 89–98.

[63] E. B. Goldstein, *Sensation and perception*. Wadsworth/Thomson Learning, 1989.

[64] J. S. Golub, J. H. Won, W. R. Drennan, T. D. Worman, and J. T. Rubinstein, "Spectral and temporal measures in hybrid cochlear implant users: On the mechanism of electroacoustic hearing benefits," *Otology and Neurotology*, vol. 33, 2012.

[65] F. R. Goodman, D. J. Disabato, T. B. Kashdan, and S. B. Kauffman, "Measuring well-being: A comparison of subjective well-being and perma," *The Journal of Positive Psychology*, vol. 13, no. 4, pp. 321–332, 2018.

[66] F. Grani, R. Paisa, J. S. Banas, I. Vogiatzoglou, and S. Serafin, "Design and evaluation of a gesture driven wavefield synthesis auditory game," in *Proc. of NIME*, 2016.

[67] E. Gunther and S. O'Modhrain, "Cutaneous grooves: Composing for the sense of touch," *International Journal of Phytoremediation*, vol. 21, 2003.

[68] L. R. Hallberg, A. Ringdahl, A. Holmes, and C. Carver, "Psychological general well-being (quality of life) in patients with cochlear implants: Importance of social environment and age," *International Journal of Audiology*, vol. 44, pp. 706–711, 12 2005.

[69] T. Hays, "Well-being in later life through music," *Australasian Journal on Ageing*, vol. 24, 2005.

[70] T. Hays and V. Minichiello, “The contribution of music to quality of life in older people: An australian qualitative study,” *Ageing and Society*, vol. 25, 2005.

[71] J. B. Hinderink, J. P. Brokx, and P. van den Broek, “Performance of prelingually versus postlingually deafened patients using single or multichannel cochlear implants,” *COCHLEAR IMPLANTS IN ADULTS*, p. 61, 1995.

[72] J. B. Hinderink, P. F. Krabbe, and P. V. D. Broek, “Development and application of a health-related quality-of-life instrument for adults with cochlear implants: The nijmegen cochlear implant questionnaire,” *Otolaryngology - Head and Neck Surgery*, vol. 123, 2000.

[73] A. Hogan, G. Hawthorne, L. Kethel, E. Giles, K. White, M. Stewart, B. Plath, and C. Code, “Health-related quality-of-life outcomes from adult cochlear implantation: a cross-sectional survey,” *Cochlear Implants International*, vol. 2, no. 2, pp. 115–128, 2001.

[74] L. K. Holden, J. B. Firszt, R. M. Reeder, R. M. Uchanski, N. Y. Dwyer, and T. A. Holden, “Factors affecting outcomes in cochlear implant recipients implanted with a perimodiolar electrode array located in scala tympani,” *Otology and Neurotology*, vol. 37, pp. 1662–1668, 2016. [Online]. Available: https://doi.org/10.1097/MAO.0000000000001241

[75] J. Huang, T. Lu, B. Sheffield, and F.-G. Zeng, “Electro-tactile stimulation enhances cochlear-implant melody recognition: Effects of rhythm and musical training,” *Ear and Hearing*, vol. 41, pp. 106–113, 2018.

[76] F. A. Huppert and T. T. So, “Flourishing across europe: Application of a new conceptual framework for defining well-being,” *Social Indicators Research*, vol. 110, 2013.

[77] R. Jack, A. Mcpherson, and T. Stockman, “Designing tactile musical devices with and for deaf users: a case study,” *Proc. of the International Conference on the Multimedia Experience of Music*, 2015.

[78] C. Kaernbach and C. Bering, “Exploring the temporal mechanism involved in the pitch of unresolved harmonics,” *The Journal of the Acoustical Society of America*, vol. 110, no. 2, pp. 1039–1048, 2001.

[79] J. Kanebako and K. Minamizawa, “Vibgrip++: Haptic device allows feeling the music for hearing impaired people,” *Lecture Notes in Electrical Engineering*, vol. 432, pp. 449–452, 2018.

[80] R. Kang, G. L. Nimmons, W. Drennan, J. Longnion, C. Ruffin, K. Nie, J. H. Won, T. Worman, B. Yueh, and J. Rubinstein, “Development and validation of the university of washington clinical assessment of music perception test,” *Ear and Hearing*, vol. 30, 2009.

[81] M. Karam and D. Fels, “Designing a model human cochlea: Issues and challenges in crossmodal audio-haptic displays,” *Proceedings of the 2008 Ambi-Sys Workshop on Haptic User Interfaces in Ambient Media Systems, HAS 2008*, 2008.

[82] M. Karam, F. Russo, and D. Fels, “Designing the model human cochlea: An ambient crossmodal audio-tactile display,” *IEEE Transactions on Haptics*, vol. 2, pp. 160–169, 2009.

[83] C. Karoui, C. James, P. Barone, D. Bakhos, M. Marx, and O. Macherey, "Searching for the sound of a cochlear implant: Evaluation of different vocoder parameters by cochlear implant users with single-sided deafness," *Trends in Hearing*, vol. 23, p. 2331216519866029, 2019, pMID: 31533581.

[84] D. R. Ketten, M. W. Skinner, G. Wang, M. W. Vannier, G. A. Gates, and J. G. Neely, "In vivo measures of cochlear length and insertion depth of nucleus cochlear implant electrode arrays." *The Annals of otology, rhinology & laryngology. Supplement*, vol. 175, 1998.

[85] M. Knobloch, J. L. Verhey, M. Ziese, M. Nitschmann, C. Arens, and M. Böckmann-Barthel, "Musical harmony in electric hearing," *Music Perception*, vol. 36, 2018.

[86] J. Kohlmorgen, S. Lemm, G. Ratsch, and K.-R. Muller, "Analysis of nonstationary time series by mixtures of self-organizing predictors," in *Neural Networks for Signal Processing X. Proceedings of the 2000 IEEE Signal Processing Society Workshop (Cat. No. 00TH8501)*, vol. 1. IEEE, 2000, pp. 85–94.

[87] K.-U. Kyung, M. Ahn, D.-S. Kwon, and M. A. Srinivasan, "Perceptual and biomechanical frequency response of human skin implication for design of tactile displays," in *Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, ser. WHC '05. USA: IEEE Computer Society, 2005, p. 96–101. [Online]. Available: https://doi.org/10.1109/WHC.2005.105

[88] L. Lassaletta, A. Castro, M. Bastarrica, R. Pérez-Mora, R. Madero, J. D. Sarriá, and J. Gavilán, "Does music perception have an impact on quality of life following cochlear implantation?" *Acta Oto-Laryngologica*, vol. 127, 2007.

[89] E. F. Lau, C. Phillips, and D. Poeppel, "A cortical network for semantics:(de) constructing the n400," *Nature reviews neuroscience*, vol. 9, no. 12, pp. 920–933, 2008.

[90] P. Laukka, "Uses of music and psychological well-being among the elderly," *Journal of Happiness Studies*, vol. 8, 2007.

[91] Y. Y. Lee, M. F. Chan, and E. Mok, "Effectiveness of music intervention on the quality of life of older people," *Journal of Advanced Nursing*, vol. 66, 2010.

[92] J. P. Lerousseau, C. Hidalgo, and D. Schön, "Musical training for auditory rehabilitation in hearing loss," *Journal of Clinical Medicine*, vol. 9, 2020.

[93] C. J. Limb and A. T. Roy, "Technological, biological, and acoustical constraints to music perception in cochlear implant users," 2014.

[94] C. Y. Lo, C. M. McMahon, V. Looi, and W. F. Thompson, "Melodic contour training and its effect on speech in noise, consonant discrimination, and prosody perception for cochlear implant recipients," *Behavioural Neurology*, vol. 2015, 2015.

[95] P. C. Loizou, "Speech processing in vocoder-centric cochlear implants," *Cochlear and brainstem implants*, vol. 64, pp. 109–143, 2006.

[96] V. Looi, H. McDermott, C. McKay, and L. Hickson, "Music perception of cochlear implant users compared with that of hearing aid users," *Ear and Hearing*, vol. 29, 2008.

[97] C. J. Maillet, R. S. Tyler, and H. Jordan, “Change in the quality of life of adult cochlear implant patients.” *The Annals of otology, rhinology & laryngology. Supplement*, vol. 165, pp. 31–48, 1995.

[98] J. Makous, R. Friedman, and C. Vierck, “A critical band filter in touch,” *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 15, no. 4, p. 2808—2818, April 1995.

[99] J. Marozeau, “Audio-tactile installations to augment music perception for people with hearing impairment,” in *24th International Congress on Acoustics*, 2022.

[100] ——, “Why people with a cochlear implant listen to music,” vol. 12631 LNCS, 2021.

[101] K. C. G. A. J. K. W. S. M. Maureen, “The effects of familiarity and complexity on appraisal of complex songs by cochlear implant recipients and normal hearing adults.” *Journal of music therapy*, vol. 40, pp. 78–112, 2003.

[102] S. McAdams, “Musical timbre perception,” *The psychology of music*, pp. 35–67, 2013.

[103] L. McArton and R. Mantie, “Music, health and well-being in ijcm articles: An integrative review,” *International Journal of Community Music*, vol. 16, 2023.

[104] J. M. McGinnis, E. Malphrus, D. Blumenthal *et al.*, “Vital signs: core metrics for health and health care progress,” 2015.

[105] T. R. McRackan, B. N. Hand, C. A. Velozo, J. R. Dubno, J. S. Golub, E. P. Wilkinson, D. Mills, J. P. Carey, N. Vorasubin, V. Brunk, M. L. Carlson, C. L. Driscoll, D. P. Sladen, E. L. Camposeo, M. A. Holcomb, P. R. Lambert, T. A. Meyer, C. Thomas, A. C. Moberly, N. H. Blevins, J. B. Larky, R. P. Herzano, M. E. Hoffer, S. M. Prentiss, R. N. Samy, S. P. Gubbels, J. Brant, J. B. Hunter, B. Isaacson, J. W. Kutz, R. K. Gurgel, D. M. Zeitler, C. A. Buchman, J. B. Firszt, R. H. Gifford, D. S. Haynes, and R. F. Labadie, “Cochlear implant quality of life (ciqol): Development of a profile instrument (ciqol-35 profile) and a global measure (ciqol-10 global),” *Journal of Speech, Language, and Hearing Research*, vol. 62, 2019.

[106] S. Merchel and M. E. Altinsoy, “Auditory-tactile music perception,” vol. 19, 2013.

[107] J. C. Middlebrooks and R. L. Snyder, “Auditory prosthesis with a penetrating nerve array,” *Journal for the Association for Research in Otolaryngology*, vol. 8, pp. 258–279, 2007.

[108] L. Migirov, J. Kronenberg, and Y. Henkin, “Self-reported listening habits and enjoyment of music among adult cochlear implant recipients,” *Annals of Otology, Rhinology & Laryngology*, vol. 118, no. 5, pp. 350–355, 2009, pMID: 19548384.

[109] D. Miranda and M. Claes, “Music listening, coping, peer affiliation and depression in adolescence,” *Psychology of Music*, vol. 37, no. 2, pp. 215–233, 2009. [Online]. Available: https://doi.org/10.1177/0305735608097245

[110] I. J. Moon, H. Byun, S.-y. Woo, G.-Y. Gwak, S. H. Hong, W.-H. Chung, and Y.-S. Cho, “Factors associated with age-related hearing impairment: A retrospective cohort study,” *Medicine*, vol. 94, no. 43, 2015.

[111] J. W. Morley and M. J. Rowe, “Perceived pitch of vibrotactile stimuli: effects of vibration amplitude, and implications for vibration frequency coding.” *The Journal of Physiology*, vol. 431, 1990.

[112] M. M. Murray, S. Molholm, C. M. Michel, D. J. Heslenfeld, W. Ritter, D. C. Javitt, C. E. Schroeder, and J. J. Foxe, “Grabbing your ear: rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment,” *Cerebral cortex*, vol. 15, no. 7, pp. 963–974, 2005.

[113] S. Nanayakkara, L. Wyse, and E. Taylor, “The haptic chair as a speech training aid for the deaf,” *Proceedings of the 24th Australian Computer-Human Interaction Conference, OzCHI 2012*, pp. 405–410, 2012.

[114] S. Nanayakkara, E. Taylor, L. Wyse, and S. Ong, “An enhanced musical experience for the deaf: Design and evaluation of a music display and a haptic chair,” 04 2009, pp. 337–346.

[115] S. C. Nanayakkara, L. Wyse, S. H. Ong, and E. A. Taylor, “Enhancing musical experience for the hearing-impaired using visual and haptic displays,” *Human-Computer Interaction*, vol. 28, 2013.

[116] D. A. Nelson, J. L. Schmitz, G. S. Donaldson, N. F. Viemeister, and E. Javel, “Intensity discrimination as a function of stimulus level with electric stimulation,” *The Journal of the Acoustical Society of America*, vol. 100, 1996.

[117] R. Okazaki, T. Hachisu, M. Sato, S. Fukushima, V. Hayward, and H. Kajimoto, “Judged consonance of tactile and auditory frequencies,” 04 2013, pp. 663–666.

[118] R. Okazaki, H. Kajimoto, and V. Hayward, “Vibrotactile stimulation can affect auditory loudness: A pilot study,” 06 2012, pp. 103–108.

[119] R. Paisa, C. Erkut, and S. Serafin, “Electronic armonica: a tangible musical interface inspired by the glass harmonica,” in *Proceedings of the Audio Mostly 2016*, 2016, pp. 251–257.

[120] R. Paisa and D. Overholt, “Enhancing the expressivity of the sensel morph via audio-rate sensing,” in *NIME 2019 New Interfaces for Musical Expression conference*, 2019, pp. 298–302.

[121] R. Paisa and S. Serafin, “Recreating an instrument inspired by the glass harmonica using fabrication techniques and physical models,” in *15th Sound and Music Computing Conference, SMC 2018*. Sound and Music Computing Network, 2018, pp. 239–242.

[122] S. Papetti and C. Saitis, *Musical Haptics: Introduction*. Springer, Cham, 2018.

[123] L. S. Pardue, K. Buys, M. Edinger, D. Overholt, and A. P. McPherson, “Separating sound from source: sonic transformation of the violin through electrodynamic pickups and acoustic actuation,” 2019.

[124] C. V. Parise, C. Spence, and M. O. Ernst, “When correlation implies causation in multisensory integration,” *Current Biology*, vol. 22, no. 1, pp. 46–49, 2012.

[125] A. Passalenti, R. Paisa, N. C. Nilsson, N. S. Andersson, F. Fontana, R. Nordahl, and S. Serafin, “No strings attached: Force and vibrotactile feedback in a virtual guitar simulation,” in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2019, pp. 1116–1117.

[126] W. Penfield and E. Boldrey, “Somatic motor and sensory representation in the cerebral cortex of a man as studied by electrical stimulation,” *Brain*, vol. 60, no. 4, pp. 389–443, 12 1937. [Online]. Available: https://doi.org/10.1093/brain/60.4.389

[127] B. Petersen, M. V. Mortensen, M. Hansen, and P. Vuust, "Singing in the key of life: A study on effects of musical ear training after cochlear implantation." *Psychomusicology: Music, Mind, and Brain*, vol. 22, 2012.

[128] B. Petry, T. Illandara, D. Elvitigala, and S. Nanayakkara, "Supporting rhythm activities of deaf children using music-sensory-substitution systems," *Conference on Human Factors in Computing Systems - Proceedings*, vol. 2018-April, 2018.

[129] L. Ping, M. Yuan, and H. Feng, "Musical pitch discrimination by cochlear implant users," *Annals of Otology, Rhinology & Laryngology*, vol. 121, no. 5, pp. 328–336, 2012.

[130] J. Pinyon, S. Tadros, K. Froud, A. Wong, I. Tompson, E. Crawford, M. Ko, R. Morris, M. Klugmann, and G. Housley, "Close-field electroporation gene delivery using the cochlear implant electrode array enhances the bionic ear. sci transl med 6: 233ra54-233ra54," 2014.

[131] H. Pongrac, "Vibrotactile perception: examining the coding of vibrations and the just noticeable difference under various conditions," *Multimedia systems*, vol. 13, no. 4, pp. 297–307, 2008.

[132] J. C. Prentice, A. B. Frakt, and S. D. Pizer, "Metrics that matter," *Journal of general internal medicine*, vol. 31, pp. 70–73, 2016.

[133] C. Prevoteau, S. Y. Chen, and A. K. Lalwani, "Music enjoyment with cochlear implantation," *Auris Nasus Larynx*, vol. 45, no. 5, pp. 895–902, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S038581461730679X

[134] D. Purves, G. Augustine, D. Fitzpatrick, L. Katz, A.-S. LaMantia, J. McNamara, and M. Williams, "Neuroscience. 2nd edition," 2001.

[135] C. M. Reed, H. Z. Tan, Y. Jiao, Z. D. Perez, and E. C. Wilson, "Identification of words and phrases through a phonemic-based haptic display," *ACM Transactions on Applied Perception*, vol. 18, 2021.

[136] B. Remache-Vinueza, A. Trujillo-León, M. Zapata, F. Sarmiento-Ortiz, and F. Vidal-Verdú, "Audio-tactile rendering: A review on technology and methods to convey musical information through the sense of touch," 2021.

[137] S. H. Rembar, O. Lind, P. Romundstad, and A. S. Helvik, "Psychological well-being among cochlear implant users: A comparison with the general population," *Cochlear Implants International*, vol. 13, 2012.

[138] L. Riecke, S. Snipes, S. van Bree, A. Kaas, and L. Hausfeld, "Audio-tactile enhancement of cortical speech-envelope tracking," *NeuroImage*, vol. 202, 2019.

[139] F. E. Rizkou, O. Benhoummad, S. Rochd, Y. Lakhdar, Y. Rochdi, and A. Raji, "Cochlear implant outcomes: Quality of life in prelingually deafened, late-implanted patients," *Otology & Neurotology*, vol. 45, no. 1, pp. e24–e27, 2024.

[140] T. Ro, J. Hsu, N. Yasar, C. Elmore, and M. Beauchamp, "Sound enhances touch perception," *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, vol. 195, pp. 135–43, 04 2009.

[141] K. Robinson, S. Gatehouse, and G. G. Browning, "Measuring patient benefit from otorhinolaryngological surgery and therapy," *Annals of Otology, Rhinology & Laryngology*, vol. 105, no. 6, pp. 415–422, 1996.

[142] A. T. Roy, P. Jiradejvong, C. Carver, and C. J. Limb, "Assessment of sound quality perception in cochlear implant users during music listening," *Otology & Neurotology*, vol. 33, no. 3, pp. 319–327, 2012.

[143] F. A. Russo, P. Ammirante, and D. I. Fels, "Vibrotactile discrimination of musical timbre," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 38, 2012.

[144] C. Saitis, H. Järveläinen, and C. Fritz, *The role of haptic cues in musical instrument quality perception*. Springer International Publishing, 2018.

[145] E. W. Sargent, B. Herrmann, C. S. Hollenbeak, and A. E. Bankaitis, "The minimum speech test battery in profound unilateral hearing loss," *Otology & neurotology*, vol. 22, no. 4, pp. 480–486, 2001.

[146] D. K. Sarko, A. R. Nidiffer, A. R. Powers III, D. Ghose, A. Hillock-Dunn, M. C. Fister, J. Krueger, and M. T. Wallace, "Spatial and temporal features of multisensory processes," *The neural bases of multisensory processes*, 2012.

[147] G. Satell, *Mapping innovation: A playbook for navigating a disruptive age*. McGraw-Hill Education New York, NY, 2017, vol. 21.

[148] M. Schürmann, G. Caetano, V. Jousmäki, and R. Hari, "Hands help hearing: Facilitatory audiotactile interaction at low sound-intensity levels," *The Journal of the Acoustical Society of America*, vol. 115, 2004.

[149] P. R. Seitz, "French origins of the cochlear implant," *Cochlear Implants International*, vol. 3, 2002.

[150] M. Seligman, "Perma and the building blocks of well-being," *Journal of Positive Psychology*, vol. 13, 2018.

[151] R. V. Shannon, "Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics." *Hearing research*, vol. 11, no. 2, pp. 157–189, 1983.

[152] A. Sharp, M.-S. Houde, M. Maheu, I. Ibrahim, and F. Champoux, "Improved tactile frequency discrimination in musicians," *Experimental Brain Research*, vol. 237, pp. 1575–1580, 2019.

[153] S. Shin, C. Oh, and H. Shin, "Tactile tone system: A wearable device to assist accuracy of vocal pitch in cochlear implant users," 2020.

[154] N. F. A. Shukor, J. Lee, Y. J. Seo, and W. Han, "Efficacy of music training in hearing aid and cochlear implant users: A systematic review and meta-analysis," *Clinical and Experimental Otorhinolaryngology*, vol. 14, pp. 15–28, 2 2021.

[155] M. C. Sierra, J. Brunskog, and J. Marozeau, "An audio-tactile artinstallation for hearing impaired people," in *Proceedings of the 2nd Nordic Sound and Music Conference, Online*, 2021, pp. 11–12.

[156] C. Solé, M. Mercadal-Brotons, S. Gallego, and M. Riera, "Contributions of music to aging adults' quality of life," *Journal of Music Therapy*, vol. 47, 2010.

[157] D. L. Sorkin and C. A. Buchman, "Cochlear implant access in six developed countries," vol. 37, 2016.

[158] F. Sorrentino, F. Gheller, N. Favaretto, L. Franz, E. Stocco, D. Brotto, and R. Bovo, "Music perception in adult patients with cochlear implant," 2020.

[159] S. Srivastava, "Women in workforce: Work and family conflict," *Management and labour studies*, vol. 32, no. 4, pp. 411–421, 2007.

[160] B. E. STEIN and M. A. MEREDITH, “Multisensory integration: Neural and behavioral solutions for dealing with stimuli from different sensory modalities,” *Annals of the New York Academy of Sciences*, vol. 608, 1990.

[161] B. E. Stein and T. R. Stanford, “Multisensory integration: Current issues from the perspective of the single neuron,” 2008.

[162] B. Stein, P. Stein, and M. Meredith, *The Merging of the Senses*, ser. A Bradford book. MIT Press, 1993. [Online]. Available: https://books.google.se/books?id=uCV9QgAACAAJ

[163] R. A. Stevenson and M. T. Wallace, “Multisensory temporal integration: task and stimulus dependencies,” *Experimental brain research*, vol. 227, pp. 249–261, 2013.

[164] Y. Suzuki, H. Takeshima, and K. Kurakata, “Revision of iso 226” normal equal-loudness-level contours” from 2003 to 2023 edition: The background and results,” *Acoustical Science and Technology*, vol. 45, no. 1, pp. 1–8, 2024.

[165] A. C. Tricco, E. Lillie, W. Zarin, K. K. O’Brien, H. Colquhoun, D. Levac, D. Moher, M. D. Peters, T. Horsley, L. Weeks, S. Hempel, E. A. Akl, C. Chang, J. McGowan, L. Stewart, L. Hartling, A. Aldcroft, M. G. Wilson, C. Garritty, S. Lewin, C. M. Godfrey, M. T. MacDonald, E. V. Langlois, K. Soares-Weiser, J. Moriarty, T. Clifford, Özge Tunçalp, and S. E. Straus, “Prisma extension for scoping reviews (prisma-scr): Checklist and explanation,” *Annals of Internal Medicine*, vol. 169, no. 7, pp. 467–473, 2018, pMID: 30178033. [Online]. Available: https://doi.org/10.7326/M18-0850

[166] U. Trivedi, R. Alqasemi, and R. Dubey, “Wearable musical haptic sleeves for people with hearing impairment,” *ACM International Conference Proceeding Series*, pp. 146–151, 2019.

[167] P. D. United Nations, Department of Economic and Social Affairs, “World Population Prospects 2019: Highlights (ST/ESA/SER.A/423),” Tech. Rep. 141, 2019.

[168] R. T. Verrillo, “Vibration sensation in humans,” *Music Perception*, vol. 9, 1992.

[169] G. F. Welch, M. Biasutti, J. MacRitchie, G. E. McPherson, and E. Himonides, “Editorial: The impact of music on human development and well-being,” 2020.

[170] S. Willemsen, R. Paisa, and S. Serafin, “Resurrecting the tromba marina: A bowed virtual reality instrument using haptic feedback and accurate physical modelling,” in *17th Sound and Music Computing Conference*. Axea sas/SMC Network, 2020, pp. 300–307.

[171] B. S. Wilson and M. F. Dorman, “Cochlear implants: A remarkable past and a brilliant future,” *Hearing Research*, vol. 242, 2008.

[172] E. C. Wilson, C. M. Reed, and L. D. Braida, “Integration of auditory and vibrotactile stimuli: Effects of phase and stimulus-onset asynchrony,” *The Journal of the Acoustical Society of America*, vol. 126, no. 4, pp. 1960–1974, 2009.

[173] ——, “Integration of auditory and vibrotactile stimuli: Effects of frequency,” *The Journal of the Acoustical Society of America*, vol. 127, no. 5, pp. 3044–3059, 2010.

[174] A. Wright, A. Davis, G. Bredberg, L. Ulehlova, and H. Spencer, “Hair cell distributions in the normal human cochlea.” *Acta oto-laryngologica. Supplementum*, vol. 444, pp. 1–48, 1987.

[175] M. Yitao and X. Li, "Music and cochlear implants," *Journal of Otology*, vol. 8, no. 1, pp. 32–38, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1672293013500043

[176] G. Young, D. Murphy, and J. Weeter, "Haptics in music: The effects of vibrotactile stimulus in low frequency auditory difference detection tasks," *IEEE Transactions on Haptics*, vol. PP, pp. 1–1, 12 2016.

[177] G. W. Young, D. Murphy, and J. Weeter, "Haptics in music: The effects of vibrotactile stimulus in low frequency auditory difference detection tasks," *IEEE Transactions on Haptics*, vol. 10, pp. 135–139, 1 2017.

[178] M. Zampini, T. Brown, D. I. Shore, A. Maravita, B. Röder, and C. Spence, "Audiotactile temporal order judgments," *Acta psychologica*, vol. 118, no. 3, pp. 277–291, 2005.

[179] F.-G. Zeng, *Compression and Cochlear Implants*. New York, NY: Springer New York, 2004, pp. 184–220. [Online]. Available: https://doi.org/10.1007/0-387-21530-1_6

[180] ——, "Celebrating the one millionth cochlear implant," *JASA Express Letters*, vol. 2, 2022.

[181] F. G. Zeng and J. J. Galvin, "Amplitude mapping and phoneme recognition in cochlear implant listeners," *Ear and Hearing*, vol. 20, 1999.

[182] T. A. Zwolan and G. Basura, "Determining cochlear implant candidacy in adults: Limitations, expansions, and opportunities for improvement," *Seminars in Hearing*, vol. 42, 2021.

Part II

Papers

Paper A

Tactile displays for auditory augmentation - a scoping review and reflections on music applications for hearing impaired users

R. Paisa, N. C. Nilsson and S. Serafin

The paper has been published in
Frontiers in Computer Science, Sec. Human-Media Interaction Vol. 5, 2023

Abstract

The field of tactile augmentation has progressed greatly over the past 27 years and currently constitutes an emerging area of research, bridging topics ranging from neuroscience to robotics. One particular area of interest is studying the usage of tactile augmentation to provide inclusive musical experiences for deaf or hard-of-hearing individuals. This article details a scoping review that investigates and organizes tactile displays used for the augmentation of music from the field of hearing assistive devices, documented in 63 scientific publications. The focus is on the hardware, software, mapping, and evaluation of these displays, to identify established methods and techniques, as well as potential gaps in the literature. To achieve this purpose, a catalog of devices was created from the available literature indexed in the Scopus database. We set up a list of 12 descriptors belonging to physical, auditory, perceptual, purpose, and evaluation domains; and each tactile display identified was categorized based on those. The frequency of use among these descriptors was analyzed and as well as the eventual relationship between them. Results indicate that the field is relatively new, with 80% of the literature indexed being published after 2009. Moreover, most of the research is conducted in laboratories, with limited industry reach. Most of the studies have low reliability due to small sample sizes, and sometimes low validity due to limited access to the targeted population (e.g., evaluating systems designed for cochlear implant users, on normal hearing individuals). When it comes to the tactile displays, the results show that the hand area is targeted by the majority of the systems, probably due to the higher sensitivity afforded by it, and that there are only a couple of popular mapping systems used by the majority of researchers. Additional aspects of the displays were investigated, including the historical distribution of various characteristics (e.g., number of actuators, or actuators type) as well as the sonic material used as input. Finally, a discussion of the current state of the tactile augmentation of music is presented, as well as suggestions for potential future research.

A.1 Background

This section presents the rationale and objectives that form the basis for the scoping review detailed in the current paper. Subsequently, the field of vibrotactile augmentation is introduced along with a collection of relevant definitions.

A.1.1 Objectives

This article presents a scoping review of vibrotactile displays in the field of hearing assistive devices documented in 63 scientific publications. The main goal is to present a window into the relatively new research field of *Tactile Music Augmentation*—a research area that has particular applications for deaf or hard-of-hearing (D/HOH) individuals [75]. The articles included are described in terms of hardware, software, mapping, and evaluation of the displays, to provide pointers toward common methods and techniques, as well as potential shortcomings and opportunities. While the primary motivation for the current work is to aid hearing impaired individuals, the majority of the technology analyzed is designed for average users. However, there is no evidence of a difference between the haptic sense of populations with hearing loss and those with normal hearing, as the most significant differences (if any) are typically thought to be perceptual, as demonstrated by [67]. Thus, we believe that the devices intended for the general public have the potential to help D/HOH individuals just as well, and we have included them in our analysis. The *Scopus®*[1] database was queried and the eligible articles were dissected to create Table A.1.

To provide an overview of the research field of vibrotactile augmentation of sound, we sought to meet the following research objectives:

- Identify the state-of-the-art and understand how research efforts have changed over the years.
- Understand the most successful and promising strategies for augmenting music with tactile stimulation
- Identify gaps in current research
- Provide a starting point with a strong foundation for designers, researchers, and practitioners in the field of vibrotactile augmentation.

By reaching these goals, this review will address the following research questions: ”What are the most successful applications of vibrotactile augmentation?”, ”What is the historical distribution?” and ”What are the most popular actuators, processing techniques, and mappings?”

More specifically, this review will summarize research findings published over 27 years, in order to learn which type of actuators, body areas, mappings, processing techniques, and evaluation practices are most common, and how these factors have evolved over the years. In addition to this, such an examination could identify potential relations between different system components (e.g., type of actuators and type of signal processing used). This would not imply that such correlations are the most successful, but it will suggest starting points for new vibrotactile augmentation applications.

This article accompanies previous reviews in the field of vibrotactile augmentation, and should be seen as complementary. The review covering a similar sample of literature was documented by [65], who examined the methods and technologies used in the tactile rendering of music. Their work focuses on music through the touch modality in the general sense and encompasses literature covering use cases extending past vibrotactile augmentation of music, and thus the authors analyze works from a different perspective than this article. Other similar work includes a review of haptic wearables [74], a

---

[1] www.scopus.com

review of wearable haptic systems for the hands [57], as well as a review of clinically validated devices that use haptic stimulation to enhance auditory perception in listeners with hearing impairment [15]. In addition to these publications, in [58], the field of *Musical Haptics* is discussed extensively. Any researcher or designer interested in the field of vibrotactile augmentation is strongly encouraged to study these publications as well.

A.1.2 Vibrotactile Music Augmentation - a short overview

Research on cutaneous augmentation has been conducted since the beginning of the 20th century, focusing on thresholds of sensitivity by using tuning forks [69]. This primitive approach was abandoned in the follow-up work in favor of electronic transducers, and by 1935, it was known that the peak skin sensitivity is somewhere in the range of 200 - 250 Hz [71]. In 1954, A. Wilska mapped the threshold for 35 areas spread across the entire body [88].

Music is complex and has been written and performed with respect to the hearing capabilities of humans. While vibrotactile augmentation can manipulate percepts, it is important to outline the musical dimensions that can be perceived through tactile stimulation alone, to better understand and evaluate the effect of multisensory integration. Rhythm—the temporal relationship of events in a musical context—is arguably the first aspect to be discussed, as it is fairly well transmitted through tactile channels [33]. Substantial research has been dedicated to understanding the vibrotactile rhythm, investigating its impact on music aesthetics [28, 78], and in regards to D/HOH individuals, its enhancement properties [1, 21] as well as its interaction with the auditory counterpart [40]. When it comes to pitch—the perceived (vibrotactile) frequency—the bandwidth is very limited compared to the ear, but this does not mean that humans cannot perceive pitch differences, only that the *just noticeable difference* between intervals must be larger, and the range is bound within 20 Hz-1000 Hz [9]. However, this has not stopped researchers from exploring the potential of vibrotactile pitch with respect to music, either as in isolation [49], or more commonly in context through *Pitch Ranking* or *Melodic Contour Identification* tasks [27]. The last musical dimension that researchers focus on in terms of its vibrotactile properties is the timbre—the tempo-spectral characteristics of the stimulation that translates into the perceived quality of the sound, anecdotally refereed to as the *color of sound or tone color*. As with pitch discrimination, spectral content discrimination is inferior to its auditory counterpart, but there is evidence that humans can identify different spectral characteristics as discrete sensations [68].

Using tactile stimulation to augment auditory signals was first explored as hearing assistive devices focusing on improving speech perception for the D/HOH communities. The first commercial device that promised such results was called *Tickle Talker* and was developed in 1985 by *Cochlear Pty. Ltd* under the supervision of G. M. Clark—inventor of the cochlear implant (CI) [11]. The *Tickle Talker* was a *multichannel electrotactile speech processor* that presented speech as a pattern of electrical sensations on 4 fingers. The stimuli presented were processed similarly to the one for early-day CIs. Several other devices emerged in the mid-late 90s that explored the possibilities of using tactile stimulation to enhance speech for the hearing impaired; these will be discussed further in the current article.

What the early devices had in common with modern ones is the fundamental principle they rely on *multisensory integration*, pioneered by B. Stein & A. Meredith [77]. This mechanism links auditory and tactile sensations and describes how humans form a coherent, valid, and robust perception of reality by processing sensory stimuli from multiple modalities [77]. The classical rules for multisensory integration demand that enhancement occurs only for stimuli that are temporally coincident and propose that enhancement is strongest for those stimuli that individually are least effective [77]. This is especially useful for CI users that are shown to be better multisensory integrators [67].

For this integration to occur, the input from various sensors must eventually converge on the same neurons. In the specific case of auditory-somatosensory stimuli, recent studies demonstrate that multisensory integration can, in fact, occur at very early stages of cognition, resulting in supra-additive integration of touch and hearing [19, 38]. This translates to a lower level of robust synergy between the two sensory apparatuses that can be exploited to synthesize experiences impossible to achieve by unisensory means. Furthermore, research on auditory-tactile interactions has shown that tactile stimulus can influence the auditory stimulus and vice versa [1, 53, 54, 66]. It can therefore be observed that auditory and haptic stimuli are capable of modifying or altering the perception of each other when presented in unison, as described by [1, 91], and studied extensively with respect to music experiences [68].

Positive results from the speech experiments, as well as advancement in transducer technology, inspired researchers to explore the benefits of vibrotactile feedback in a musical context. Most of the works fall into two categories: *Musical Haptics*, which focuses mainly on the augmentation of musical instruments, as presented by [58], and *vibrotactile augmentation* of music listening, generally aimed at D/HOH. The focus of this article will be on the latter. A common system architecture can be seen in Figure A.1, with large variability for each step, depending on the goal. For example, in [12] a 4 actuator system is used to enhance music discrimination in a live concert scenario by creating a custom mapping scheme between the incoming signal and the frequency and amplitude of the transducers. A contrasting goal is presented by [64], where phoneme identification in speech is improved by using a total of 24 actuators. These two examples have different objectives, thus the systems have different requirements, but the overall architecture of both follows the one shown in Figure A.1. Throughout the article, each block will be analyzed in detail, and each system description can be seen in Table A.1.

Currently, research into tactile displays is expanding from speech to var-

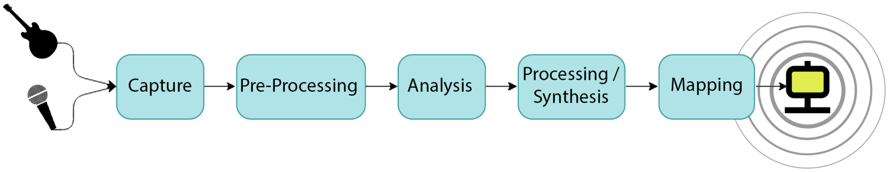


Fig. A.1: Archetype of a vibrotactile augmentation system

ious aspects of music enhancement, although the field of research is still relatively new; for example, over 80% of the articles included in this review are newer than 2009. Nevertheless, the technology is slowly coming out of research laboratories and into consumers' hands, with bands such as *Coldplay* offering *SubPacs*[2] for their D/HOH concert audience.

A.1.3 Definitions

Before presenting the objectives, it is worth introducing some general terminology and describing the interpretation used throughout this article. The first clarification is with respect to how words such as *tactile, vibrotactile and haptic* are used in this article:

- **Haptics** refer to *"the sensory inputs arising from receptors in skin, muscles, tendons, and joints that are used to derive information about the properties of objects as they are manipulated"*, explains [34]. It is worth highlighting that haptic sensations involve tactile systems and proprioceptive sensory mechanisms.

- **Tactile** refers to the ability of the skin to sense various stimulations, such as physical changes (mechanoreceptors), temperature (thermoreceptors), or pain (nociceptors), according to [44]. There are six different types of mechanoreceptors in the skin, each with an individual actuation range and frequency, and together they respond to physical changes, including touch, pressure, vibration, and stretch.

- **Vibrotactile** refers to the stimulation presented on the skin that is produced by oscillating devices. The authors of [44] present evidence that two types of mechanoreceptors respond to vibrotactile stimulation; namely, Pacinian receptors and, to a lesser extent, Meissner corpuscles. These receptors have been frequently analyzed and characterized in terms of their frequency and amplitude characteristics, but it is not excluded that there are other mechanoreceptors responsible for tactile percepts.

---

[2] https://subpac.com/

These definitions should provide a basic understanding of the taxonomy necessary to interpret this article; nevertheless, an individual study is recommended for a better understanding of the field of physiology and neuroscience. Furthermore, the authors recommend choosing the most descriptive terminology when discussing augmentation in order to avoid potential confusion (e.g., the use of *vibrotactile augmentation* instead of *haptic augmentation* is preferred when describing a system that involves vibrations). Finally, the term **augmentation** will be used as *the process of increasing the cognitive, perceptual or emotional, value, or quality* of the listing experience. Throughout this article, augmentation generally involves the usage of dedicated, specialized hardware (HW) and software (SW) systems.

Tactile augmentation of music is a fairly new multidisciplinary research field; therefore, it is paramount to achieve consensus on terminology and definitions.

A.2 Methodology

The methodology used to select and analyze the literature will be presented in the following section, starting with the system used to include articles and followed by an explanation of the process used to extract data from the article pool. No a priori review protocol was applied throughout the data collection method. The section ends by presenting known limitations in the methods described. The entire review process followed the *PRISMA-ScR* checklist and structure [81].

A.2.1 Identifying relevant studies

The *Scopus*® database was queried due to its high Scientific Journal Rankings required for inclusion, as well as significance for the topic. The inclusion selection was a 4 step process:

Step 1 included deciding upon the selection of keywords (below) that were used to browse the database.The keywords were designed to cover various aspects of tactile music augmentation.

1. Audio-haptic sensory substitution
2. Audio-tactile sensory substitution
3. Cochlear implant music
4. Cochlear implant (vibro)tactile
5. Cochlear implant haptic (display)
6. Electro-haptic stimulation
7. Hearing impaired music augmentation

8. Hearing impaired (vibro)tactile
9. (Vibro)tactile music
10. (Vibro)tactile display
11. (Vibro)tactile augmentation
12. (Vibro)tactile audio feedback

As of 23.02.2022 a total of 3555 articles were found that contain at least one of the items present in the list above in their title, abstract or keywords section. There was no discrimination between the types of documents that were included, but due to the database's inclusion criteria, PhD thesis and other potentially non-peer-reviewed works are not present.

Step 2 was a selection based on the title alone. Throughout steps 2-4, the eligibility criteria were as follows:

1. Written in English
2. Reporting primary research
3. Must describe devices that are, or could be used for tactile augmentation of music for D/HOH*
4. Must be designed for an audience (as opposite to a performer)

*A device that can be used for tactile augmentation of music for the hearing impaired (as item 3 describes) can be a system that was designed for laboratory studies that focused on the augmentation of speech and not music. Furthermore, a system for musical augmentation of normal hearing people could be used for D/HOH individuals since the tactile receptors do not differ depending on hearing capabilities; more-so there is evidence that the perception of congruent tactile stimulation is elevated in CI users [67]. Since the focus is on the technological aspects of these tactile displays and not on their efficiency, it was deemed relevant to include vibrotactile systems that are designed for music augmentation for normal hearing individuals as well.

The selection process was completed with the inclusion of 144 articles for further investigation.

Step 3 represented the reading of abstracts and evaluating their relevance; 102 articles were selected according to the eligibility criteria.

Multiple articles from the same authors describing the same setups were included only once, selecting the most recent publication and excluding the older ones. This would be the case where the authors evaluated multiple hypotheses using the same HW/SW setup.

Step 4 constituted of reading the articles selected in Step 3; the same inclusion criteria were used. This step coincided with data extraction, but articles that had been wrongfully included based on the abstract alone were discarded. The entire inclusion process resulted in 63 articles that were used to construct this review.

A.2.2 Data items

The relevant articles were studied and a selection of descriptors was noted for each system presented; as mentioned in Section A.1.1, the focus is on hardware, software and evaluation practices, making the analysis an agnostic process with respect to their applications, target groups, or success. If the article discussed more than one system, all relevant ones were included and analyzed. The features used to analyze and compare the systems were:

1. **Purpose of the display** represents the end goal of the device, irrespective of the eventual evaluation conducted. All documented purposes for each display were included. A common purpose would be ″Speech Enhancement″.

2. **Listening situation** refers to the context where the display has been used. A frequent situation is the ″Laboratory Study″.

3. **Number of actuators** indicates the total number of actuators in the display, regardless of their type.

4. **Actuators type** enumerates the type of actuators used, from a hardware perspective. For example, eccentric rotating mass (ERM) vibration motors are one of the most widely used haptic technologies.

5. **Signals used to feed the actuators** represents the type of audio material used to excite the transducers. In the case of ERMs, most displays will indicate a ″sine wave″, unless otherwise noted. This is due to the hardware nature of the actuator, capable of producing only sinusoidal oscillations, while the actual signal used is a DC voltage. Some ERM systems can reproduce harmonically complex signals as well, but the cases are few and far apart.

6. **Type of signal processing** (generally called DSP) enumerates the processes applied to an audio input signal to extract relevant information or prepare it for the tactile display. Fundamental frequency (F0) extraction is a common signal processing technique used for tactile displays.

7. **Mapping scheme** describes the features from the auditory input that are mapped to the tactile output.

8. **Area of the body actuated** presents where the actuators are placed on the human body (e.g., hands and chest). Throughout this article *″hand″* is used a combination of 2 or more sub-regions (e.g.,palm and fingers).

9. **Whether it is a wearable device or not** (binary Yes/No)

10. **Evaluation measurement** presents the measurement criteria assessed in the evaluation (if applicable). Most items in this descriptor column

have been grouped into a meta-category; for example ″vocal pitch accuracy″ seen in [73] and ″pitch estimation accuracy″ from [15] have been grouped into ″Music Listening Performance″.

11. **Evaluation population** describes the hearing characteristics of the individuals participating in the evaluation (if applicable).
12. **Number of participants** presents the total number of participants in the evaluation described by the previous two items.

These features were chosen in order to create an objective and complete characterization of each system while allowing a high degree of comparability. Furthermore, features regarding evaluation were included to better understand the research field as a whole. A detailed explanation for some relevant categories can be found in Section A.3.1. Based on these features, Table A.1 was constructed that contains all the systems analyzed.

A.2.3 Delimitation

To ensure that the inclusion process was feasible, we imposed several constraints on the process, which may have influenced some aspects of the review. First, only one database was used to browse for articles, by only one author. Although inquiring several databases is recommended, the Scopus® database already provided a large sample of publications, and includes many if not all the relevant journal publishers (e.g., IEEE). Furthermore, using the articles indexed in an academic database results in the inevitable exclusion of artistic work that might have limited exposure but with a potentially valuable contribution. Since music is inherently an artistic expression and not an academic work, this limitation could greatly impact the results of the analysis presented in this article. Similarly, articles written in other languages than English were excluded, resulting in many works not having a chance of inclusion.

A.3 Results

This section provides a thematic analysis of the systems included in the reviewing process. The data is first presented as a table of characteristics that describes each system and its usage, and it will be succeeded by a graphical representation of the most important findings.

A.3.1 Table explanation

Due to the large variation in the hardware design and evaluation methods used in the literature presented in Table A.1, some categories contain meta-descriptors that encapsulate similar features. This section will identify and clarify these situations, as well as provide an additional explanation that would allow readers an easier interpretation of the Table A.1.

The first notion worth explaining is the usage of the *"N/A"* acronym. Although most of the time it should be understood as *"Not Available"*—information that is not present in the cited literature—some situations fit the interpretation more to *"not applicable"*. One example can be encountered in the *"Nr. part."* category, where *"N/A"* generally means "not applicable" if the tactile display has not been evaluated at all. A more detailed version of the table, as well as the analysis software, can be found at Tactile Displays Review Repository[3].

Second, the columns *"Evaluation Population"* and *"Nr. Part."* are linked, and the number of participants separated by commas represents each population described in the previous category. For example, in [72], *Evaluation Population* contains: *Normal Hearing, Hearing aid(HA) Users, CI Users, D/HOH* while *"Nr. Part."* contains *10, 2, 6, 2*; this should be red as 10 normal hearing participants, 2 hearing aid users, 6 cochlear implant users, and 2 deaf or hard-of-hearing individuals.

Third, in the *"Mapping"* column several items called *"Complex Mapping"* that encapsulate all the more elaborate mappings, as well as mappings that have been specially designed for the tactile display in question.

Fourth, the tactile displays analyzed usually fall under more than one category, with respect to those ones mentioned in Section A.2.2. For example, Figure 10 shows *purpose of devices*, where several tactile displays fall under HAD as a secondary purpose, on top of their main goal (e.g., Music Enhancement). This means that analyzed items in most of the figures presented in Section A.3.2 are not exclusive, resulting in a greater total amount of data points than number of displays included.

Next, a brief description of the actuator types encountered in the review and shown in Figure A.4 is presented below. It is worth noting that the list ignores auxiliary systems necessary to operate these actuators (converters, amplifiers, etc.), and describes them based on their practical tactile applications:

- **Voice Coil** actuators get their name from the most common application: moving the paper cone in a speaker and are also known as non-commutated DC linear actuators. They consist of a permanent magnetic element (sometimes replaced by an electromagnet with the same role) and a suspended coil attached to a mobile mass. A variation in this architecture exists, where the permanent magnet is the moving mass, and the coil is static. The current from an amplifier that flows through the coil, creates an electromagnetic field that interacts with the permanent magnet, moving the mass (or the paper cone) accordingly. Voice coil actuators come in various sizes and forces available, have a rela-

---

[3] https://github.com/razvysme/TactileDisplaysReview

tively wide frequency response (in the KHz range), and provide high acceleration.

- **Subwoofers** are a type of voice coil actuators that are optimized to reproduce sound at low frequencies, commonly below 80Hz. Their construction and size vary radically depending on the application, but they do imply the properties described above. One drawback of subwoofers is that they usually require generous amplification to operate. Furthermore, they are optimized for sound reproduction, therefore their tactile characteristics are usually a byproduct of high amplitude playback.

- **Solenoids** are somewhat similar to voice coils, but instead of providing a permanent magnet interacting with an electromagnet, they have a coil creating a magnetic field in order to move a ferrous shaft. Solenoids are generally used to open or close locks, valves, or to apply a constant force on a surface and are not necessarily suitable for oscillating behavior, resulting in a limited frequency response.

- **Piezo-actuators** are mechanisms that vibrate based on the change in the shape of a piezoelectric material and belong to the category of *"resonating actuators"*, which have an efficient operating frequency embedded in their mechanical design. Because of that, piezo-actuators have limited frequency response (generally within 80% of the resonant frequency), but they can be designed to be tiny or in complex shapes, as opposed to the ones presented above. The tactile feedback produced by piezo-actuators is relatively modest, for a given current.

- **Linear Resonant Actuators (LRA)** are mass-spring systems that employ a suspended mass attached to an electromagnetic coil that vibrates in a linear fashion due to the interaction with the permanently magnetized enclosure. Being a resonant system, they need to be driven with signals close to the peak frequency response, similar to piezo-actuators and ERMs described below. Some advanced LRAs have auto-resonance systems that detect the optimal frequency for producing highest amplitude tactile feedback, trading tactile frequency accuracy for perceived intensity.

- **Electro-dynamic shakers (EDS)** is an industrial name given to vibration systems with excellent frequency response characteristics, generally necessary in vibration analysis and acoustics industries. They come in two categories, voice coils and electro-hydraulic shakers, but the systems described in Table A.1 only use voice the former type. These can be seen more like bass shakers described below, but with much better frequency response as well as a much higher cost.

- **Eccentric Rotating Mass (ERMs)** are another type of resonating actuators that operate by attaching an unbalanced mass to the shaft of a DC motor. Rotating the mass produces vibrations of different frequencies and amplitudes, typically linked to the amount of current fed to the motor. These types of actuators are very popular due to their low cost, and relatively strong vibration force, but they respond slower than other resonators to a chance in the current (a lag of 40-80ms is commonly expected). Another limitation of ERMs, as well as LRAs and piezo-actuators is that the frequency and amplitude reproduced are correlated due to their resonating design, and thus are generally suggested when limited tactile frequency information is necessary.

- **Dual-Mode Actuators (DMAs)** are relatively new types of actuator that are similar to LRAs, but are designed to operate at two different frequencies simultaneously, usually out of phase with each other. Due to their novelty, the amount of variation and experimentation with them is limited, but citepHwang2013 provides evidence that DMAs outperform LRAs in tactile displays in music as well as HCI applications.

- **Contact speakers** are a sub-category of voice coils that are primarily designed to excite hard surfaces in order to produce sound. They work by moving a suspended coil that has a shaft with a contact surface at the end. The contact surface is usually glued or screw on the desired surface, thus vibrating as the coil oscillates. Contact speakers vary largely in size and power requirements, but generally provide a wide frequency response. In the context of tactile stimulation, they usually have a poor low-frequency representation (below 100Hz) and are always producing sound (sometimes very loud), due to their focus on the auditory reproduction.

- **Bass shakers** are another sub-category of voice coils that are grouped by marketed applications rather than physical properties. They work by having a large mass attached to the moving coil, usually in a protective enclosure. These devices are suggested to be used in an audio listening scenario (be it films or games) and should be attached to seating furniture (sofa, chairs, etc). Some vendors provide mounting hardware for drum stools or vibrating platforms designed for stage musicians that complement in-ear headphones to monitor band activity. Bass shakers are generally large, heavy, and require abundant power to operate, but provide a relatively good frequency response up to approx. 350Hz.

Device Name	*Purpose*	*Listening Situation*	*Nr. Act.*	*Actuators Type*	*Signals Used*	*DSP*	*Mapping Scheme*	*Body Area Actuated*	*Wearable*	*Evaluation Measure*	*Evaluation Population*	*Nr. Part.*
Pump-and-Vibe [25]	Music Enhancement	Lab Study	8	Mixed	Sawtooth wave	F0 extraction, Beat extraction	Complex Mapping, Pitch to Position	Arm	Yes	Experience	Normal Hearing	20
Tactile Phonemic Sleeve (TAPS) [64]	Speech Enhancement	Lab Study	24	Contact speakers	Sine wave	N/A	Complex Mapping	Arm	Yes	Discrimination	Normal Hearing	7
N/A [59]	HAD, Speech Enhancement	Lab Study	3	ERM	Sine wave	Phoneme extraction	Complex Mapping	Hands	Yes	N/A	N/A	N/A
mosaicOne_C [15]	HAD, Music Enhancement	General	4	ERM	Sine wave	N/A	N/A	Wrist	Yes	N/A	N/A	N/A
Syntacts [62]	SW/HW platform	General	N/A	Unspecified	Various Waveforms	N/A	N/A	N/A	N/A	N/A	N/A	N/A
N/A [12]	Music Enhancement	Lab Study	4	LRA	Sine wave	Spectrum isolation, Envelope extraction	Pitch to Position	Wrist	Yes	Discrimination	Normal Hearing	3
N/A [83]	Music Enhancement	Artistic	4	ERM	Sine wave	N/A	Complex Mapping	Arm	Yes	Experience	Normal Hearing	20
N/A [17]	HAD, Sound Localization	Lab Study	2	EVT	Sine wave	Multiband compression, ILD Enhancement, Envelope extraction	Auditory Frequency to Tactile Frequency	Wrist	No	Discrimination	Normal Hearing	32
mosaicOne_B [16]	HAD, Pitch Discrimination	Lab Study	12	ERM	Various Waveforms	F0 extraction	Pitch to Position	Arm	Yes	Music Listening Performance	Normal Hearing	12
Tactile Tone [73]	HAD, Music Training	Singing along music	9	ERM	Sine wave	F0 extraction	Complex Mapping	Hands	Yes	Music Listening Performance	CI Users	2
N/A [72]	HAD, Music Enhancement	Lab Study	6	Voice coil	Music	N/A	N/A	Hands	Yes	Discrimination	Normal Hearing, HA Users, CI Users, D/HOH	10, 2, 6, 2
EarVR [47]	HAD, Sound Localization	VR, Lab Study	2	ERM	Square wave	Amplitude Thresholding	Auditory Amplitude to Tactile Amplitude	Ears	Yes	Discrimination	Normal Hearing, D/HOH	20, 20
Tasbi [63]	AR/VR Human-Computer Interaction	N/A	6	LRA	Sine wave	N/A	N/A	Wrist	Yes	N/A	N/A	N/A
N/A [42]	HAD, AV Enhancement	Lab Study, Film	3	ERM	Square wave	Note onset detection	N/A	Hands	Yes	Experience	Normal Hearing, D/HOH, HA Users, CI Users	9, 4, 1, 2
Sound Forest [20]	Digital Musical Instrument (DMI)	Public Exhibition	1	Voice coil, Bass shaker	Sine wave, Noise	N/A, LPF	N/A	Whole Body, Feet	No	Experience	Normal Hearing	4

Continues on the next page...

Device Name	*Purpose*	*Listening Situation*	*Nr. Act.*	*Actuators Type*	*Signals Used*	*DSP*	*Mapping Scheme*	*Body Area Actuated*	*Wearable*	*Evaluation Measure*	*Evaluation Population*	*Nr. Part.*
N/A [43]	HAD, Music Enhancement	Lab Study	1	ERM	Sine wave	F0 extraction	Auditory Frequency to Tactile Frequency	Wrist	Yes	Music Listening Performance	Normal Hearing	8
GLOS [22]	HAD, Speech Enhancement	Lab Study	5	ERM	Sine wave	Speech-to-Text	Complex Mapping	Fingers	Yes	Speech Performance	Normal Hearing	3
N/A [82]	HAD, Music Enhancement	Lab Study	4	Voice coil	Piano	Frequency Thresholding	Pitch to Position	Leg	Yes	Music Listening Performance	Normal Hearing	5
N/A [28]	Music Enhancement	General	1	Subwoofer, Bass shaker	Music	LPF	N/A	Back	No	Experience	Normal Hearing	40
body:suit:score [87]	Music Composition	N/A	60	ERM	Sine wave	Envelope extraction, F0 extraction	Amplitude to Position, Pitch to Position	Whole Body	Yes	N/A	N/A	N/A
N/A [10]	HAD, Speech Enhancement	Lab Study, Concert	2	Voice coil	Speech	F0 extraction, Envelope extraction	Auditory Frequency to Tactile Frequency	Fingers	No	Discrimination	Normal Hearing	12
N/A [50]	Music Training	Lab Study	1	Voice coil	Sine wave	N/A	Auditory Frequency to Tactile Frequency	Hands	Yes	N/A	N/A	N/A
N/A [13]	Music Enhancement	Lab Study	3	Voice coil	Sine wave	N/A	N/A	Flank, Fingers	No	Discrimination	Normal Hearing	18
MuSS-Bits++ [60]	Music Training	Lab Study, Music teaching for DHOH	1	Voice coil	Music	Frequency extraction, Envelope extraction	Auditory Frequency to Tactile Frequency	Hands	Yes	Music Listening Performance	D/HOH	11
N/A [29]	HAD, Music Enhancement	Lab Study	1	Voice coil	Music	LPF	N/A	Fingers	Yes	Discrimination	CI Users	17
LIVEJACKET [23]	Music Enhancement	Lab Study	22	Mixed	Music	N/A	Instrument to Position	Whole Body	Yes	Experience	Normal Hearing	12
VibGrip++ [35]	Music Enhancement	Lab Study	5	Piezo	Music	N/A	N/A	Hands	Yes	N/A	N/A	N/A
Hedonic Tactile Player [84]	Exploratory	General	3	ERM	Square wave, Sine wave, Triangle wave	N/A	N/A	Whole Body	Yes	Experience	Normal Hearing	3
Basslet [41]	Music Enhancement, Gaming	General	1	Voice coil	Various Waveforms	LPF	N/A	Wrist	Yes	N/A	N/A	N/A
N/A [30]	HAD, Speech Enhancement	Lab Study	1	Voice coil	Speech	F0 extraction, Envelope extraction	N/A	Fingers	Yes	Discrimination	CI Users	10
N/A [18]	HAD, Music Enhancement, Dance	Lab Study	5	ERM	Sine wave	Envelope extraction	N/A	Arm	Yes	Experience	D/HOH	45

Continues on the next page...

Device Name	*Purpose*	*Listening Situation*	*Nr. Act.*	*Actuators Type*	*Signals Used*	*DSP*	*Mapping Scheme*	*Body Area Actuated*	*Wearable*	*Evaluation Measure*	*Evaluation Population*	*Nr. Part.*
N/A [80]	HAD, Music Enhancement, Dance	Lab Study	1	Bass shaker	Music	LPF	N/A	Whole Body, Feet	No	Music Listening Performance	Normal Hearing, D/HOH	14, 7
SmartFingerBraille [56]	HAD, Speech Enhancement	Lab Study	6	ERM	Sine wave	N/A	N/A	Fingers	Yes	Experience	Normal Hearing	3
Auris System [2]	HAD, Music Enhancement	Lab Study	11	Mixed	Music	F0 extraction, Envelope extraction	Auditory Frequency to Tactile Frequency	Whole Body	No	Experience	Normal Hearing, D/HOH	13
MuSS-Bits [61]	HAD, Music Training, Music Playing	Exploratory	1	ERM	Sine wave	Envelope extraction	Auditory Amplitude to Tactile Amplitude	Wrist	Yes	Experience	D/HOH	11
Mood Glove [45]	Film Enhancement	Film	8	ERM	Sine wave	N/A	Complex Mapping	Hands	Yes	Experience	Normal Hearing	10
N/A [26]	Exploratory	Lab Study	1	EVT	Sine wave	N/A	N/A	Fingers, Feet	No	Discrimination	Normal Hearing, D/HOH	70, 14
N/A [3]	Music Enhancement, Music Composition	Artistic	16	ERM	Square wave	Binaural auditory rendering	Auditory Spatialization to Tactile Spatilization	N/A	Yes	N/A	N/A	N/A
Silent Rave Furniture [32]	Music Enhancement	Concert	9	Voice coil	Sine wave, Noise	Vocoding	Pitch to Position	Back, Hands, Feet	No	Experience	D/HOH	1
N/A [24]	Music Enhancement	Artistic	6	Mixed	Square wave, Music	N/A	Pitch to Position	Whole Body	No	N/A	N/A	N/A
CollarBeat [70]	Music Enhancement	Lab Study	2	Voice coil	Sine wave, Music	N/A	N/A	Whole Body	No	Experience	Normal Hearing	6
N/A [92]	Exploratory	Lab Study	6	Voice coil	Various Waveforms	N/A	N/A	Hands	Yes	Discrimination	Normal Hearing	30
N/A [39]	Music Playing	Lab Study	4	Voice coil	Various Waveforms	F0 extraction	Gesture to Tactile Parameters	Fingers	Yes	Discrimination, Experience	Normal Hearing	5
N/A [7]	Film Enhancement	Lab Study	16	Voice coil	Music, Noise	Envelope extraction, Spectrum isolation	Pitch to Position	Back	No	Experience	Normal Hearing	59
[31]	Music Enhancement	Lab Study	1	DMA	Sine wave	Spectrum isolation	Auditory Amplitude to Tactile Amplitude	Fingers	No	Experience	Normal Hearing	24
Haptic chair [52]	HAD, Music Enhancement	Lab Study	4, 8	Voice coil	Music	N/A	N/A	Whole Body	No	Experience	D/HOH	43, 12
N/A [90]	DMI	Lab Study	6	Voice coil	Various Waveforms	N/A	N/A	Fingers	No	Discrimination	Normal Hearing	10
Haptic chair [51]	Speech Enhancement	Lab Study	4	Contact speakers	Speech	N/A	N/A	Whole Body	No	Speech Performance	D/HOH	20
Emoti-chair [4]	Music Enhancement	Lab Study, Workshop	16	Voice coil	Music	Spectrum isolation	Pitch to Position	Back	No	Experience	Normal Hearing, D/HOH	12, 6
N/A [79]	Sound Localization	Lab Study	4	ERM	Sine wave	N/A	N/A	Head	No	N/A	N/A	N/A

Continues on the next page...

Device Name	*Purpose*	*Listening Situation*	*Nr. Act.*	*Actuators Type*	*Signals Used*	*DSP*	*Mapping Scheme*	*Body Area Actuated*	*Wearable*	*Evaluation Measure*	*Evaluation Population*	*Nr. Part.*
Emoti-chair [36]	Music Enhancement	Lab Study	16	Voice coil	Music	Spectrum isolation, Instrument isolation, Pitch shift	Pitch to Position, Instrument to Position	Back	No	N/A	N/A	N/A
N/A [6]	Music Enhancement	Lab Study	1	Voice coil	Square wave	N/A	N/A	Waist	No	Discrimination	Normal Hearing	4
N/A [5]	HAD, Speech Enhancement	Lab Study	5	ERM	Sine wave	Spectrum isolation	Pitch to Position	Fingers	Yes	Speech Performance	Normal Hearing	N/A
N/A [85]	HAD, Speech Enhancement	Lab Study	64	Piezo	Sine wave	Spectrum isolation, Envelope extraction, Amplitude modulation	Complex Mapping	Fingers	Yes	Discrimination, Speech Performance	Unspecified	Unspecified
N/A [76]	HAD, Speech Enhancement	Lab Study	2	Unspecified	Unspecified	Envelope extraction	N/A	Hands	Yes	Speech Performance	Normal Hearing	Unspecified
N/A [46]	HAD, Speech Enhancement	Lab Study	1	Piezo	Custom AM, FM coding	Amplitude Thresholding, Spectrum isolation	Complex Mapping	Wrist	Yes	Speech Performance	D/HOH	5
N/A [48]	HAD, Speech Enhancement	Lab Study	2	Piezo	Unspecified waveform	Spectrum isolation	Pitch to Position	Sternum	Yes	Speech Performance	CI Users	10
N/A [55]	HAD, Speech Enhancement	Lab Study	7	Piezo	Unspecified waveform	Formant tracking	Complex Mapping	Sternum	Yes	Speech Performance	D/HOH	2
N/A [86]	HAD, Speech Enhancement	Lab Study	1	Unspecified	Unspecified	N/A	N/A	Wrist	Yes	Discrimination, Speech Performance	D/HOH	8
N/A [86]	HAD, Speech Enhancement	Lab Study	16	LRA	Sine wave	Spectrum isolation	Pitch to Position	Arm, Abdomen	Yes	Discrimination	Normal Hearing	3
N/A [89]	HAD, Speech Enhancement	Lab Study	16	Solenoids	Square wave	F0 extraction	Pitch to Position, Pitch to Amplitude	Arm	Yes	N/A	N/A	N/A
N/A [14]	HAD, Speech Enhancement	Lab Study	16	Solenoids	Square wave	Spectrum isolation	Pitch to Position	N/A	N/A	N/A	N/A	N/A

Table A.1: Description of the tactile displays analyzed in this study. A complete explanation of the labels can be found in section A.3.1

A.3.2 Synthesis of results

Besides Table A.1, plots and histograms highlighting the most interesting relationships between characteristics are discussed.

When doing the literature search for scientific publications that included the phrases described in section A.2.1, we found that 84% of tactile displays have been introduced after 2010, as shown in Figure A.2. This is strong evidence that interest in integrating tactile stimulation into the music-listening activities is blooming.

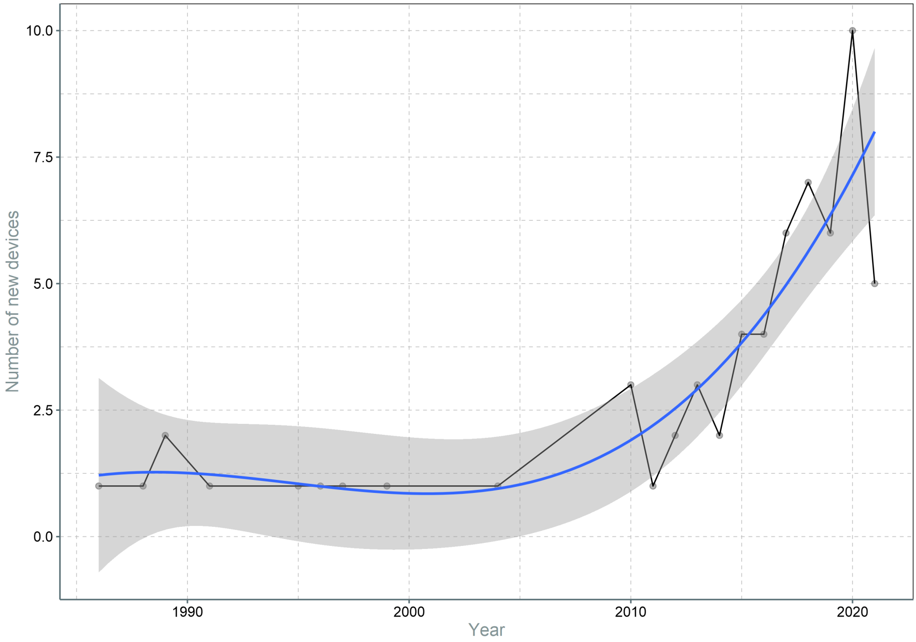


Fig. A.2: Distribution of the number of new systems described in publications every year; black line is the empirical value, the blue line is the predicted value and the grey area represents the 95% confidence interval for the estimated number of studies. The prediction model uses a cubic regression that was chosen due to the lowest AIC and BIC scores (as described by [8]) when compared to other models of orders 1 - 5.

Distribution of types of actuators used over time

One of the main descriptors of a system is the number of actuators it uses, and it could easily be (wrongfully) assumed that the advancement in transducer technology would encourage researchers to use more actuators in their studies. As shown in Figure A.3, the average number of actuators decreases slightly over the years, and the predictor model (blue line) suggests that this number will not increase in the near future. In addition to the occasional outliers with more than 60 actuators, most systems use less than 20 transducers and more than half use less than the average of 8.

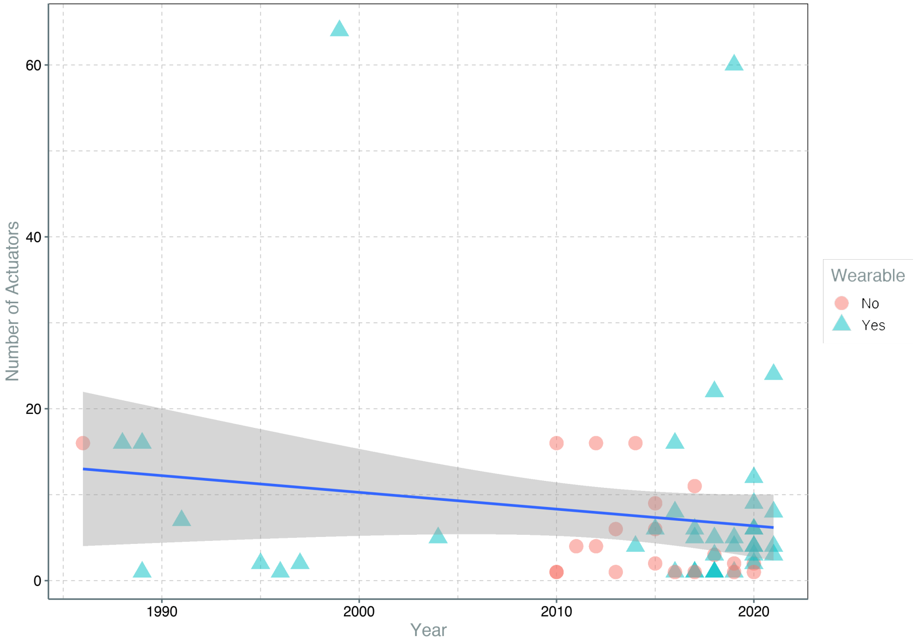


Fig. A.3: Distribution of the number of actuators in a device over years; dashed orange line is the mean over years, the blue line is predicted value, and the grey area represents the 95% confidence interval for the estimated number of actuators; a blue triangle indicates a wearable device, and a red circle indicates a fixed device. The prediction model uses a linear regression that was chosen due to the lowest AIC and BIC scores (as described by [8]) when compared to other models of orders 1 - 5.

Plotting the distribution of actuators over time in Figure A.4, shows that voice coils have been used since the 2010s, and are generally preferred for applications that require higher frequency and amplitude accuracy. The drawbacks of this type of actuators are that they are generally larger, more expensive, and would require a more complex Digital-to-Analog (DAC) to operate since they use bipolar signals. Including subwoofers, contact speakers and bass shakers in the ”voice coil” category would create a cluster representing 34%, signifying their importance. Another popular choice is eccentric rotating mass (ERM) actuators, which are smaller, cheaper, and simpler to operate than voice coils, but provide limited frequency response, and the amplitude is coupled to the frequency. It is also interesting to observe that older systems used piezo and solenoids–technologies that have a very limited application range with small amplitude, or small frequency response. Lastly, it’s important to highlight that one category of electro-dynamic systems is never used for tactile augmentation—electro-hydraulic shakers. These devices can provide large displacement and could be deployed to actuate very large surfaces, but have a limited frequency response, generally below 200Hz.

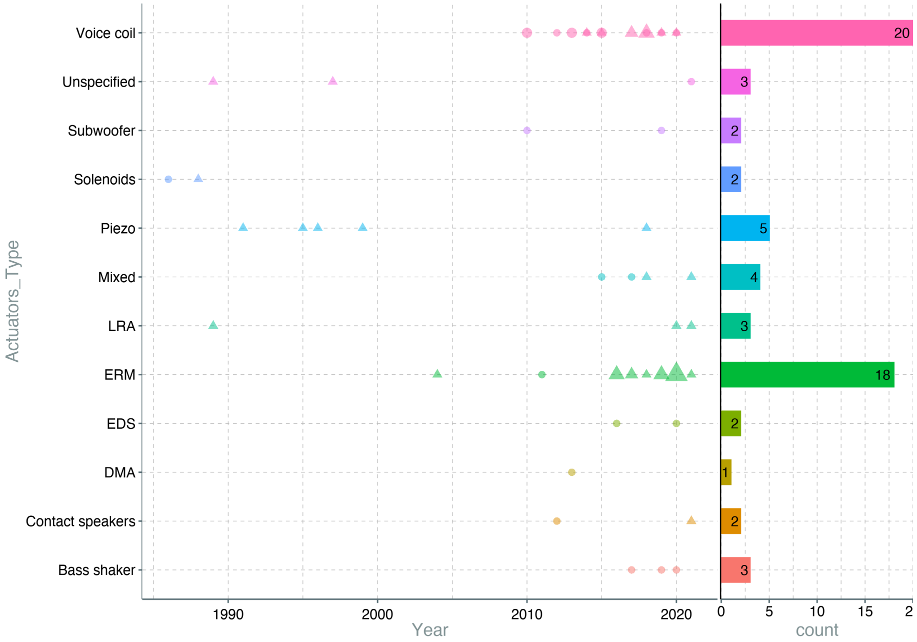


Fig. A.4: Distribution of types of actuators used over time; a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices

Mappings

Figure A.5 shows that before 1990 mapping schemes were rather simple (e.g., mapping pitch-to-position or amplitude). Tactile frequency has only recently been brought into discussion probably because of the high computational power required for the analysis stage, but also because of recent advancements in voice coil actuators. This technical progress allowed researchers to explore the tactile frequency as a method of encoding the auditory frequency. The pitch-to-position mapping (the idea of cochlear implants) is the one that is used the longest, probably because the creator of the CI and *Tickle Talker* used the same mapping for both. This could have been an inspiration for further research to produce incremental improvements in these systems, rather than revolutionary approaches, as can be seen in the work of [51] and [37]. Around 2015, new mappings start to be explored, and the popularity of ”Pitch-to-position” starts decreasing.

Looking at the relationship between the mapping schemes and the number of actuators in Figure A.6 it is interesting to highlight the fact that only

Fig. A.5: Distribution of mapping schemes used over time; a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices

one mapping scheme utilises tactile frequency, in combination with voice coil actuators. Furthermore, *something*-to-position mapping is popular, taking advantage of the larger surface area the body can afford. This hypothesis is reinforced by Figure A.11 showing that areas of high sensitivity are excited with a small number of actuators, probably constrained by the actuator size.

Evaluation practices

Observing Figure A.7 it can be seen that speech research was the main focus before 2013, while music has been the topic of more investigation since. This can be explained by the advancement in hearing aids and CI technology that solved the speech intelligibility problem to a satisfactory degree. Nevertheless, the need to use tactile augmentation remains present when music is played for D/HOH or CI users due to the problems shown by all hearing assistive devices have with multi-stream, complex signals, as well as with timbre and melody recognition and sound localization in the case of CIs. Simultaneously, subjective experiences of users have become a topic of interest in the second decade of the century, as seen in the work of [12] or [47], even though tactile augmentation research has been exploring that topic for more than 27 years. A further look at the plot shows a large amount of research

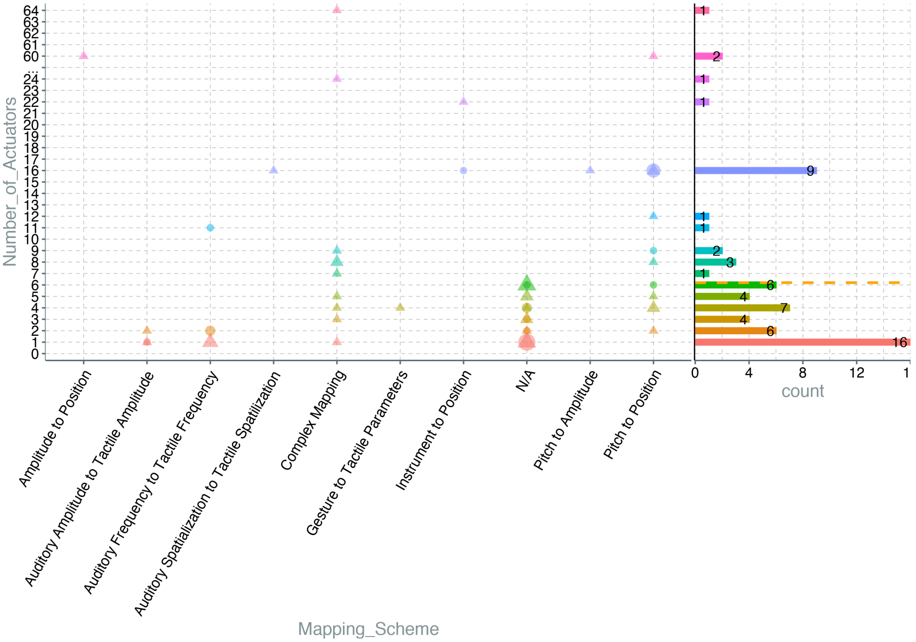


Fig. A.6: Distribution of the number of actuators for each mapping scheme; a triangle indicates a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices. The dashed orange line represents the mean.

on users' experiences between 2000 and 2010. At the same time, there is still interest in finding the limits of the physiological and cognitive systems involved in tactile perception and integration, as evidenced by the large number of studies involving discrimination tasks conducted in the past decade. An extended version of Figure A.7, reinforces the fact that speech enhancement has been a focus since the beginning of tactile augmentation. Although interest in it has decreased in recent years, there are still researchers working on it. The second wave of tactile augmentation research started around 2010 and it focused mostly on music, but new technologies and media consumption methods show interest into tactile stimulation as well (e.g., gaming and enhancement of film and AV).

The overall number of participants is very small, with a mean of 11, resulting in generally low-reliability studies, as displayed in Figure A.8. Furthermore, we can see that the hearing aid and CI users studies are mostly below this average. This is a strong indicator that most of the research is preliminary and underlines the necessity for studies to adhere to clinical guidelines (longitudinal, more participants, better control), as well as the need for replication of prior studies. Nevertheless, CI users are mostly evaluating systems focusing on music and speech, while HA users only focus on music as seen

Fig. A.7: Distribution of evaluation measurements over years, a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices

in Figure A.9. This could be explained by the fact that the speech needs for HA users are fulfilled by the current state-of-the-art in HA tech. However, it could be argued that tactile augmentation could also be interesting for normal hearing and hearing aid users as well; this can be seen in the push for multimodal cinema experiences. Nevertheless, it cannot be overlooked that devices aimed at groups with a particular set of requirements (CI and HA users) are evaluated using a normal hearing population, sometimes exclusively. This practice, while common, might result in studies with lower validity, because the requirements for the target population are not met.

In Figure A.10 the relationship between listening situations and evaluation measures is presented. On one hand, the large number of discrimination studies further highlight the incipient state of the research field, which looks to outline the "playing field" by testing the threshold and *just-noticeable differences*. On the other hand, studies focusing on music are increasingly more frequent, indicating that researchers bring forward new systems and ideas that have a more applied research angle to them. This aspect is supported as well by the large emphasis on the users' experience, which is evaluated in various listening situations, not only in research laboratories.

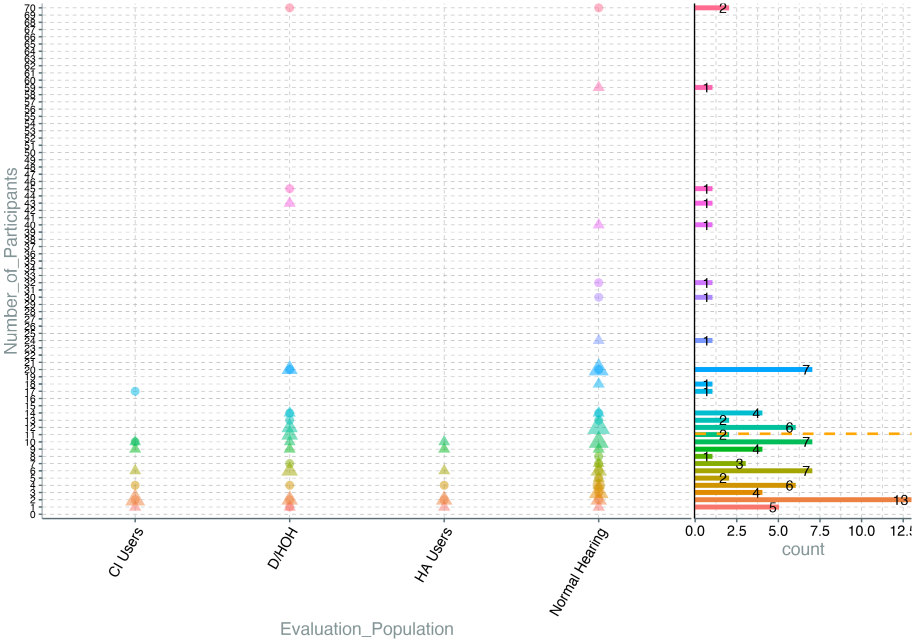


Fig. A.8: Distribution of the number of participants for each of the populations evaluated; a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices.*D/HOH* represents persons suffering from a hearing disability, but without any hearing assistive device. The dashed orange line represents the mean.

Body regions used for stimulation

Figure A.11 shows that systems designed for fingers use a few actuators, usually one actuator per finger. Similarly, for the wrists, it seems that a small number of actuators is preferred, probably because of physical limitations. On the other hand, when the whole body is used, the number of actuators increases; the same is true for arms. This clearly indicates that the size of the actuator, as well as the spatial resolution of the skin, is a constraining factor in designing complex systems, and advancements in actuator technologies will allow designers to insert more actuators aiming at the high sensitivity areas (hands, wrist, etc). It is known that there is a positive correlation between 2-point discrimination and the contact size of the actuator, so small actuators would require a smaller space between them.

The wrist and fingers are used with mappings that require greater accuracy in terms of frequency, while mappings that rely on simpler encoding (such as amplitude and custom encoding) are generally used with a higher number of actuators, as shown in Figure A.11; a potential explanation for this is that the lower physiological accuracy in certain areas is compensated for by a higher number of actuators. Going back to the number of actuators in

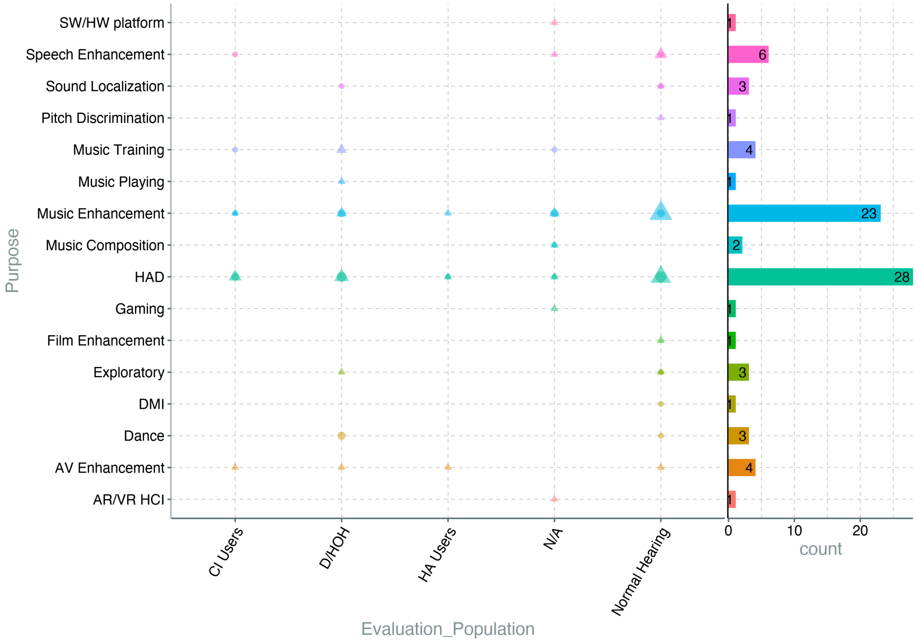


Fig. A.9: Distribution of the purpose for each evaluation population; a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices. textitD/HOH represents persons suffering from a hearing disability, but without any hearing assistive device; items on Y axis are not exclusive.

areas that are using *"auditory frequency to tactile frequency"* mapping, it seems that one or two actuators are sufficient for this type of mapping.

A.4 Discussion

A.4.1 Summary of evidence

In this scoping review, we identified 63 primary articles that describe unique vibrotactile displays used for audio augmentation, published from 1986 to 2021. Within this specific research pool, our findings highlight that most of the work in the field of vibrotactile augmentation of sound can be categorized as preliminary, missing the large-scale studies usually associated with clinical research. This conclusion is supported by the low reliability of evaluations presented derived from a low number of participants, as well as the occasional low validity of the said evaluations. The latter is evidenced by experiments conducted with poorly sampled individuals; for example, tactile displays designed for D/HOH are evaluated on normal hearing users. Finally, it should be underlined that much of the available literature covers

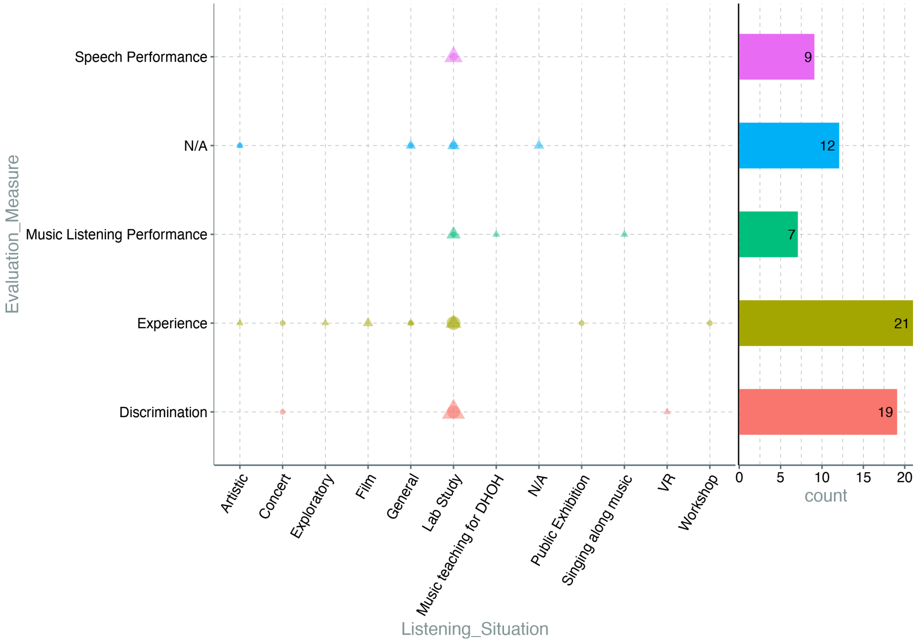


Fig. A.10: Distribution of the evaluation measure goals for each listening situation; a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices; items on X axis are not exclusive.

research conducted under laboratory conditions and not in ecologically valid environments. As such, this contradicts the idea that most long-term benefits are obtained when participants use hearing assistive devices in daily life scenarios. For these reasons we can see an gap in the evaluation and experimental protocols conducted in most studies included, and we suggest that researchers start focusing on larger scale, longitudinal studies that are more akin to clinical ones, when evaluating their audio augmenting tactile displays.

Looking at the hardware aspect, the majority of included studies present tactile displays designed for some regions of the hand. This is expected as hands provide the highest fine motor skills as well as very good spatio-temporal resolution, but it might not be the most practical for devices that are designed to be used extensively, especially when daily activities are to be executed while using the tactile displays. Furthermore, other areas could present different advantages in terms of duration of stimulation or intensity tolerance for use cases where the finest discrimination properties are not vital. Therefore, we see a great opportunity to branch out and encourage designers and researchers to create displays that afford similar perceptual characteristics and are to be sensed by different body regions. Furthermore,

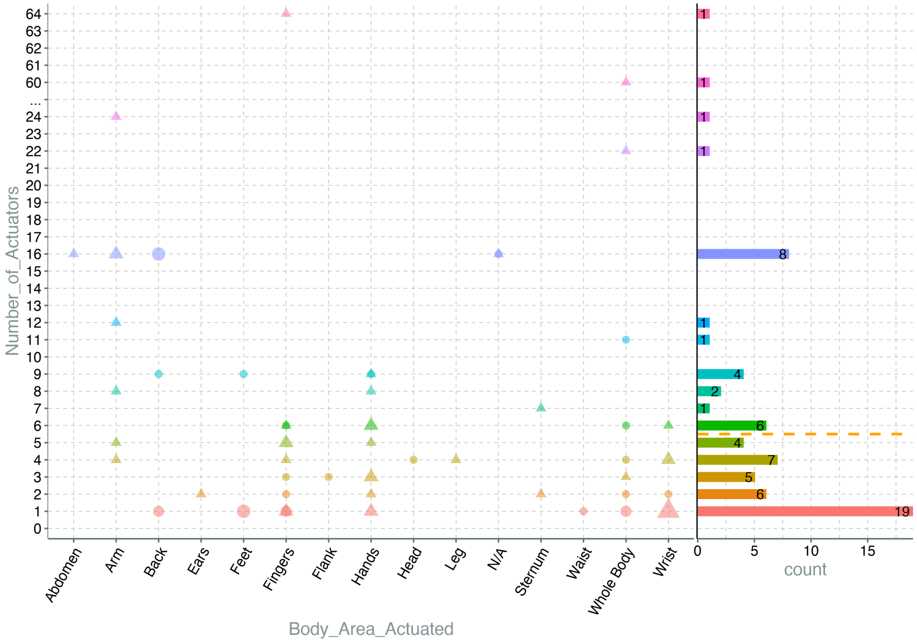


Fig. A.11: Distribution of number of actuators over the body area they actuate; a triangle indicate a wearable device, a circle indicates a fixed device and the size of the shape represents the amount of similar devices. The dashed orange line represents the mean.

versatile designs are strongly encouraged (in terms of mode of interaction, HW/SW and mapping), in order to be adaptable to the inter-user needs; this is especially important for CI users, where the variation in hearing abilities is largest.

On a similar note, our finding indicates that researchers present a large variety of designs in terms of type of actuators, mapping, signal processing, etc. further solidify the exploratory phase of the entire research field, characteristic of an early development stage. There seems to be a consensus on the upper limit of the number of actuators necessary for vibrotactile augmentation, although more than half of the displays identified use less than the average number of transducers, with the mode being 1 actuator. This could be attributed to cost reduction strategies, but since most of the devices are researched almost exclusively in laboratories and are generally far from commercialisation, we are confident to suggest that a high number of actuators might not provide substantial benefits to tactile augmentation of audio. With this in mind, we must emphasize the importance of mapping strategies used, both in the time/frequency domain and in the psychoperceptual space, in order to design the best tactile displays for vibrotactile augmentation. Our research pool shows that almost half of the devices studied do not imply any form of mapping between the auditory signal and tactile stimulation, while the mode is *pitch-to-position*—a mapping scheme introduced with the very first audio-tactile augmentation device. We see a great potential for exploration into creative mapping schemes that could have roots in the most commonly encountered ones, as well as radical new ideas that could be generic or case-specific (e.g., bespoke for concert or film scenarios). These new mappings should be carefully designed and evaluated primarily with respect to the target group's hearing profiles(CI, HA, etc.) as well as signal processing used and the eventual acoustic stimulation; if possible, all these aspects should be co-created involving the end users, in order to produce a coherent multisensory experience.

A.4.2 Limitations

This is the first scoping review focusing on the technological aspect behind the vibrotactile augmentation of music. Although mainly concerned with the hardware and software characteristics of the tactile displays described in Table A.1, this article has addressed some elements of the dissemination and evaluation of the devices described. Nevertheless, the scoping nature of this review rules out a detailed description of implementation for each study or evaluates the quality and effectiveness of the included tactile display. Therefore, it is impossible to recommend specific techniques or strategies that would predict better music perception, training, or adjacent metrics for D/HOH and CI implanted people.

While a comprehensive search has been conducted on one of the most relevant databases, this process was carried out by a single reviewer and there was no forward-citation search on the included studies. Furthermore, there was no review of the reference list of included articles or a manual search protocol to scan relevant journals, as it was concluded that most of the articles are indexed by Scopus®. This resulted in the exclusion of any gray literature, as the process of searching for relevant unpublished material was of considerable difficulty.

A.4.3 Conclusions

The purpose of this scoping review was to investigate and report the current technological state in the field of vibrotactile augmentation, viewed from the perspective of music enhancement for hearing impaired users. A total of 3555 articles were considered for eligibility from the Scopus® database, resulting in the inclusion of 63 studies. The vibrotactile devices in each article was analyzed according to a pre-defined set of characteristics, focusing on hardware and software elements, as well as the evaluation and experiment design, regardless of the hearing profile of their users. The evidence gathered indicates that this research field is in an early phase, characterized by an exploratory approach and preliminary results. A secondary objective of this article was to identify the gaps and trends in the literature that can guide researchers and designers in their practice, and a list of suggestions and recommendations has been presented, based on graphical representations of statistics analysis. The data and the system used to synthesize the review are publicly accessible, and we recommend that readers explore them and generate their own graphs and interpretations.

References

[1] S. C. Aker, H. Innes-Brown, K. F. Faulkner, M. Vatti, and J. Marozeau, "Effect of audio-tactile congruence on vibrotactile music enhancement," *The Journal of the Acoustical Society of America*, vol. 152, no. 6, pp. 3396–3409, 2022.

[2] F. A. Araujo, F. L. Brasil, A. C. L. Santos, L. D. S. B. Junior, S. P. F. Dutra, and C. E. C. F. Batista, "Auris system: Providing vibrotactile feedback for hearing impaired population," *BioMed Research International*, vol. 2017, 2017.

[3] J. Armitage and K. Ng, "Feeling sound: Exploring a haptic-audio relationship," pp. 146–152, 2016.

[4] A. Baijal, J. Kim, C. Branje, F. A. Russo, and D. Fels, "Composing vibrotactile music: A multi-sensory experience with the emoti-chair," *Haptics Symposium 2012, HAPTICS 2012 - Proceedings*, pp. 509–515, 2012.

[5] H.-H. Bothe, A. F. Mohamad, and M. R. Jensen, "Flipps: Fingertip stimulation for improved speechreading," pp. 1152–1159, 2004.

[6] C. Branje, M. Maksimowski, M. Karam, D. Fels, and F. A. Russo, "Vibrotactile display of music on the human back," *3rd International Conference on Advances in Computer-Human Interactions, ACHI 2010*, pp. 154–159, 2010.

[7] C. Branje, G. Nespoil, F. A. Russo, and D. Fels, "The effect of vibrotactile stimulation on the emotional response to horror films," *Computers in Entertainment*, vol. 11, 2014.

[8] K. P. Burnham and D. R. Anderson, "Multimodel inference: Understanding aic and bic in model selection," *Sociological Methods and Research*, vol. 33, 2004.

[9] C. Chafe, "Tactile audio feedback," in *Opening a New Horizon: Proceedings of the 1993 International Computer Music Conference, ICMC*

1993, Tokio, Japan, September 10-15, 1993. Michigan Publishing, 1993. [Online]. Available: http://hdl.handle.net/2027/spo.bbp2372.1993.013

[10] K. Cieśla, T. Wolak, A. Lorens, B. Heimler, H. Skarżyński, and A. Amedi, "Immediate improvement of speech-in-noise perception through multisensory stimulation via an auditory to tactile sensory substitution," *Restorative Neurology and Neuroscience*, vol. 37, pp. 155–166, 2019.

[11] R. S. C. Cowan, P. J. Blamey, K. L. Galvin, J. Z. Sarant, J. I. Alcántara, and G. M. Clark, "Perception of sentences, words, and speech features by profoundly hearing-impaired children using a multichannel electrotactile speech processor," *The Journal of the Acoustical Society of America*, vol. 88, 1990.

[12] N. DeGuglielmo, C. Lobo, E. J. Moriarty, G. Ma, and D. E. Dow, "Haptic vibrations for hearing impaired to experience aspects of live music," in *Bio-Inspired Information and Communications Technologies*, T. Nakano, Ed. Sugimoto, Japan: Springer International Publishing, 2021, pp. 71–86.

[13] D. Egloff, M. Wanderley, and I. Frissen, "Haptic display of melodic intervals for musical applications," *IEEE Haptics Symposium, HAPTICS*, vol. 2018-March, pp. 284–289, 2018.

[14] A. Engebretson and M. O'Connell, "Implementation of a microprocessor-based tactile hearing prosthesis," *IEEE Transactions on Biomedical Engineering*, vol. BME-33, pp. 712–716, 1986.

[15] M. Fletcher, "Using haptic stimulation to enhance auditory perception in hearing-impaired listeners," *Expert Review of Medical Devices*, vol. 18, pp. 63–74, 2021.

[16] M. Fletcher, N. Thini, and S. Perry, "Enhanced pitch discrimination for cochlear implant users with a new haptic neuroprosthetic," *Scientific Reports*, vol. 10, 2020.

[17] M. Fletcher and J. Zgheib, "Haptic sound-localisation for use in cochlear implant and hearing-aid users," *Scientific Reports*, vol. 10, 2020.

[18] H. Florian, A. Mocanu, C. Vlasin, J. Machado, V. Carvalho, F. Soares, A. Astilean, and C. Avram, "Deaf people feeling music rhythm by using a sensing and actuating device," *Sensors and Actuators, A: Physical*, vol. 267, pp. 431–442, 2017.

[19] J. Foxe, G. Wylie, A. Martinez, C. Schroeder, D. Javitt, D. Guilfoyle, W. Ritter, and M. Murray, "Auditory-somatosensory multisensory processing in auditory association cortex: An fmri study," *Journal of neurophysiology*, vol. 88, pp. 540–3, 08 2002.

[20] E. Frid and H. Lindetorp, "Haptic music - exploring whole-body vibrations and tactile sound for a multisensory music installation," *Proceedings of the Sound and Music Computing Conferences*, vol. 2020-June, pp. 68–75, 2020.

[21] S. A. Gilmore and F. A. Russo, "Neural and behavioral evidence for vibrotactile beat perception and bimodal enhancement," *Journal of Cognitive Neuroscience*, vol. 33, 2020.

[22] C. Giulia, D. Chiara, and H. Esmailbeigi, "Glos: Glove for speech recognition," *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, vol. 2019, pp. 3319–3322, 2019.

[23] S. Hashizume, S. Sakamoto, K. Suzuki, and Y. Ochiai, "Livejacket: Wearable music experience device with multiple speakers," pp. 359–371, 2018.

[24] L. Hayes, "Skin music (2012): An audio-haptic composition for ears and body," *C and C 2015 - Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, pp. 359–360, 2015.

[25] A. Haynes, J. Lawry, C. Kent, and J. Rossiter, "Feelmusic: Enriching our emotive experience of music through audio-tactile mappings," *Multimodal Technologies and Interaction*, vol. 5, 2021.

[26] C. Hopkins, S. Maté-Cid, R. Fulford, G. Seiffert, and J. Ginsborg, "Vibrotactile presentation of musical notes to the glabrous skin for adults with normal hearing or a hearing impairment: Thresholds, dynamic range and high-frequency perception," *PLoS ONE*, vol. 11, 2016.

[27] ——, "Perception and learning of relative pitch by musicians using the vibrotactile mode," *Musicae Scientiae*, 2021.

[28] M. Hove, S. Martinez, and J. Stupacher, "Feel the bass: Music presented to tactile and auditory modalities increases aesthetic appreciation and body movement," *Journal of Experimental Psychology: General*, 2019.

[29] J. Huang, T. Lu, B. Sheffield, and F.-G. Zeng, "Electro-tactile stimulation enhances cochlear-implant melody recognition: Effects of rhythm and musical training," *Ear and Hearing*, vol. 41, pp. 106–113, 2018.

[30] J. Huang, B. Sheffield, P. Lin, and F.-G. Zeng, "Electro-tactile stimulation enhances cochlear implant speech recognition in noise," *Scientific Reports*, vol. 7, 2017.

[31] I. Hwang, H. Lee, and S. Choi, "Real-time dual-band haptic music player for mobile devices," *IEEE Transactions on Haptics*, vol. 6, pp. 340–351, 2013.

[32] R. Jack, A. Mcpherson, and T. Stockman, “Designing tactile musical devices with and for deaf users: a case study,” *Proc. of the International Conference on the Multimedia Experience of Music*, 2015.

[33] N. T. Jiam and C. J. Limb, “Rhythm processing in cochlear implant-mediated music perception,” *Annals of the New York Academy of Sciences*, vol. 1453, no. 1, pp. 22–28, 2019.

[34] L. A. Jones, *Haptics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 1808–1811. [Online]. Available: https://doi.org/10.1007/978-3-540-29678-2_2140

[35] J. Kanebako and K. Minamizawa, “Vibgrip++: Haptic device allows feeling the music for hearing impaired people,” *Lecture Notes in Electrical Engineering*, vol. 432, pp. 449–452, 2018.

[36] M. Karam, C. Branje, G. Nespoli, N. Thompson, F. Russo, and D. Fels, “The emoti-chair: An interactive tactile music exhibit,” *Conference on Human Factors in Computing Systems - Proceedings*, pp. 3069–3074, 2010.

[37] M. Karam and D. Fels, “Designing a model human cochlea: Issues and challenges in crossmodal audio-haptic displays,” *Proceedings of the 2008 Ambi-Sys Workshop on Haptic User Interfaces in Ambient Media Systems, HAS 2008*, 2008.

[38] C. Kayser, C. I. Petkov, M. Augath, and N. K. Logothetis, “Integration of touch and sound in auditory cortex,” *Neuron*, vol. 48, 2005.

[39] H. Knutzen, T. Kvifte, and M. M. Wanderley, “Vibrotactile feedback for an open air music controller,” vol. 8905, 2014.

[40] A. P. Lauzon, F. A. Russo, and L. R. Harris, “The influence of rhythm on detection of auditory and vibrotactile asynchrony,” *Experimental Brain Research*, vol. 238, 2020.

[41] Lofelt, “Basslet,” https://eu.lofelt.com/, 2017, accessed: 2018-06-30.

[42] M. J. Lucía, P. Revuelta, Á. García, B. Ruiz, R. Vergaz, V. Cerdán, and T. Ortiz, “Vibrotactile captioning of musical effects in audio-visual media as an alternative for deaf and hard of hearing people: An eeg study,” *IEEE Access*, vol. 8, pp. 190 873–190 881, 2020.

[43] X. Luo and L. Hayes, “Vibrotactile stimulation based on the fundamental frequency can improve melodic contour identification of normal-hearing listeners with a 4-channel cochlear implant simulation,” *Frontiers in Neuroscience*, vol. 13, 2019.

[44] A. Marzvanyan and A. F. Alhawaj, *Physiology, Sensory Receptors*. StatPearls Publishing LLC., 2019.

[45] A. Mazzoni and N. Bryan-Kinns, "Mood glove: A haptic wearable prototype system to enhance mood music in film," *Entertainment Computing*, vol. 17, pp. 9–17, 2016.

[46] P. Milnes, J. Stevens, B. Brown, I. Summers, and P. Cooper, "Use of microcontroller in a tactile aid for the hearing impaired," *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pp. 413–414, 1996.

[47] M. Mirzaei, P. Kán, and H. Kaufmann, "Earvr: Using ear haptics in virtual reality for deaf and hard-of-hearing people," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, pp. 2084–2093, 2020.

[48] R. Miyamoto, A. Robbins, M. Osberger, S. Todds, A. Riley, and K. Kirk, "Comparison of multichannel tactile aids and multichannel cochlear implants in children with profound hearing impairments," *American Journal of Otology*, vol. 16, pp. 8–13, 1995.

[49] J. W. Morley and M. J. Rowe, "Perceived pitch of vibrotactile stimuli: effects of vibration amplitude, and implications for vibration frequency coding." *The Journal of Physiology*, vol. 431, 1990.

[50] K. Nakada, J. Onishi, and M. Sakajiri, "An interactive musical scale presentation system via tactile sense through haptic actuator," *HAVE 2018 - IEEE International Symposium on Haptic, Audio-Visual Environments and Games, Proceedings*, pp. 13–16, 2018.

[51] S. Nanayakkara, L. Wyse, and E. Taylor, "The haptic chair as a speech training aid for the deaf," *Proceedings of the 24th Australian Computer-Human Interaction Conference, OzCHI 2012*, pp. 405–410, 2012.

[52] S. C. Nanayakkara, L. Wyse, S. H. Ong, and E. A. Taylor, "Enhancing musical experience for the hearing-impaired using visual and haptic displays," *Human-Computer Interaction*, vol. 28, 2013.

[53] R. Okazaki, T. Hachisu, M. Sato, S. Fukushima, V. Hayward, and H. Kajimoto, "Judged consonance of tactile and auditory frequencies," 04 2013, pp. 663–666.

[54] R. Okazaki, H. Kajimoto, and V. Hayward, "Vibrotactile stimulation can affect auditory loudness: A pilot study," 06 2012, pp. 103–108.

[55] M. Osberger, A. Robbins, S. Todd, and C. Brown, "Initial findings with a wearable multichannel vibrotactile aid," *American Journal of Otology*, vol. 12, pp. 179–182, 1991.

[56] O. Ozioko, W. Taube, M. Hersh, and R. Dahiya, "Smartfingerbraille: A tactile sensing and actuation based communication glove for deafblind people," *IEEE International Symposium on Industrial Electronics*, pp. 2014–2018, 2017.

[57] C. Pacchierotti, S. Sinclair, M. Solazzi, A. Frisoli, V. Hayward, and D. Prattichizzo, "Wearable haptic systems for the fingertip and the hand: Taxonomy, review, and perspectives," *IEEE Transactions on Haptics*, vol. 10, 2017.

[58] S. Papetti and C. Saitis, *Musical Haptics: Introduction*. Springer, Cham, 2018.

[59] A. Pavlidou and B. Lo, "Artificial ear - a wearable device for the hearing impaired," *2021 IEEE 17th International Conference on Wearable and Implantable Body Sensor Networks, BSN 2021*, vol. 2021-January, 2021.

[60] B. Petry, T. Illandara, D. Elvitigala, and S. Nanayakkara, "Supporting rhythm activities of deaf children using music-sensory-substitution systems," *Conference on Human Factors in Computing Systems - Proceedings*, vol. 2018-April, 2018.

[61] B. Petry, T. Illandara, and S. Nanayakkara, "Muss-bits: Sensor-display blocks for deaf people to explore musical sounds," *Proceedings of the 28th Australian Computer-Human Interaction Conference, OzCHI 2016*, pp. 72–80, 2016.

[62] E. Pezent, B. Cambio, and M. Ormalley, "Syntacts: Open-source software and hardware for audio-controlled haptics," *IEEE Transactions on Haptics*, vol. 14, pp. 225–233, 2021.

[63] E. Pezent, M. O'Malley, A. Israr, M. Samad, S. Robinson, P. Agarwal, H. Benko, and N. Colonnese, "Explorations of wrist haptic feedback for ar/vr interactions with tasbi," *Conference on Human Factors in Computing Systems - Proceedings*, 2020.

[64] C. M. Reed, H. Z. Tan, Y. Jiao, Z. D. Perez, and E. C. Wilson, "Identification of words and phrases through a phonemic-based haptic display," *ACM Transactions on Applied Perception*, vol. 18, 2021.

[65] B. Remache-Vinueza, A. Trujillo-León, M. Zapata, F. Sarmiento-Ortiz, and F. Vidal-Verdú, "Audio-tactile rendering: A review on technology and methods to convey musical information through the sense of touch," *Sensors*, vol. 21, 2021.

[66] T. Ro, J. Hsu, N. Yasar, C. Elmore, and M. Beauchamp, "Sound enhances touch perception," *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, vol. 195, pp. 135–43, 04 2009.

[67] J. Rouger, S. Lagleyre, B. Fraysse, S. Deneve, O. Deguine, and P. Barone, "Evidence that cochlear-implanted deaf patients are better multisensory integrators," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, 2007.

[68] F. A. Russo, "Multisensory processing in music," *The Oxford Handbook of Music and the Brain*, 2019.

[69] A. Rydel and W. Seiffer, "Untersuchungen über das vibrationsgefühl oder die sog. knochensensibilität (pallästhesie)," *Archiv für Psychiatrie und Nervenkrankheiten*, vol. 37, 1903.

[70] R. Sakuragi, S. Ikeno, R. Okazaki, and H. Kajimoto, "Collarbeat: Whole body vibrotactile presentation via the collarbone to enrich music listening experience," *International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments, ICAT-EGVE 2015*, pp. 141–146, 2015.

[71] W. Setzepfand, "Zur frequenzabhängigkeit der vibrationsempfindung des menschen," *Z. Biol*, vol. 96, pp. 236–240, 1935.

[72] A. Sharp, B. Bacon, and F. Champoux, "Enhanced tactile identification of musical emotion in the deaf," *Experimental Brain Research*, vol. 238, pp. 1229–1236, 2020.

[73] S. Shin, C. Oh, and H. Shin, "Tactile tone system: A wearable device to assist accuracy of vocal pitch in cochlear implant users," 2020.

[74] P. B. Shull and D. D. Damian, "Haptic wearables as sensory replacement, sensory augmentation and trainer - a review," *Journal of NeuroEngineering and Rehabilitation*, vol. 12, 2015.

[75] F. Sorgini, R. Caliò, M. C. Carrozza, and C. M. Oddo, "Haptic-assistive technologies for audition and vision sensory disabilities," *Disability and Rehabilitation: Assistive Technology*, vol. 13, 2018.

[76] K.-E. Spens, C. Huss, M. Dahlqvist, and E. Agelfors, "A hand held two-channel vibro-tactile speech communication aid for the deaf: Characteristics and results," *Scandinavian Audiology, Supplement*, vol. 26, pp. 7–13, 1997.

[77] B. Stein, P. Stein, and M. Meredith, *The Merging of the Senses*, ser. A Bradford book. MIT Press, 1993. [Online]. Available: https://books.google.se/books?id=uCV9QgAACAAJ

[78] B. A. Swerdfeger, J. Fernquist, T. W. Hazelton, and K. E. MacLean, "Exploring melodic variance in rhythmic haptic stimulus design," 2009.

[79] B. Tessendorf, D. Roggen, M. Spuhler, T. Stiefmeier, G. Tröster, T. Grämer, M. Feilner, and P. Derleth, "Design of a bilateral vibrotactile feedback system for lateralization," *ASSETS'11: Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 233–234, 2011.

[80] P. Tranchant, M. Shiell, M. Giordano, A. Nadeau, I. Peretz, and R. Zatorre, “Feeling the beat: Bouncing synchronization to vibrotactile music in hearing and early deaf people,” *Frontiers in Neuroscience*, vol. 11, 2017.

[81] A. C. Tricco, E. Lillie, W. Zarin, K. K. O’Brien, H. Colquhoun, D. Levac, D. Moher, M. D. Peters, T. Horsley, L. Weeks, S. Hempel, E. A. Akl, C. Chang, J. McGowan, L. Stewart, L. Hartling, A. Aldcroft, M. G. Wilson, C. Garritty, S. Lewin, C. M. Godfrey, M. T. MacDonald, E. V. Langlois, K. Soares-Weiser, J. Moriarty, T. Clifford, Özge Tunçalp, and S. E. Straus, “Prisma extension for scoping reviews (prisma-scr): Checklist and explanation,” *Annals of Internal Medicine*, vol. 169, no. 7, pp. 467–473, 2018, pMID: 30178033. [Online]. Available: https://doi.org/10.7326/M18-0850

[82] U. Trivedi, R. Alqasemi, and R. Dubey, “Wearable musical haptic sleeves for people with hearing impairment,” *ACM International Conference Proceeding Series*, pp. 146–151, 2019.

[83] L. Turchet, T. West, and M. M. Wanderley, “Touching the audience: musical haptic wearables for augmented and participatory live music performances,” *Personal and Ubiquitous Computing*, 2020.

[84] A. Vallgårda, L. Boer, and B. Cahill, “The hedonic haptic player,” *International Journal of Design*, vol. 11, pp. 17–33, 2017.

[85] C. Wada, H. Shoji, and T. Ifukube, “Development and evaluation of a tactile display for a tactile vocoder,” *Technology and Disability*, vol. 11, pp. 151–159, 1999.

[86] J. Weisenberger and S. Broadstone, “Evaluation of two multichannel tactile aids for the hearing impaired,” *Journal of the Acoustical Society of America*, vol. 86, pp. 1764–1775, 1989.

[87] T. West, A. Bachmayer, S. Bhagwati, J. Berzowska, and M. Wanderley, “The design of the body:suit:score, a full-body vibrotactile musical score,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11570 LNCS, pp. 70–89, 2019.

[88] A. Wilska, “On the vibrational sensitivity in different regions of the body surface,” *Acta Physiologica Scandinavica*, vol. 31, 1954.

[89] E. Yeung, A. Boothroyd, and C. Redmond, “A wearable multichannel tactile display of voice fundamental frequency,” *Ear and Hearing*, vol. 9, 1988.

[90] G. Young, D. Murphy, and J. Weeter, “Audio-tactile glove,” *DAFx 2013 - 16th International Conference on Digital Audio Effects*, 2013.

[91] ——, "Haptics in music: The effects of vibrotactile stimulus in low frequency auditory difference detection tasks," *IEEE Transactions on Haptics*, vol. PP, pp. 1–1, 12 2016.

[92] ——, "Vibrotactile discrimination of pure and complex waveforms," 2015, pp. 359–362.

Paper B

The Relationship between Frequency and Hand Region Actuated

R. Paisa, J. Andersen, N. C. Nilsson and S. Serafin

The paper has been published in
11th International Workshop on Haptic and Audio Interaction Design(HAID), London, UK Conference Proceedings, 2022.

The layout has been revised.

Abstract

While the frequency response of the skin is described at macro level, there is a need to explore discrete areas of interest. The experiments described in this paper are part of a project that aims to build devices for cochlear implant (CI) users that meet music listening needs. The aim is to demonstrate that constant amplitude vibrotactile stimuli with distinct frequencies excite different areas of the hand with varying perceived intensity. 65 subjects took part in two within-subject experiments investigating the areas of the hand with most intense perceived sensation when exposed to various stimuli. Multinomial logistic regression was performed on the data and it was concluded that particular signals will elicit stronger sensations on some regions of the hand, and weaker on others. This indicates that there is a correlation between frequency of the stimuli and the area of the hand mostly stimulated.

B.1 Introduction

The skin is the largest organ of the human body, and besides its protective and temperature regulation functions, it affords the sense of touch - together with the somatosensory system. The psycho-perceptual properties of the skin, with respect to touch, have been studied extensively from multiple divergent angles. For the sake of conciseness, this article approaches the topic from the standpoint of musical augmentation for hearing impaired and cochlear implant (CI) users. Specifically, it demonstrates that distinct areas of the hand elicit a frequency dependent perception of vibrotactile stimuli, when stimulated from the same location. These findings allow tactile display designers to create novel devices with less technical complexity implied by multiple actuators.

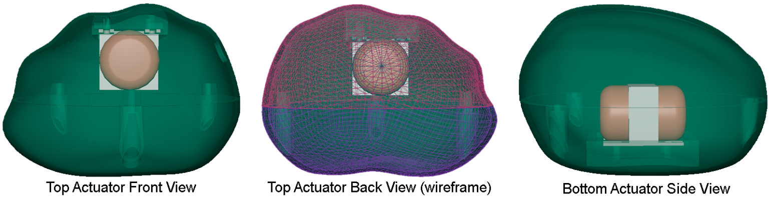

Top Actuator Front View
Top Actuator Back View (wireframe)
Bottom Actuator Side View

Fig. B.1: See-through renders of the two versions of the VAM

B.2 Background and related work

Tactile stimuli are perceived through a series of complex mechanisms part of the somatosensory system; the fundamental agents are four types of mechanoreceptors: slowly adapting types 1 and 2, the Meissner corpuscle (colloquially known as Rapid Adapting, RA) and the Pacinian corpuscle. These types of receptors respond differently to stimuli and elicit different sensations. In a musical context, the Meissner corspuscles and the Pacinian ones are known to be most useful [3, 6, 14]. The former ones have a high innervation density, and respond mostly to the range of 1Hz - 100Hz, with a peak sensitivity around 40Hz; the latter have a broader range from 40Hz to 1000Hz, with maximum sensitivity around the 250Hz mark. Due to their larger size, Pacinian receptors have a lower spatial resolution compared to the Meissner corpuscles, and are known to respond mostly to vibrations [3, 6, 14]. While the parallel to the ear can be drawn, the frequency discrimination of the skin is far more complex and it is influenced by the amplitude, duration or the location of the stimuli, and not least by the multi-channel nature of the sensing system [1, 10]. Because there is such a high variation, it would be incorrect to discuss of a single frequency response, therefore one should think of a range of responses, and focus on isolated body locus if possible [3, 5]. In this paper we focus on the hand. The hand is one of the regions with the highest number of mechanoreceptors, specifically, the fingers are sensitive to a very wide frequency range from 0.5Hz to 900Hz [5, 15]. However, these values are frequently challenged by new research, and there appears to be unknown confounding variables that affect these limits. A similar disagreement seems relevant to discussions about frequency discrimination thresholds, with researchers claiming that a difference of 4.6% would be enough to perceive different stimuli, while others suggest numbers closer to 30% [5, 7, 11]. When it comes to the spatial resolution measured by the classic cutaneous two-point discrimination test, the hand has thresholds between are 2mm-8mm, depending on the area [8]. Although sound and vibrotactile stimuli both are repeated changes in mechanical pressure captured by different receptors, it is hard to equate the two perception processes due to the radical differences in the sensing apparatus, as well as the lack of consensus in the literature. The mechanism linking auditory and tactile sensations is called multisensory integration, pioneered by [13]. It describes how humans form coherent, valid, and robust perception of reality, by processing sensory stimuli from various modalities. Multisensory integration suggests that enhancement can occur only for stimuli that are temporally coincident and propose that enhancement is strongest for those stimuli that individually are least effective. This mechanism seems to be playing a crucial role in the rehabilitation process of hearing loss and CI process, and it has been proven that deaf and CI users end up being better multisensory integrators than their hearing peers [12].

B.3 Materials

The tactile display used in this project is build around a Tactuator BM1C transducer, build by Tactile Labs, and was named VAM: *Vibrotactile Actuator for Music*. Other transducers like Haptuator Mk1 and Mk2 from the same manufacturer were tested in the setup, but their amplitude was too low.

The Tactuator BM1C provides a much wider frequency response than can be perceived cutaneously, and the distortion (if any) is controlled due to the dampers encapsulating the actuator. These characteristics made it suitable for this project.

The tactile display was made exclusively for left hand usage, and had an ovoid shape as seen in figure B.1 with the following dimensions: 84mm wide, 58mm tall and 89mm deep. The shape was inspired by the resting hand position when fixed with an orthopedic splint. This pose should minimize the strain on the wrist while allowing the rest of the hand to relax, ensuring a similar holding pressure across the entire tactile display area. The enclosure was first modeled in clay ensuring that each digit has an ergonomic socket to rest into. The clay artefact was 3D scanned using Autodesk ReCap, by analyzing 40 still images of the subject, taken from multiple angles with a Fujifilm X-T1 camera and a Fujinon XF 35mm @ f2.0 lens. Throughout this process, the clay prototype was suspended in mid air using transparent fishing line, in order to allow images to be captured from all angles. The resulting scan resulted in a high fidelity model but with chaotic topology, therefore a new 3D model was created, using the scan as an outline. This resulted in an accurate digital replica of the intended surface, ready for 3D printing. The model was split horizontally in two halves, that were hold together by three M3x16 bolts, and printed on an Ultimaker3 using PLA material. Two versions were constructed, one with the actuator resting on the bottom half, and the other with the actuator attached to the top half. This was done in order to be able to run experiments that account for the location of the actuator in respect to areas measured as shown in figure B.2.

In order to ensure that the display had no resonant peaks that could influence the perception, its frequency response was measured on the interval 0-1000kHz. The result was an average of 9 measured on 3 different locations, using a single axis from an ADXL355 analog accelerometer: tip of the middle finger socket, tip of thumb socket and lower palm . After measuring the frequency response, an approximate filter with inverse response was designed as a cascade of second order bandpass biquad filters with independent gains and bandwidths. The detailed filter parameters can be found in the table B.1 below.

Moreover, an attempt to account for the frequency response of the skin as reported by Merchel et.al was made with the same cascading biquad filters approach [9]. The goal was to ensure constant supra-threshold stimulation across the entire band of interest.

Frequency (Hz)	15	30	60	100	150	200	250	300	500	800	1000
Bandwidth (Hz)	30	30	40	50	50	50	100	150	300	200	400
Gain (dB)	31	31	25	15	7	2	0	2	8	15	22

Table B.1: Filter parameters for compensation of resonant frequencies

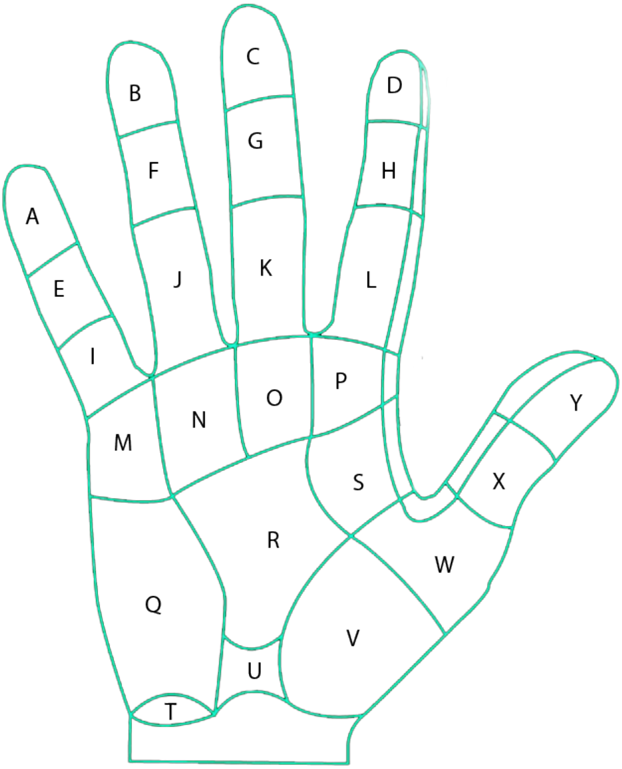


Fig. B.2: Palm areas segmentation

B.4 Evaluation

B.4.1 Goal

The experiments are part of a study on single-actuator, high-fidelity vibrotactile devices to be used by cochlear implant persons when listening to music. These experiments are at pre-paradigm stage, and they aim to explore the informal findings from previous observations and reports: the relationship between the frequency of the vibrotactile actuator and the area of hand where users perceived said vibrations. This goal translates into the following hypothesis: *people perceive distinct frequencies in different areas of the hand (thumb, fingers, palm, wrist, etc), when stimulated by single-actuator vibrotactile handheld device with a constant supra-threshold amplitude.*

Moreover, the relation between frequencies and their areas of maximal perception was explored. The aim was to determine a potential mapping between vibrotactile frequency and excited hand area.

B.4.2 Experiments

Two within-subjects designed experiments were conducted with 35 (25M, 10F) and 32 (23M, 9F) participants respectively, that required them to indicate the areas of the hand where they felt the vibration the most. The only difference between the two experiments was the position of the actuator inside the VAM. Both designs can be seen in figure B.1; experiment 1 was conducted with the actuator in the bottom half. The experiment had a multi-alternative (28) three forced choice design, that represents areas of the hand as seen in figure B.2. The sinusoidal stimuli presented were first compensated according to the frequency response of the handheld device and then calibrated for constant supra threshold level (reference = 250Hz) with a cascade of 2nd order biquad filters with F0 described in [9]. The 11 different stimuli presented had fixed frequencies ranging from from 15Hz to 1000Hz, separated by a major third intervalB.1 Each stimuli lasted for 2 seconds, as there is evidence for no change in vibrotactile threshold of detection for stimuli longer that 1 second, although threshold could be higher for simulations shorter than 1 second [4, 14].

B.4.3 Procedure

Participants were recruited among students and staff at Aalborg University Copenhagen. Participation was voluntary and no compensation was offered. The experiments lasted about 15 minutes and consisted of an accommodation phase followed by 3 simulations for each of the 11 frequencies. In the accommodation phase, the participants were instructed on the posture; they had to hold their hand on a soft, spongy surface throughout the experiment, to avoid the VAM touching the table, and potentially actuating it. Furthermore, they were exposed to an accommodation vibrotactile stimuli lasting one minute - a snipped of electronic music, in order to introduce the range of sensations possible, and allow the users to find a comfortable posture and relax the hand. When the music stimuli finished playing, the participant could proceed to the next phase.

Throughout the experimental phase, the participants were presented with one of the 11 stimuli, and were requested to rank the top 3 areas where they perceived the stimulation the most, by using the mouse to select the zones from the user interface (identical to figure B.2). The participants had the opportunity to change the areas after initial selection, but they could not experience the stimuli again; this was done to ensure consistent exposure times between participants. Once the top 3 areas were selected, the users could continue to the next stimulation. The users had the possibility to select a "Not sensed" category, marked with "Ø" in figure B.4. During the experiment, participants wore a pair of Bose QC700 noise canceling headphones playing pink noise at a comfortable level, with noise cancellation enabled on the highest setting. The data collected was anonymized, without the possibility of matching answers sets with participants. The following was logged: trial number (1-33), stimuli frequency, 1st, 2nd, 3rd selected area, and the log file creation time.

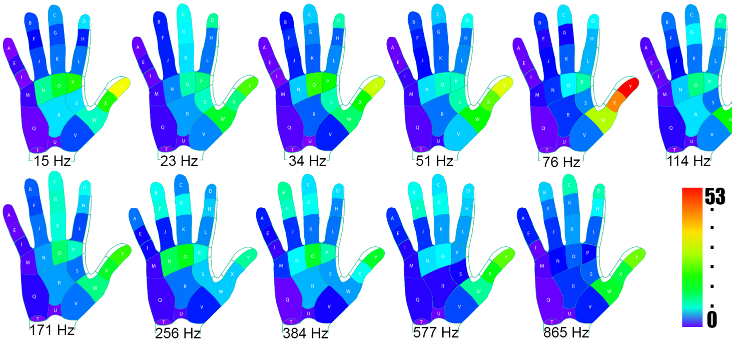
B.5 Results

The data from the two experiments was treated as categorical, and analyzed separately. First, the distribution of preferences was analysed, and then the data was inserted into a predictive model to compare tendencies and determine if the frequency of the stimuli (as the independent variable) can explain the differences in those trends.

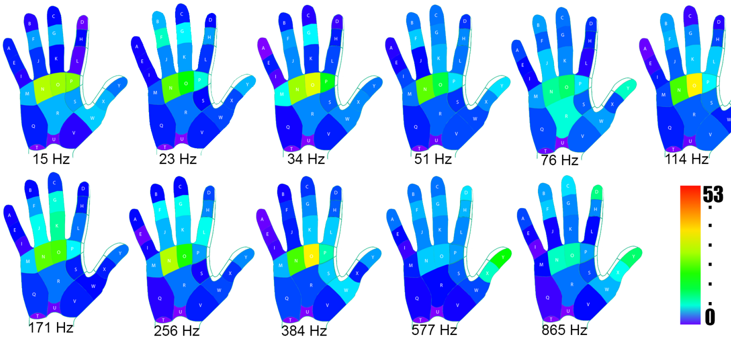
Descriptive Statistics: Figure B.3 shows the sum of preferences for each frequency. We present the sum instead of the top 3 individual rankings for two reasons: (1) For experiment 1, the highest ranking areas across the conditions was always option ″Y″ - the closest to the physical location to the actuator. This is assumed to be a bias towards that location, but a further test is planned to confirm it. (2) During post-experiment interviews, participants reported not ranking the sensations based on perceived intensity, but instead the presence of sensation in the selected locations. These two factors pointed towards a low validity of reported ranks, thus each answer was treated with equal importance. Initial results show that the perception of all frequencies seem to peak towards finger tips, as well as the palmar digital area, and be almost not existent closest to the wrist (areas, ″T″,″U″, ″Q″); a fact that can be explained by the shape, size and grip of the VAM. Besides that, it's noticeable that different frequencies excited distinct areas of the hand, especially at 171Hz, 256Hz and 384Hz for experiment 1 and 51Hz, 76Hz, and 577Hz for experiment 2. These areas are felt around the proximal phalanges while other frequencies are perceived on areas closest to the actuator's position (X-Y for experiment 1, and N-O for experiment 2). The three frequencies from experiment 1 are closest to the peak response of the Pacinian Corpuscles [2], indicating that even though there could be an amplitude based bias (due to the physical proximity to the actuator), some frequencies elicit vibrotactile sensations in different areas of the hand. However, this cannot be concluded based on a frequency-location relation alone.

Multinomial regression model with mixed effects: To understand if there is a difference between the areas, a multinomial regression model was created that predicts the probability of selecting any of the areas by the explanatory variable: the frequency, that was treated as numerical and continuous. This is a matter of trying to parameterize the probability of selecting any of the areas in terms of baseline probability and the effect of the frequency. However, first it is important to account for the repeated measurement, by introducing a user-based random effect, and only after account for the fixed effect (frequency).

In order to run the multinomial regression model without over-parameterizing it, a reference area was selected, and the probability of choosing another area than the the reference was computed, for each frequency. In the case of ex-



(a) Experiment 1 (actuator in the bottom half)

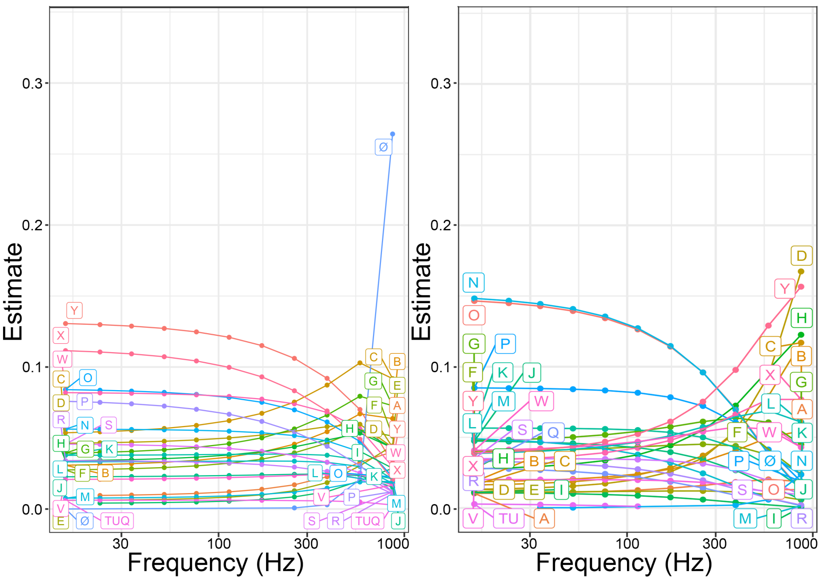


(b) Experiment 2 (actuator in the top half)

Fig. B.3: Distribution of summed preferred areas

periment 1, the reference area was ″Y″, while for the experiment 2 it was ″O″; this was done to challenge the bias introduced by the difference in amplitude as a result of distance-based energy dissipation within the VAM. In order to reach convergence of the model, some areas with too limited samples were aggregated into a zone; for experiment 1 areas ″T″, ″U″, ″Q″ were combined to create area ″TUQ″, and for experiment 2, area ″T″ and ″U″ were aggregated into zone ″TU″, as seen in figure B.4.

After the prediction model was created, another identical model was created without accounting for frequency. The two models were compared with a *F test* that clarifies whether adding the frequency term is a significant contribution to the prediction model. In the case of both experiments there was a significant difference between the model accounting for frequency and the one that did not, with significance level (p-value) smaller than machine precision (e^{-16} and e^{-14} respectively). This results confirms the hypothesis stated in B.4 and the variability observed in the choice of area can be explained by frequency, as predicted by the models described above.



(a) Experiment 1 Predictions **(b)** Experiment 2 Predictions

Fig. B.4: Predicted distribution of areas over frequency

B.6 Discussion

The results related to located perception of stimuli with different frequencies suggest that there is a significant difference between the stimuli frequencies, in terms of area where they are perceived the most. Nevertheless, there is not enough data relative to the complexity of the model to allow for accurate pairwise analysis (e.g., the chance of selecting middle finger tip compared to the index finger tip, for any given frequency), but the data represented in figure B.4 clearly show that areas closer to the actuator for both experiments (”X”,”Y”,”W” and ”N”,”O”, ”P”) have a much greater chance for lower frequencies, than higher ones. Conversely, some areas closer to fingertips have higher chances for being selected as the frequency increases; this behavior is in line with previous studies investigating the amplitude sensing properties of the skin [5, 15]. As a result, there is evidence that the fingertips have the broadest sensing capacity of the body, not only in terms of amplitude, but also frequency. Moreover, the results suggests that there is a decay in frequency sensing, starting from the tip of the middle finger and decreasing radially.

While both experiments suggest that there are ares that sense differently, the results hint that these differences depend highly on the design of tactile display. The position of the actuator, and probably more important, the contact point for it, influences the distribution of perceived frequencies. Throughout this study the position and the contact point were correlated, therefore it's impossible to identify which of the two explains the variation of the data between experiment one and two. One interesting observation is that the "Not sensed (Ø)" category is more prominent in the first experiment than the second one, especially at the highest frequencies tested, indicating that there is a somewhat higher sensitivity to the upper spectrum when the actuator is closer to the metacarpal area.

To our knowledge, this study is among the first that evaluate the characteristics of the skin on multiple points, with a single actuator, as opposite to the traditional approach of evaluating the sole point of contact. Of course, there is a potential limitation in measuring the frequency response accurately with a single actuator, therefore there's a need to challenge the findings from this article with a different setup that could offer greater control over the stimuli delivered at each area of the hand. Nevertheless, given the nature of the larger project this study belongs to, it is valuable to explore haptic perception in relation to a device designed explicitly for CI users' music listening. This is important to underline as there is a need for further rigorous investigation through the same paradigm, in order to fully understand the vibrotactile properties of the skin.

B.7 Conclusions

This study proposes a single-actuator handheld tactile display that allows users to sense stimuli across the entire spectrum of the skin. Two separate devices were constructed with different actuator placements. The displays used in two users studies that evaluated the perception of vibrotactile stimuli with various frequencies, over 27 discrete areas. The results indicate that there is a difference in perception area that is dependent of the frequency of the stimuli. While the data is insufficient for accurate comparison of each frequency-area combination, at least by using multinomial regression, the probability of selecting areas closest to the actuator is higher at lower frequencies and the finger tips show highest sensitivity as the frequency increases. For years, there has been an interest in the vibrotactile properties of the human skin. This research contributes to this debate, by showing, not only that distinct areas of the hand sense frequency with variate intensity, but also that it's possible to create a single actuator tactile display that will stimulate the human hand in multiple areas.

References

[1] D. M. Birnbaum and M. M. Wanderley, "A systematic approach to musical vibrotactile feedback," 2007.

[2] A. J. Brisben, S. S. Hsiao, and K. O. Johnson, "Detection of vibration transmitted through an object grasped in the hand," *Journal of Neurophysiology*, vol. 81, 1999.

[3] C. Chafe, "Tactile audio feedback," in *Opening a New Horizon: Proceedings of the 1993 International Computer Music Conference, ICMC 1993, Tokio, Japan, September 10-15, 1993*. Michigan Publishing, 1993. [Online]. Available: http://hdl.handle.net/2027/spo.bbp2372.1993.013

[4] C. Hopkins, S. Maté-Cid, R. Fulford, G. Seiffert, and J. Ginsborg, "Vibrotactile presentation of musical notes to the glabrous skin for adults with normal hearing or a hearing impairment: Thresholds, dynamic range and high-frequency perception," *PLoS ONE*, vol. 11, 2016.

[5] L. A. Jones and N. B. Sarter, "Tactile displays: Guidance for their design and application," *Human Factors*, vol. 50, 2008.

[6] K.-U. Kyung, M. Ahn, D.-S. Kwon, and M. A. Srinivasan, "Perceptual and biomechanical frequency response of human skin implication for design of tactile displays," in *Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, ser. WHC '05. USA: IEEE Computer Society, 2005, p. 96–101. [Online]. Available: https://doi.org/10.1109/WHC.2005.105

[7] D. A. Mahns, N. M. Perkins, V. Sahai, L. Robinson, and M. J. Rowe, "Vibrotactile frequency discrimination in human hairy skin," *Journal of Neurophysiology*, vol. 95, 2006.

[8] F. Mancini, A. Bauleo, J. Cole, F. Lui, C. Porro, P. Haggard, and G. Iannetti, "Whole-body mapping of spatial acuity for pain and touch," *Annals of Neurology*, vol. 75, 06 2014.

[9] S. Merchel and M. E. Altinsoy, "Psychophysical comparison of the auditory and tactile perception: a survey," *Journal on Multimodal User Interfaces*, vol. 14, 2020.

[10] J. W. Morley and M. J. Rowe, "Perceived pitch of vibrotactile stimuli: effects of vibration amplitude, and implications for vibration frequency coding." *The Journal of Physiology*, vol. 431, 1990.

[11] G. H. Mowbray and J. W. Gebhard, "Sensitivity of the skin to changes in rate of intermittent mechanical stimuli," *Science*, vol. 125, 1957.

[12] J. Rouger, S. Lagleyre, B. Fraysse, S. Deneve, O. Deguine, and P. Barone, "Evidence that cochlear-implanted deaf patients are better multisensory integrators," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, 2007.

[13] B. Stein, P. Stein, and M. Meredith, *The Merging of the Senses*, ser. A Bradford book. MIT Press, 1993. [Online]. Available: https://books.google.se/books?id=uCV9QgAACAAJ

[14] R. T. Verrillo, "Vibration sensation in humans," *Music Perception*, vol. 9, 1992.

[15] A. Wilska, "On the vibrational sensitivity in different regions of the body surface," *Acta Physiologica Scandinavica*, vol. 31, 1954.

Paper C

A Comparison of Audio-to-Tactile Conversion Algorithms for Melody Recognition

R. Paisa, J. Andersen, N. C. Nilsson and S. Serafin

The paper has been published in
EUROREGIO / Baltic Nordic-Acoustic Meetings Conference Proceedings, 2021.

The layout has been revised.

Abstract

Besides language, music is one of the two major acoustic channels for expression of human nature and it is ubiquitous to all cultures. Due to a tight correlation between auditory and haptic stimuli, more and more attention is focused on the importance of the latter sensation in a musical context [21]. For the hearing impaired especially, tactile feedback has been investigated extensively for it's musical applications and hearing assistive devices, as early as 1983 [3]. This study compares three common audio-to-haptic signal processing algorithms designed for full range vibrotactile transducers used for tactile augmentation of music. The focus is on melody discrimination over three instruments: double bass, digital subtractive synthesizer with a sawtooth oscillator and trumpet. The transducer used is a high fidelity Tactuator BM1C, enclosed in a custom anatomical handheld case, inspired by an orthopedic resting hand splint. An evaluation was conducted on 34 participants and used a within-group design with three alternative forced choice task assessing the participants ability to match melodies to tactile stimuli. The results indicate that no algorithm performs better than others, which is in line with the literature regarding the overall poor frequency discrimination of the skin. Nevertheless, post experiment interviews suggest that some participants perceived multiple frequencies simultaneously, on different areas of their hand, similar to auditory polyphony.

C.1 Introduction

In recent years haptic feedback has received increasing attention from the sound and music community, mainly because of the strong connection between the auditory and haptic experiences. This has given birth to *musical haptics* field of research [21]. The mechanism linking auditory and tactile sensations is called multisensory integration, pioneered by Barry Stein and Alex Meredith. It describes how humans form coherent, valid, and robust perception of reality, by processing sensory stimuli from various modalities [24]. The classical rules for multisensory integration demand that enhancement occurs only for stimuli that are temporally coincident and propose that enhancement is strongest for those stimuli that individually are least effective [24]. For this integration to occur, the input from various sensors must eventually converge on the same neurons. In the specific case of auditory-somatosensory stimuli, recent studies demonstrate that multisensory integration can in fact occur at very early stages of cognition, resulting in supra-additive integration of touch and hearing [9, 12]. This translates to a robust synergy between the two sensory apparatuses, than can be exploited to synthesize experiences impossible to achieve by unisensory means. Furthermore, research within auditory-tactile interactions has shown that tactile stimulus can influence auditory stimulus and vice-versa [19, 20, 22]. It can therefore be observed that auditory and haptic stimuli are capable of modifying or altering the perception of each other when presented in unison [29].

This study is the first in a project that has as long term goal to help partially impaired hearing individuals and cochlear implant users to appreciate music. With that in mind, the aim of this particular study is to compare three signal processing methods that convert full spectrum music into vibrotactile haptic feedback suitable for the proprieties of skin receptors, namely the Meissner's corpuscles and Pacinian ones, while preserving the melodic information encoded in the original signal. The three processing methods were chosen from existing literature [7, 8, 15, 23, 27, 30]. The experiment revolved around a handheld device designed to be comfortable to hold for longer periods of time, and capable of reproducing full spectrum audio signal. The hand was identified as the most sensitive body region for touch, due to a very high density of receptors [17, 28].

This paper describes the device built for the study followed by a detailed presentation of each signal processing technique used to convert music to vibrotactile stimuli. Subsequently, it is presented the experimental study evaluating user performance when tasked to match the haptic stimuli to a coherent auditory one. The aim of the study was twofold: (1) to evaluate the three signal processing methods in terms of their ability to convey the melodic structure existing in the original signal.(2) To evaluate the proposed hardware in terms of it's ergonomics and ease of use, as well as it's ability to produce a satisfying haptic experience. Specifically it was considered relevant to determine if a satisfactory experience can be elicited with a single, high-fidelity actuator.

C.2 Background

Live concerts, especially amplified ones, as well as movie scores are know to create haptic sensations coupled with the sound, conveying valuable information such as articulation and timing. Several studies have tried to replicate and quantify this phenomena with compelling results [10, 11, 17, 18].

Merchel approached the topic from an architectural acoustics point of view, aiming to prove that concert halls with a strong haptic feedback improve the overall quality of the concert experience [18]. His studies proposes several signal processing techniques to be used for the haptic channel, indicating that in music with a rich low end, the audio signal passed through a low pass filter is enough to improve the experience. Furthermore, he suggests that simple sinusoidals with frequencies not related to the audio signal will produce an enhanced listening experience, but the frequency of these haptics oscillators will have an impact in the overall perception [18].

Other group of authors suggested to account for haptic feedback at the composition stage, creating a coherent audio-haptic experience, instead of approaching haptics as an afterthought [1, 11]. Gunther and O'Modhrain coined the term *tactile composition* as a *system that facilitates the composition and permeation of intricate, musically structured spatio-temporal patterns on the surface* *of the body*, emphasizing the importance of a compositional language for the sense of touch [11]. Their 2001 *Concerts for the skin* experiments surface some important notions like selective haptic attention - the ability to selectively direct attention into different stimuli, if several body areas are actuated at the same time [11]. On top of that they suggest that the music-haptic relationship does not need to produce congruent stimuli at all times, and the composer should engage into a parallel multimodal composition that inter-plays between the two sensory channels.

Listening for pitch is almost always dependant on the frequency of the audio content, while the timbre and amplitude rarely have an impact on pitch perception [2]. In contrast, the perception of frequency from a vibrotactile stimuli is more complicated due to the multi-channel nature of the cutaneous sensing organ - the skin [2]. Moreover, perception of tactile frequency is amplitude and time dependant, and it varies significantly depending on the position on the body. Nevertheless, there is one important similarity between auditory and tactile pitch perception: within certain frequencies, the discrimination fits a critical band model [16]. Specifically, certain frequency ranges are perceived as distinct sensations, indicating that with enough exposure, tactile pitch perception can be interpreted similarly to the auditory one - a fact proven by many hearing impaired people [4]. This does not mean that understanding music through vibrotactile stimuli is equivalent to hearing it, but the experience, while different, could be just as enjoyable.

Music usually uses a wider frequency spectrum than the skin can provide, and the tactile pitch-amplitude coupling only makes understanding music without hearing it more complicated. Unlike the ear, with its single receptor capable of 20Hz-20000Hz frequency range perception, the skin has multiple types of receptors, each with its own frequency and temporal characteristics. For music perception, two of them prove to be useful: the Meissner corpuscles and the Pacinian receptors [14]. The Meissner corpuscles, also known as Rapid Adapting (RA) receptors, have a very high innervation density and have a limited frequency range of 10Hz-100Hz, with a peak sensitivity around 40Hz. The Pacinian receptors are larger than the RA ones, have a low spatial resolution, and a frequency response between 40Hz and 1000Hz, and are most sensitive around 250Hz [14]. In an attempt to describe the tactile music properties, Erp & Spapé conducted an experiment on the perceptual attributes of vibrotactile melodies [6]. Their results indicate that users can perceive and evaluate multiple characteristics from the tactile stimuli(f.ex. aggressive, soft, alarming, bombastic, etc) and that melodies generally land in one of four clusters, on a two dimensional tempo-intrusiveness map [6]. In a similar fashion, Ternes & MacLean designed a large set of distinguishable tactile rhythms, further highlighting the potential of tactile melodies [25].

C.3 Implementation

C.3.1 Hardware

The hardware device is an ovoid shape with the following dimensions: 84mm wide, 58mm tall and 89mm deep and can be seen in figure C.1.

Fig. C.1: Side and front view of the haptic device

The shape was inspired by the resting hand position when fixed with an orthopedic splint. This pose should minimize the strain on the wrist, and allow the fingers to relax in their natural rest position. The initial shape was created using modelling clay, aiming to ensure the finger position is anatomic, each digit having its own socket. The clay artefact was 3d scanned using Autodesk ReCap[1], by analyzing 40 still images of the subject, taken from multiple angles with a Fujifilm X-T1 camera and a Fujinon XF 35mm @ f2.0 lens. The artefact was suspended in midair with fishing line, affording visibility from all angles. The 3D scan resulted in a very high fidelity digital model, but in order to improve topology, a new 3d model was created using the scan as an outline. The final shape was split in half horizontally, to have access inside where the electronics would eventually lie. The two halves were held together by three M3x16 bolts that have been incorporated in the design show in figure C.1.

For the actuator, a socked was created on the bottom half of the device, and a 3.5mm female jack opening has been installed on top half to connect the transducer to the amplifier. The jack was oriented towards the left side of the device, allowing for cable connection that should not interfere with the user while holding it. The haptic device halves were fabricated with 2mm wall thickness using an Ultimaker 3 and PLA material.

The device was for left hand only, as it was intended to have the users navigate the experiment's questionnaire with the computer mouse, which is generally used with the right hand. Initial informal tests showed that people unfamiliar with the device tend to hold it in unintended ways, thus for the

---

[1] https://www.autodesk.com/products/recap/

experiment finger positioning visual signifies have been painted on. When it comes to the transducer, a Tactuator BM1C vibrotactile actuator manufactured by Tactile Labs[2] was used. Haptuator Mk1 and Mk2 were also tried, but the Tactuator BM1C proved to have the highest amplitude in the current setup, and the distortion (if any) was non disruptive, as would be the case with Mk2 and M1 that rattle rather loud when overdriven. All transducers tested offer full spectrum reproduction. The tactuator requires amplification to achieve desirable amplitude therefore a high gain Behringer HA8000 headphone mixing and distribution amplifier was used.

C.3.2 Tactile signal processing

In an attempt to improve the perception of pitch through tactile sensing, three signal processing methods that convert arbitrary auditory signal into a tactile one were compared. Each of the processing methods was inspired from existing literature, and was re-implemented to exploited one physical or perceptual trait relevant for music listening.
Method 1: Compression of frequency spectrum
The first method focused on compressing the musically relevant frequency spectrum defined between 40Hz and 2093 Hz into a narrower ”tactile range” one up to 1046 Hz, to address the Pacinian receptors exclusively, since these are the most sensitive to vibrotactile stimuli [2, 14]. The lower limit represents the crossover between RA receptors and Pacinian receptors, and the high frequency represents the top range of the Pacinian ones. The frequency compression was implemented as seen in Figure C.2 as following:

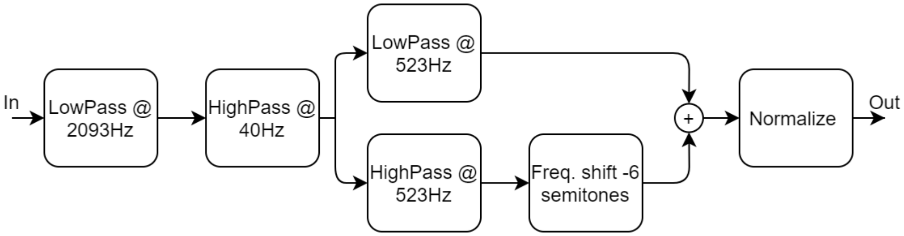


Fig. C.2: Signal processing for first condition

1. Apply a lowpass FIR filter with 60 dB/octave attenuation at 2093Hz - the corresponding frequency of the fundamental for a C7 note with A4 = 440 Hz tuning. This meant that only the highest octave available on a piano was ignored. Nevertheless, that the majority of instruments, including the human voice, have the high limit considerably lower than C7, Since the upper harmonics are not contributing much to melody

---

[2] www.tactilelabs.com

perception, the frequency limitation was not consider to be a practical problem [5].

2. Apply a highpass FIR filter with 60 dB/octave attenuation filter at 40Hz to limit the actuation of RA receptors.
3. Split the frequency band at 523 Hz (C5 note) in two spectra, using a lowpass and a high pass filter. The lower one (40Hz - 523Hz) will be called Spectrum A, and the higher on B. The C5 note was chosen in relationship to the fundamental frequencies of the melodies used, and described in C.3.3.
4. Pitch shift down spectrum B 6 semitones to shift the high frequency content into the tactile sensible range
5. Add the Spectrum A (original) and Spectrum B (pitch shifted)
6. Normalize to 1 to avoid clipping, and ensure equal amplitude throughout the melodies selection.

Method 2: Sinusoidal oscillators
The second method focused on Pacinian receptors as well, and it used sinusoidal with the frequency equal to the fundamental one of the actual tone, instead of the original signal as suggested by Merchel [8, 18, 23, 27]. This was done with the aim of avoiding higher frequency content from masking or diminishing the fundamental harmonic perception, since tactile spectral masking works similar to auditory one [16]. The sinusoidals were generated using the same MIDI information as the auditory signal, using Xfer Serum[3] wavetable synthesizer with *Basic Shapes* table, on position one and a square envelope(0 attack, max sustain, 0 decay). In order to ensure amplitude coherence between the auditory and tactile stimuli, the contour/envelope was extracted from the original file and applied to the haptic one. The last step was to apply normalization, similar to method 1.

Method 3: Tactile transient reinforcement
The last tactile signal processing tried to make use of both the RA and Pacinian receptors. The tactile signal combined the auditory stimuli with a haptic reinforcement one, aimed at the RA receptors in order to emphasize changes in pitch, practically working as an exciter or *transient emphasizer*. This feedback approach was inspired by the way frets provide guitar players feedback about the note selection, as described it [13]. The haptic signal was created by adding a haptic reinforcement component, to the signal generated with method 2. The haptic reinforcement was generated similarly to the sinusoidal described above, but 3 octaves lower than the auditory signal. This meant that the frequencies lied in the peak frequency response of the RA

---

[3] https://xferrecords.com/

receptors [14]. An attack-decay (AD) envelope was used for the haptic reinforcement signal, with 10ms attack time, in order to reduce artefacts(clicks), and 500ms logarithmic decay time to avoid temporal masking over the higher frequency signal. The two signals were summed with amplitudes of 0.8 for the haptic reinforcement, and 0.2 for the original, unprocessed signal, followed by normalization. The large difference in volume between the two signals is due to the lower amplitude response of the Tactuator BM1C below 40Hz.

All processing was done in Matlab unless specified otherwise. Highpass and lowpass filters had a steepness of 0.8 (default in Matlab). The amplitude contour was computed as the moving RMS envelope of the unprocessed melody every 5000 samples, in order to avoid artefacts introduced by abrupt changes in loudness. Sampling frequency used was 48kHz, and all files were exported uncompressed (wav).

C.3.3 Melodies

The 75 musical melodies composed for the project, were all of a duration of three to eight seconds and spread across a randomized selection of different major- and minor keys; figure C.3 shows the distribution of notes across all melodic phrases. The melodies were simple and kept in a melodic style easily recognizable for listeners familiar with western music. They all represented a small musical progression with a beginning and an end. Tempo was 120bpm and, rhythmically, there were a mixture of whole- quarter- eighth and sixteenth- notes. For each of the 75 true melodies, two false were added. The false melodies always had the same rhythmic content as the true one, but at least 75 percent of the tones were changed. In the false melodies, the musical progression would not be perceived as natural, since the selection of notes were not following western melodic tradition. Figure C.4 shows the average number of semitones deviating from the correct melodies.

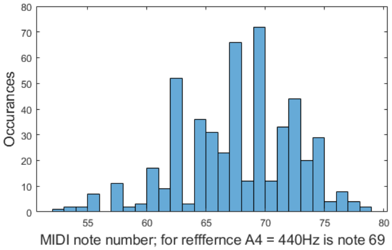


Fig. C.3: Distribution of notes in the correct melodies

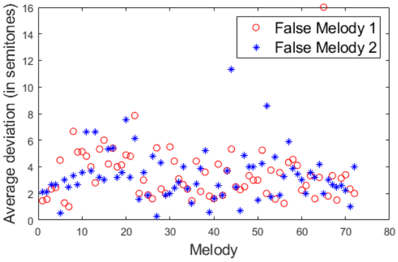


Fig. C.4: False melodies deviation from the correct ones

C.4 Evaluation

The aim of the study was to (1) evaluate three signal processing techniques for converting audio material to tactile stimuli, and (2) to evaluate the usability of the haptic device itself. To meet this aim a within-subjects study was performed, comparing four conditions that varied in term of the tactile feedback provided through the haptic device, when listening to melodic phrases. **The hypothesis was:** *There is a difference in terms of tactile melody discrimination between an unprocessed signal and a processed one when presenting congruent bi-modal melodic phrases through a single-actuator handheld vibrotactile device and headphones.*

C.4.1 Task and Stimuli

The task for the participant was to select the haptic feedback that matched the auditory signal played through the headphones. A three alternative forced choice design was used, with only one correct option; for each trial the participants were presented with three types of haptic feedback. The experiment had four conditions, with different singal processing techniques described in section C.3: [1]Control condition with no processing, [2]Frequency compression, [3]Sine wave at the fundamental frequency and [4]Tactile reinforcement of transients.

There were 72 trials in total: 18 for each processing technique, plus 18 for unprocessed acting as control condition. Melodies where chosen and presented randomly out of pool of 75 possibilities, distributed equally among the three instruments. A total of 900 possible trials were used for the whole experiment: 75 melodic lines * 4 conditions * 3 instruments, ensuring a high level of validity. For each instrument, condition and melodic line there was one correct haptic stimulus, and two incorrect ones. The order for melodies, order of conditions and choice of instruments were assigned randomly, in real time, for each participant. In order to ensure similar exposure levels for all participants, they were allowed to experience each stimuli/melody combination only once. The experiment took place in the Multisensory Experience Lab at Aalborg University campus in Copenhagen.

C.4.2 Participants

Participation in the experiment was voluntary and the majority of participants were students of *Sound and Music Computing* and *Medialogy* programs, that are affiliated with the Multisensory Experience Lab, and the Tonmeister program from the Royal Danish Academy of Music. Some participants had musical experience, but this was not a selection criteria. The participants have been encouraged to partake in the experiment at their convenient time and there was no reward for doing it. The data collected has been anonymous, without the possibility of matching answers sets with the participant. There were a total of 34 participants (24 male, 10 female). Although the ultimate goal is to create a device for hearing impaired users, current COVID-19 restrictions prevented us to test on the relevant target group. The experiment was conducted with 3 participants in parallel, in 3 different rooms, that were briefed and debriefed together, by the first author.

C.4.3 Setup and Equipment

The hardware setup consisted of a Windows computer running the experiment application with a Behringer HA8000 headphone distribution amplifier connected to it. The left audio channel contained the haptic melodies, and the right channel carried the auditory signal. The distribution amplifier routed the auditory signal to both headphones channels. The headphones used were different due to availability, but had similar price and quality level: *Creative Aurvana Live!, AKG K240 and Sennheiser HD240 Pro*. The level balance between the haptic and auditory signal was set by the second author, and calibrated to have a natural balance between the two sensory inputs to assure that not one would overpower the other. First the headphones were adjusted to a comfortable playback-level and then the haptics were added up close to the distortion limit of the transducer. The mix was constant for all participants.

C.4.4 Procedure

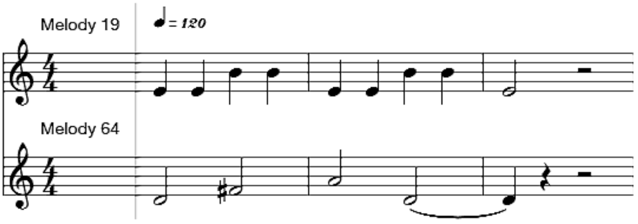
As mentioned, three participants partook in the experiment at the same time. They were welcomed and introduced to the task, emphasizing that it is not required from them to over-analyze the vibrotactile stimuli, but instead they should answer based on their intuition. The participants were then guided to the setup rooms and required to experience Queen's "Don't stop me now", as training and accommodation with the system. The song was chosen due to its popularly, but also because it features many combinations of instruments and intensities: from low intensity voice only, to high intensity full band playing. This should provide the users with most of the possible stimuli expected throughout the experiment. After the accommodation phase was finished, the users were required to click on "Start" button to begin the experiment. Each trial consisted of listening to the same auditory melody three times, with different haptic stimuli for each as described in C.3.2. A visual indicator was signaling what exposure was playing at all times, and the "Select the haptic stimuli that matches best the melodic phrase you heard" message was permanently displayed in the bottom of the page, followed by the trial number. There was a 2 seconds gap between exposures within the same trial. Since all potential melodies were fairly similar in length, the experiment was completed in 18-19 minutes. After all 3 participants in a series finished, they were gathered for a post-experiment debriefing discussing about their experience, comfort, amplitude and potential suggestions. The setup, similar to an ad-hoc focus group, facilitated interesting discussions between participants,

h!

Fig. C.5: Best (19) and worst (64) performing melodies

as well as between conductors and participants.

C.4.5 Data Collection

The data collected has been anonymous, without the possibility of matching answers sets with the participant. The following information was logged: trial number(1-72), melody number(1-72), condition(1-4), instrument(1-3), correct answer, user answer, and inevitably, the log file creation time.

C.5 Results

The data collected from the 34 participants was treated as nominal and was analyzed using Friedman tests. The main test was run to determine if there were significant differences between the three proposed signal processing techniques with respect to number of correct identification of the matching haptic stimuli. In addition to this analysis, the data was analysed on a per-instrument basis, as well as instrument performances against each-other. The number of correct answers were not statistically significantly different among conditions $\chi^2(3) = 0.885, p = .829$. Similarly, no significant differences were found when the instruments were analyzed independently: bass trials ($\chi^2(3) = 2.590, p = .459$), synth trials ($\chi^2(3) = 4.528, p = .210$), and trumpet trials ($\chi^2(3) = 1.401, p = .701$).

No effect on the number of correct answer was observed, but while inspecting the descriptive statistics, it seemed that one instrument (synthesizer) stands out therefore an exploratory analysis was ran, comparing the sets of trials for each instrument against each other. The results are $\chi^2(2) = 1.746, p = .418$. Furthermore, when looking at the best and worst performing three melodies in terms of correct answers, it was discovered that the ones with multiple short notes had a slightly higher average number of correct answers, while the ones with longer, sustained notes had a lower number of correct answers, even when harmonic content is very similar. The correlation between average note length in the melody and number of correct answers is ρ = -0.19 with p = 0.1. Figure C.5 show the best performing melody with 14 correct answers out of 34 (41.1%) and worst performing one with 2 correct answers (5.8%). Lastly, there was no correlation found between the average deviation from correct melody as show in C.4 and the number of correct notes: $\rho < 0.07$.

C.5.1 Post-experiment interview

The post experiment interview highlighted some interesting facts regarding the physical design, the experiment as well as potential directions for further experimentation. Several participants remarked that the experiment is too long and repetitive, loosing focus towards the end. Regarding the physical properties, some subjects reported that they experimented with different arm resting positions (arm resting on the knee facing up/down, arm resting on the table, crossed arms) noticing that each position will produce slightly different results. Out of those who mentioned position, there seemed to be a consensus that palm facing up feels the best, with one mention that it felt stronger. Probably the most interesting feedback was that some participants felt different frequencies in different areas of the hand, one participant mentioning that sometimes it could sense two frequencies at the same time. This has been expressed in various forms, some claiming that higher frequencies feel too strong, especially for the fingers, but the lower frequencies feel good.

Regarding the hardware, the feedback has been generally good, but some participants suggested that the device was either too big or too small for their hands. At the same time, few participants reported that it is a slightly uncomfortable to hold for long time while most mentioned it was comfortable.

C.6 Discussion

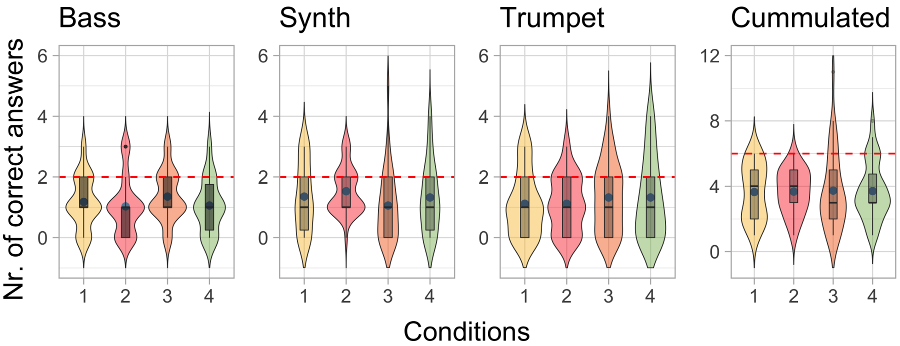


Fig. C.6: Distribution of number correct answers (max. possible = 6/instrument/condition),for each instrument, for each of the four conditions: 1 = control condition; 2 = frequency compression; 3 = sine wave at the fundamental frequency; 4 = haptic reinforcing of transients. Box plots inside the violin plots represent interquartile range, blue circles represent the mean and red line represents the expected chance level.

The results related to *number of correct answers* suggest that the proposed processing techniques do not result in statistically significant different performances. Furthermore, looking at the nature of the haptic stimuli, it is observed that the *bass, synthesizer and trumpet* perform similarly among the 4 conditions. Even though the performed tests does not permit us to conclude that the results are statistically equivalent, the descriptive statistics does seem to indicate that the effect of processing methods was negligible, and possibly non-existing. This can be seen in figure C.6 showing the median for all conditions is 1 and the variance is consistent, regardless of the instrument presented. A potential explanation for this similarity can be found in the fact that a higher number of shorter notes are easier to recognize, opposite to longer, sustain ones as seen in Figure C.5. A similar behavior was observed by Tommerdahl et. al in 2005 in his studies on the vibrotactile discrimination capacity of skin for various stimuli lengths, concluding that the cerebral cortex undergoes a profound inhibition withing 1-2 seconds from the start of a 200Hz stimuli [26]. That being said, this study did not investigate the impact of legato or staccato as all harmonic events were rhythmically independent, therefore no conclusion can be reached on whether the rapid melodic changes or the short duration explain this phenomena. Nonetheless, if condition 1 - control and condition 3 - sinusoidal oscillators, are analysed in isolation, the results do not align with the findings of Merchel, that suggest using sine signals with the frequency matching the fundamental one from the auditory signal produces a better tactile experience - at least in terms of melodic content identification [18].

When it comes to the comfort and performance of the physical device, the results are mixed. Some users claimed it was comfortable and provided appropriately strong vibrotactile stimuli, while other complained that it can be too strong at times or that it becomes uncomfortable to use for longer periods of time. These findings indicate a preferences for individual customization of device size as well as control over the haptic intensity.

Lastly, an interesting phenomena was describe by several users, claiming that different frequencies are felt in different areas of the hand. This is a direction worth exploring further, since it can indicate that single actuator devices can address different areas of the hand, providing an extra dimension for communication.

C.7 Conclusion

In this paper it was proposed a system that allows musical signals to be converted to vibrotactile stimuli. The system was evaluated in a user experiment exploring impact of three signal processing techniques used for audio to haptic conversion, in terms of user's ability to identify the melodic information. The stimuli used for the experiment was composed of short melodies played on double bass, synthesizer and trumpet. The results indicated that there was no significant difference between the 3 proposed techniques and no processing at all, when it comes to melody identification, underlining the need for new algorithms that can be empirically validated. However, there was an indication that users do perform better at the identification task when the haptic stimuli contains shorter notes, regardless of processing algorithm or instrument played. Finally, it was surfaced that different areas of the hand can sense separate frequencies, but further research needs to be conducted in order to fully understand the phenomena.

References

[1] J. Armitage and K. Ng, "Feeling sound: Exploring a haptic-audio relationship," pp. 146–152, 2016.

[2] D. M. Birnbaum and M. M. Wanderley, "A systematic approach to musical vibrotactile feedback," 2007.

[3] P. L. Brooks and B. J. Frost, "Evaluation of a tactile vocoder for word recognition," *The Journal of the Acoustical Society of America*, vol. 74, 1983.

[4] L. Bruns, D. Mürbe, and A. Hahne, "Understanding music with cochlear implants," *Scientific Reports*, vol. 6, 2016.

[5] J. M. Eargle, *Frequency Ranges of Musical Instruments and the Human Voice*. Boston, MA: Springer US, 2002, pp. 324–325. [Online]. Available: https://doi.org/10.1007/978-1-4615-2027-6_156

[6] J. V. Erp and M. Spapé, "Distilling the underlying dimensions of tactile melodies," *Proceedings of Eurohaptics 2003*, 2003.

[7] M. D. Fletcher, N. Thini, and S. W. Perry, "Enhanced pitch discrimination for cochlear implant users with a new haptic neuroprosthetic," *Scientific Reports*, vol. 10, 2020.

[8] M. Fletcher, S. Mills, and T. Goehring, "Vibro-tactile enhancement of speech intelligibility in multi-talker noise for simulated cochlear implant listening," *Trends in Hearing*, vol. 22, 2018.

[9] J. Foxe, G. Wylie, A. Martinez, C. Schroeder, D. Javitt, D. Guilfoyle, W. Ritter, and M. Murray, "Auditory-somatosensory multisensory processing in auditory association cortex: An fmri study," *Journal of neurophysiology*, vol. 88, pp. 540–3, 08 2002.

[10] M. Giordano, J. Sullivan, and M. M. Wanderley, *Design of Vibrotactile Feedback and Stimulation for Music Performance*, 2018.

[11] E. Gunther and S. O'Modhrain, "Cutaneous grooves: Composing for the sense of touch," *International Journal of Phytoremediation*, vol. 21, 2003.

[12] C. Kayser, C. I. Petkov, M. Augath, and N. K. Logothetis, "Integration of touch and sound in auditory cortex," *Neuron*, vol. 48, 2005.

[13] H. Knutzen, T. Kvifte, and M. M. Wanderley, "Vibrotactile feedback for an open air music controller," vol. 8905, 2014.

[14] K.-U. Kyung, M. Ahn, D.-S. Kwon, and M. A. Srinivasan, "Perceptual and biomechanical frequency response of human skin implication for design of tactile displays," in *Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, ser. WHC '05. USA: IEEE Computer Society, 2005, p. 96–101. [Online]. Available: https://doi.org/10.1109/WHC.2005.105

[15] M. J. Lucía, P. Revuelta, Á. García, B. Ruiz, R. Vergaz, V. Cerdán, and T. Ortiz, "Vibrotactile captioning of musical effects in audio-visual media as an alternative for deaf and hard of hearing people: An eeg study," *IEEE Access*, vol. 8, pp. 190 873–190 881, 2020.

[16] J. Makous, R. Friedman, and C. Vierck, "A critical band filter in touch," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 15, no. 4, p. 2808—2818, April 1995.

[17] A. Mazzoni and N. Bryan-Kinns, "How does it feel like? an exploratory study of a prototype system to convey emotion through haptic wearable devices," 2015.

[18] S. Merchel and M. E. Altinsoy, "Auditory-tactile music perception," vol. 19, 2013.

[19] R. Okazaki, T. Hachisu, M. Sato, S. Fukushima, V. Hayward, and H. Kajimoto, "Judged consonance of tactile and auditory frequencies," 04 2013, pp. 663–666.

[20] R. Okazaki, H. Kajimoto, and V. Hayward, "Vibrotactile stimulation can affect auditory loudness: A pilot study," 06 2012, pp. 103–108.

[21] S. Papetti and C. Saitis, *Musical Haptics: Introduction*. Springer, Cham, 2018.

[22] T. Ro, J. Hsu, N. Yasar, C. Elmore, and M. Beauchamp, "Sound enhances touch perception," *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, vol. 195, pp. 135–43, 04 2009.

[23] S. Shin, C. Oh, and H. Shin, "Tactile tone system: A wearable device to assist accuracy of vocal pitch in cochlear implant users," 2020.

[24] B. Stein, P. Stein, and M. Meredith, *The Merging of the Senses*, ser. A Bradford book. MIT Press, 1993. [Online]. Available: https://books.google.se/books?id=uCV9QgAACAAJ

[25] D. Ternes and K. E. MacLean, "Designing large sets of haptic icons with rhythm," vol. 5024 LNCS, 2008.

[26] M. Tommerdahl, K. Hester, E. Felix, M. Hollins, O. Favorov, P. Quibrera, and B. Whitsel, "Human vibrotactile frequency discriminative capacity after adaptation to 25 hz or 200 hz stimulation," *Brain research*, vol. 1057, no. 1-2, p. 1—9, September 2005. [Online]. Available: https://doi.org/10.1016/j.brainres.2005.04.031

[27] U. Trivedi, R. Alqasemi, and R. Dubey, "Wearable musical haptic sleeves for people with hearing impairment," *ACM International Conference Proceeding Series*, pp. 146–151, 2019.

[28] A. Wilska, "On the vibrational sensitivity in different regions of the body surface," *Acta Physiologica Scandinavica*, vol. 31, 1954.

[29] G. Young, D. Murphy, and J. Weeter, "Haptics in music: The effects of vibrotactile stimulus in low frequency auditory difference detection tasks," *IEEE Transactions on Haptics*, vol. PP, pp. 1–1, 12 2016.

[30] ——, "Vibrotactile discrimination of pure and complex waveforms," 2015, pp. 359–362.

Paper D

A Real-Time Cochlear Implant Simulator - Design and Evaluation

C. Steinhauer, T. Sønderbo and R. Paisa

The paper has been published in
The 20th Sound and Music Computing Conference Conference Proceedings,
2023.

Abstract

This article describes the implementation of a flexible real-time Cochlear Implant (CI) simulator, and it's preliminary evaluation set to investigate if a specific set of parameters can simulate the musical experience through CIs using Normal Hearing (NH) subjects. A Melodic Contour Identification (MCI) test is performed with 19 NH subjects to identify melodic contours processed by the simulator. The results showed that the participants had a decrease in precision in determining musical contours as the intervals between notes decreased, showing that the reduced spectral resolution increases the difficulty to identify smaller changes in pitch. These results fall in line with other studies that perform MCI tests on subjects with CI, suggesting that the real-time simulator can mimic the reduced spectral resolution of a CI successfully. This study validates that the implemented simulator, using a pulse-spreading harmonic complex as a carrier for a vocoder, can partially resemble the musical experience had by people with hearing loss using CI hearing technology. This suggests that the simulator might be used to further examine the characteristics that could enhance the music listening experience for people using CIs.

D.1 Introduction

A cochlear implant (CI) is a device that restores part of the auditory sensations for those who suffer severe-to-profound hearing loss. The implant stimulates the auditory nerve directly in the cochlea, based on auditory input captured by the external sound processor [29], and it can provide good speech comprehension in quiet environments [11]. However, CI listeners experience poor music perception, both in terms of self-reported music enjoyment and objective perceptual abilities, scoring significantly lower than NH subjects [18, 23, 24]. This stems from a deficit of pitch and timbre perception, which is limited by the spectral resolution of the CI [20, 24], as well as an insufficient providence of input dynamic range to cover the wide amplitude range of music [19]. Proper pitch perception is crucial to understanding the harmonic complex tones produced by musical instruments or when distinguishing sources, for example in situations with multiple talkers [20]. Since music is an important part of social gatherings and mood regulation, the limitations of music perception in CI individuals may affect their well-being and quality of life [17], demanding that further research is needed to improve upon this.

The end results of auditory performance of the CI receiver can be affected by multiple factors such as the choice of device, surgical quality, duration of hearing loss (HL), whether HL occurred pre- or postlingual, the quality of recovery from surgery, rehabilitation practices, and even more [14]. This results in a large variation in the individual experience that applies to the perception and appreciation of music. For instance, studies suggests that early deafened, late implant users appreciates music more than postlingual users, although there seemed to be no significant difference between the two groups in identifying musical contours [7].

This study proposes an implementation of a CI simulator application that processes sound in real-time, and can be adjusted with user-configurable parameters with immediate effect. This flexibility encourages further exploration of various auditory perceptions experienced by CI users and discussion about how the combination of parameters affects the experience. The simulator incorporates three carriers: sine, noise, and pulse-spreading harmonic context (PSHC), the latter also being evaluated in a melodic contour identification (MCI) test. Here, we assessed nineteen NH subjects' ability to perceive musical contours by the simulator, with similar configured parameters to those outlined in Karoui et. al. [16]. In doing so, we wish to draw parallels to an MCI test conducted with CI subjects Galvin et al. [10] and to discuss if the parametric setup can infer anything about the musical perception of CI users.

D.2 Background

To get a better understanding of the different parameters and shortcomings of CI technology, there is an increasing interest in simulating the auditory perception of CI listeners. Channel vocoders are typically used to acoustically simulate CIs, usually with a number of sinusoidal or noise-band carriers that simulate the electrical pulse trains from the implanted electrodes [16]. The vocoder approximates the signal chain found in cochlear implants, as electrodes can be thought as limited bandwidth carriers, coupled directly to the auditory nerve.

A vocoder (voice coder-decoder) analyzes a signal by separating it into a number of channels and then extracts the low frequency envelope of each channel. The vocoder recreates the original signal using the extracted envelopes to modulate either noise or carrier waves for each separate channel and then summarizes these into a final, re-synthesized signal.

Vocoder based simulators are not a novelty, having been used in several studies, these have been mostly evaluated in terms of speech intelligibility spatial localization [2, 4, 26]. In 2006 Poissant et al. [26] started to look at the effects of reverberation on speech in quiet scenario with multiple vocoder configurations, concluding that systems with a lower number of channels exhibit an exaggerated vulnerability to reverberation. Similarly, Whitmal et al. [28] compares two vocoder channels in terms of sentence identification, concluding that noise based vocoders perform worse than tone based ones, mainly due to their intrinsic temporal modulation that can interfere with the temporal fluctuations carrying speech cues. More recently, Jain and Ghosh [15] set to investigate the effect of several simulator parameters on speech quality and intelligibility using both subjective and objective tests.

Another approach to investigate the validity of CI simulators proposed by several studies [6, 25] is to compare the electric and acoustic stimulation in CI subjects with single-sided deafness (SSD). Both these studies found substantial variability in selection of the most similar sounding simulation to the CI ear. A recently introduced pulsatile carrier, pulse spreading harmonic complex (PSHC [13]), aims to mitigate the limitations of sinusoidal and noise carriers for CI simulations outlined by [28]; specifically, that sinusoids cannot simulate the broad spread of excitation produced by a CI electrode, [21], and that noise contains intrinsic modulations that are absent in CIs. PSHC is broadband and intrinsic modulations can be minimized by its pulse rate [13]. When evaluated by SSD-CI subjects using tone, noise and PSHC-based vocoders, it was found that the PSHC ones were judged more similar to the CI, than the other two cases [1, 16].

One thing that all aforementioned simulators have in common is that they do not allow for real-time control of parameters, or some work only with pre-recorded sounds, severely limiting the possibility of exploration and customization. Cavdir et al. [5] mentioned that when working with accessible technology it is important to acknowledge the individual characteristics and engage in an open discussion with the target group, thus, in the case of CI simulators it is important to be able to adjust both input signals and the technical characteristics of the system.

D.3 Materials and Methods

In this study, we developed a CI simulation application which aims to simulate the auditory experience of CI users in real time. It may be configured based on several parameters, allowing us to consider several different perceptions of CI. These parameters include amount of channels, carrier type, frequency range of the simulated implant, individual channel gain and overall compression (threshold and makeup gain). The main layout of the application and parameters are displayed in Figure D.1.

D.3.1 Software

The CI Simulation application was based on *JUCE*[1] as the framework due to its run-time performance benefits, which were required for real-time processing. - and an interactive UI allowing configuration of several parameters in real-time. The MCI test was implemented in Windows Forms (*WinForms .NET*). The MCI test results was analysed in MATLAB and visualized with PowerBI.

---

[1] https://juce.com/

☐ Use media as input

☐ Loop

Stopped

Select number of channels: 6

Frequency range: 250 Hz - 4500 Hz

☐ Sine processor

☐ Noise processor

☑ PSHC processor

C1 1.00 · C2 1.00 · C3 1.00 · C4 1.00 · C5 1.00 · C6 1.00

Fig. D.1: Screenshot of the application in action with a microphone as input source device. The upper section of the layout consists of two displays, visualizing the frequency spectrum of the raw input and processed output; with each channel displayed individually in the output spectrum. The lower section of the layout constitute the user-configurable parameters, and can be modified to adjust the output in real-time. Use **(A)** to select number of channels, **(B)** to enable and disable carrier types, and control the gain of each carrier, **(C)** to control the gain of each individual channel, and **(D)** to adjust the upper and lower cut-offs of the frequency range in which the band-pass filters will be distributed between.

D.3.2 Cochlear Implant Simulation Processor

The processor is based on a vocoder to represent the CI, with each channel corresponding to a subset of the electrodes inserted into the cochlear. The base design of the vocoder is implemented as described by Karoui et al. [16]. The implementation of the CI simulation processor is designed to be configurable, which can be directly interacted using the UI components of the application. As the application executes the simulation in real time, the processor fetches the parameter values from a state that contains the most recent settings based on the configurable components. When the processor receives an *audio block*, it is duplicated into N amount of channels, where the number (N) is user configurable. The channels are then processed through three stages: *preprocessing*, *analysis* and *reconstruction* as shown in Figure D.5. After being processed through these three stages, the output signal is then compressed with a configurable threshold. Furthermore, there exist additional options that allows controlling the gain of each channel and overall output signal.

Preprocessing

In the *preprocessing* stage, the signal is filtered with a bandpass filter to limit the frequency range. The upper and lower frequencies are configurable, but have a default value of 250Hz-4500Hz to match the ones in Karoui et. al. [16] and ensure comparability between studies. In the other study, the limited frequency range served the purpose of being able to control the input stimuli passed to SSD-CI subjects, by low-pass filtering all CI stimuli so that the stimuli delivered to the two ears always had the same bandwidth. Furthermore, the upper limit is just over the highest fundamental frequency available on a piano, thus affording a broad selection of musical notes to be reproduced. An example of an input signal used in the experiment described in D.3.3 can be seen in Figure D.2; it is a simple piano melody that consists of five tones, starting in A4 and increasing in pitch by two semitones each time. The pre-filtered signal in the time domain (A) has approximately even amplitude peaks for all tones, in the frequency domain (B) the fundamental frequency (f_0) of the five tones span approximately between 400Hz and 750Hz. Post-filtered signal has been filtered with a sixth-order Butterworth band-pass described in Figure D.3, which in this case represents a single channel in the vocoder; this will be elaborated upon in D.3.2. As a result, on time (C) and frequency domain (D) the amplitude peaks decrease for each tone, as the frequencies pass beyond the upper cutoff frequency of the filter (500Hz). The envelope (E) is extracted with half-wave rectification followed by a second order Butterworth low-pass filter.

Analysis

In the *Analysis* stage, the temporal information for each channel is processed and extracted. Sixth-order Butterworth band-pass is used to filter each channel which greatly reduce spectral information, but preserves the temporal envelope cues in each band [27].

An example of a sixth-order Butterworth is shown in Figure D.3 with cutoff frequencies at 250Hz and 500Hz. The effect of the filter used on a signal of a simple piano melody can be seen in Figure D.2.

The lower and upper frequencies of the Butterworth filters used in the *analysis* stage are calculated using the Greenwood function [12] within the frequency range used in the *preprocessing* stage (D.3.2). The N^* channels covers their own section of the frequency spectrum with the theoretical width of excitation along the basilar membrane [22] (see Figure D.4). The envelope is extracted with half-wave rectification, followed by a second order Butterworth low-pass filter. Similar to the approach taken by Mesnildrey et al. [21], the cutoff frequency for the low-pass filter is equivalent to the pulse rate of the PSHC carrier divided by two. This will be further elaborated in the Reconstruction stage (D.3.2). The cutoff frequency for the low-pass filter has a upper limit of $200Hz$ for cases wherein half the frequency of the pulse rate is greater than the aforementioned value.

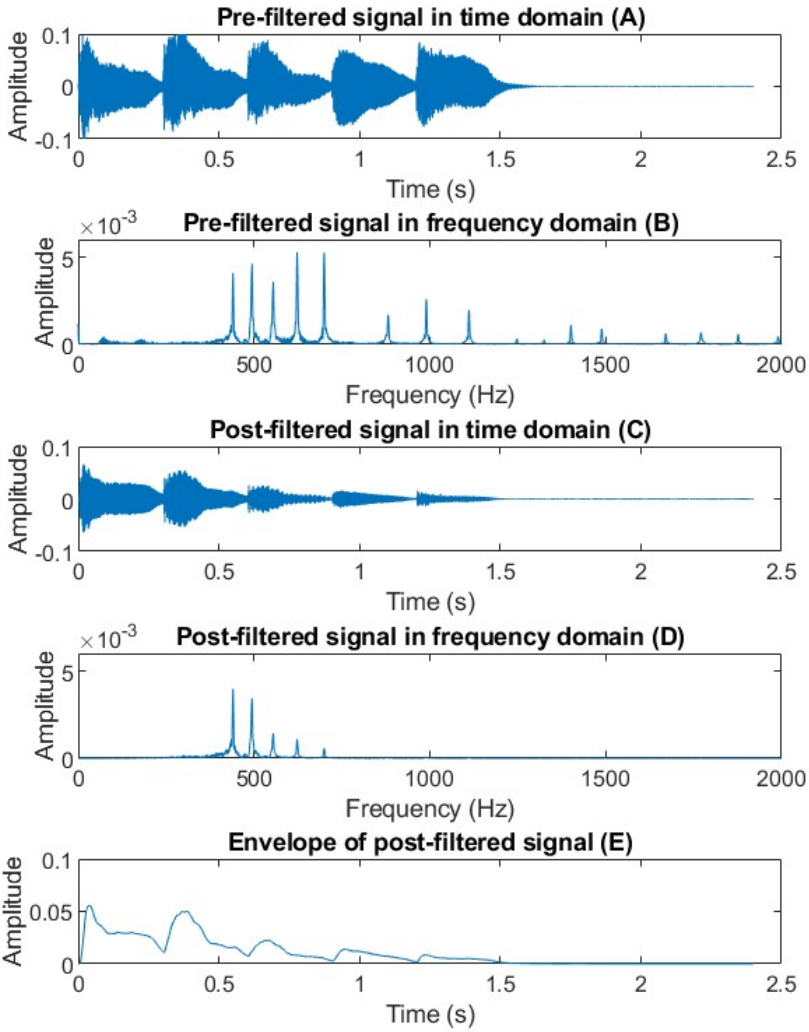


Fig. D.2: Input signal in different pre-processing stages in both time and frequency domains

Reconstruction

In the *Reconstruction* stage, each channel is synthesized to produce acoustic simulation. The envelope of each channel is modulated using one or more

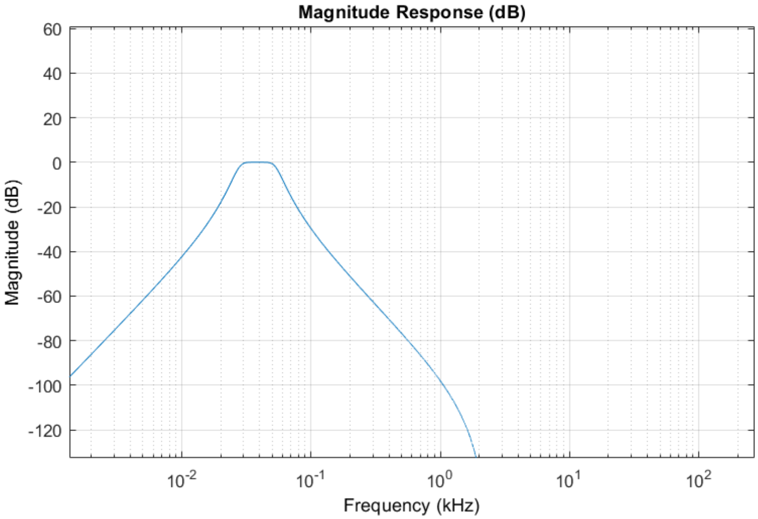


Fig. D.3: Magnitude response of a sixth-order Butterworth bandpass filter with 250 Hz and 500 Hz as lower and upper cutoff frequencies respectively. The Frequency (kHz) axis is logarithmic.

of three carriers: sinusoidal (SINE), Gaussian noise (NOISE) or PSHCs. The carrier modulation resembles the procedure used by Karoui et al. [16]. As such, NOISE carrier generates noise based on Gaussian distribution. SINE carrier generates a sinusoid with the center frequency of the corresponding bandpass. The PSHC carrier is implemented as described in [21]. Therefore, all PSHC carriers have a fundamental frequency of 0.3 Hz and the pulse rate is frequency dependant for each channel, in order to limit intrinsic modulations. The optimal pulse rate is calculated using equation D.1, which is a second order polynomial fit between center frequency and pulse rate derived from *Table 1* in [21]:

$$y = 37 + 151x + 0.17x^2 \tag{D.1}$$

where x is the center frequency and y is the optimal pulse rate.

Finally, all PSHCs were filtered with a gammatone filter as implemented in [3]), which were initialized with the same center frequencies of the band of the analysis. This optimizes the carrier such that their intrinsic modulations after auditory filtering, with the optimal pulse rate, showed smaller internal crest factors than other carriers with equivalent bandwidths [13]. After amplitude modulation, the output signal is filtered once again using the corresponding Butterworth band-pass filters from the Analysis stage. The implementation of the simulator does not include configurable ERB (equivalent rectangular bandwidth) mismatch.The authors in [16] found that participants overall had a preference for no ERB mismatch, as such this configuration was omitted from this simulator. For the experiment described in D.3.3,

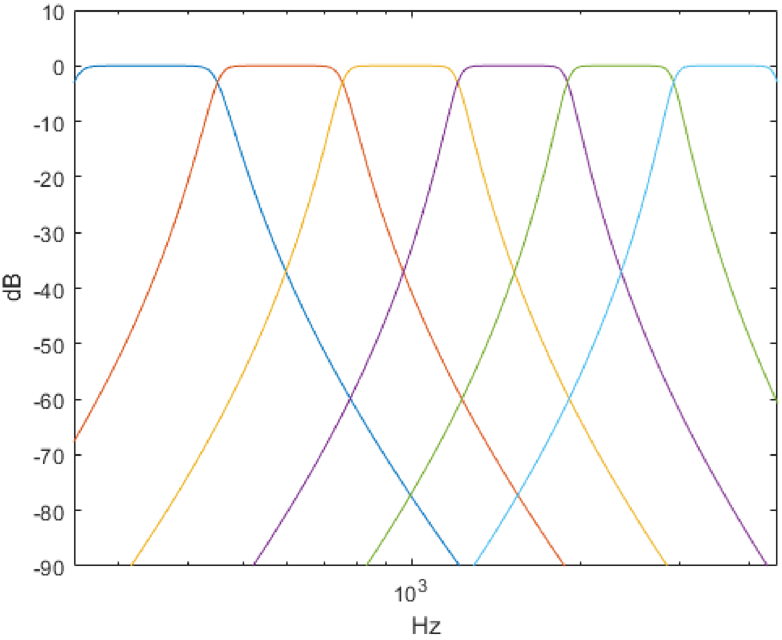


Fig. D.4: Analysis Butterworth bandpass filters for 6 channels spanning from 250-4500Hz, distributed with the greenwood function. The x-axis (Hz) is logarithmic.

only PSHC carries have been used, as this configuration is documented to be closer to CI's as highlighted in [1, 16].

D.3.3 Melodic Contour Identification Evaluation

The MCI test was based on the procedure described by Galvin Et al. [8–10], with the simulator parameters set to match those of the experiment described by [16]. The goal was to investigate whether our simulator can perform similarly to the one validated by bi-modal users, as presented in [16], when listening to music.

Subjects

There were 19 participants in the MCI test. All participants volunteered for the test, and no information regarding participants was recorded. Furthermore, none of the participants had any hearing loss based on self report.

Stimuli

Stimuli consisted of equal amplitude virtual grand piano (Ableton Live 11 Grand Piano collection[2]) triggered by MIDI notes which has been processed using the CI simulator. Piano sounds were used as they produce overtones giving the notes harmonic components. This makes some notes audible even if their fundamental frequency is below cutoff frequency of the band-pass filter in the simulator. The simulator was configured with the parameters described in [16], due to the PSHC vocoder being judged most similar compared to SINE and NOISE vocoders with these configurations. It was processed through six channels spanning from 250Hz-4500Hz and synthesized with PSHC as carrier type. In order to obtain results as comparable as possible with the ones in [16], no compression was added to the stimuli.

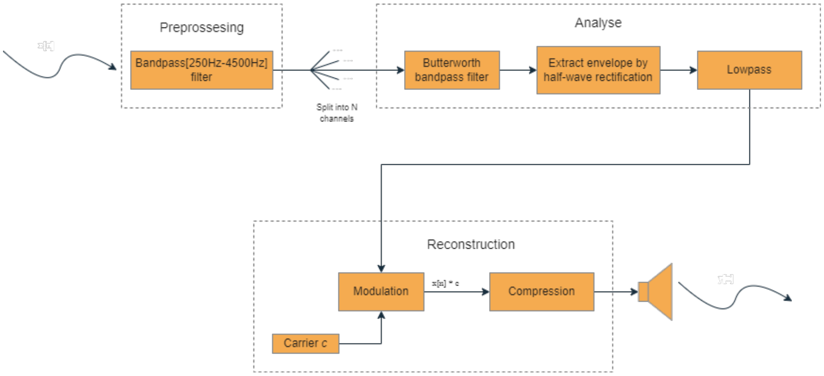


Fig. D.5: Block diagram of the three stages of the CI simulation processor

Procedure

Similar to previous studies investigating contour identification for CI users [8–10], the experiment described here used five notes of equal duration following musical intervals. Each note was played for the duration of 250 ms with 50 ms of rest between notes. Each set of notes used either two, three or five semitones separation between each other, which differs from the procedure by Galvin et al. [10] as their evaluation used all combinations of one to five semitone intervals. This was done to simplify the experiment, as an initial evaluation of the CI simulator and PSHC carrier type. The melody played for a set is either *Rising*, *Falling* or *Flat* as depicted in Figure D.6. The root note was the starting note for the contours in three different octaves (A3,

---

[2] www.ableton.com

A4 and A5). This leads to a total of twenty-one different combinations of melodies. Each combination was repeated three times in a random order for each participant, resulting in a total sixty-three test cases. The participants were not informed about the number of melodies and contours, or the number of repetitions for each case, and they were only allowed to hear each contour only once.

A custom PC application was used to perform the test. The application played each contour based on an input from the participant. The participant sat in front of a laptop computer in an empty classroom, and could select the interpreted contour from the three options presented (Falling, Flat, Rising) as seen in Figure D.6. It took each participant five to ten minutes to complete the test. The MCI test results of all participants were recorded and stored in a database with relevant metadata (contour type, semitone interval, octave and selected contour). The sounds were played to the participant at a comfortable level using a set of Bose QC700 noise-cancelling headphones, with noise cancellation set to the highest level.

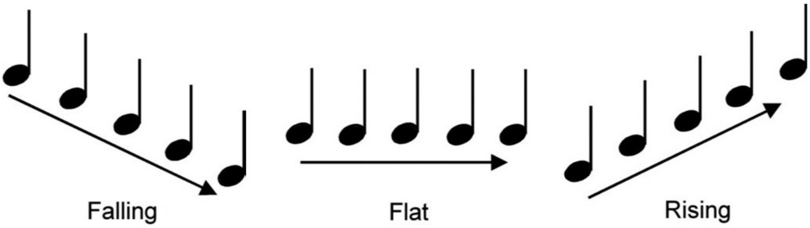


Fig. D.6: Depiction of Falling, Flat, and Rising melodic contours.

In summary, 21 different combinations of contours were tested: *contour*{Falling, Rising} * *octave*{A3, A4, A5} * *semitone*{2, 3, 5} and *contour*{Flat} * *octave*{A3, A4, A5}.

D.4 Results

This section will cover the results of the MCI test. The test included 19 participants for a total of 1197 answers. Of the answers submitted 511 of them were incorrect and 686 of them were correct resulting in an average performance of 57.3% accuracy in determining contours with a standard deviation of 12.8%[3]. A detailed overview of the amount of correct answers per contour

---

[3] A bug was discovered in the program used to test the participants. This caused one contour to be played four times instead of three - and a random other contour to be played two times instead of three. Because of this, a set of seventeen participants had an extra flat contour in their test, replacing a falling contour for a subset of eleven participants, and replacing a rising contour the remaining subset of six participants. Every participant has been exposed to each combination of contour, octave and semitone interval, and has evaluated a total of sixty-three contours.

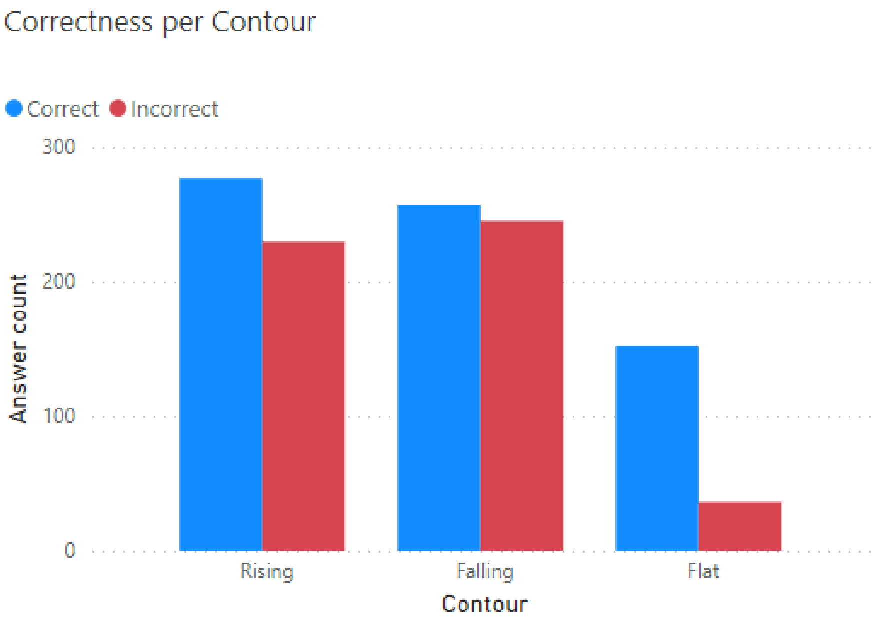


Fig. D.7: Results of the MCI test shown as correct and incorrect answers by contour. The most significant difference in amount of correct answers is the Flat contour with 152 correctly identified and 36 incorrect.

in percentage is detailed in Table D.1.

A Wilcoxon signed-rank test was conducted to determine the effect of accuracy based on observations of pairs in their respective category (Octaves and Semitones). The nineteen participants' answers were summarized based on each category and paired up to test if there were any significant difference between the pairs. The significance level was set to $\alpha = 0.05$ and compensated using Bonferroni correction $\alpha/6 = 0.008$. The pairs and results are shown in Table D.2.

All participants performed equal or above chance level ($\geq 33\%$) with the worst two participants settled just at chance level at 33% accuracy and the best one was 74.6% correct. Distribution of correct and incorrect answers in their respective category can be seen in Figure D.7 for the contour types, Figure D.8 for octaves and Figure D.9 for semitone intervals.

D.5 Discussion

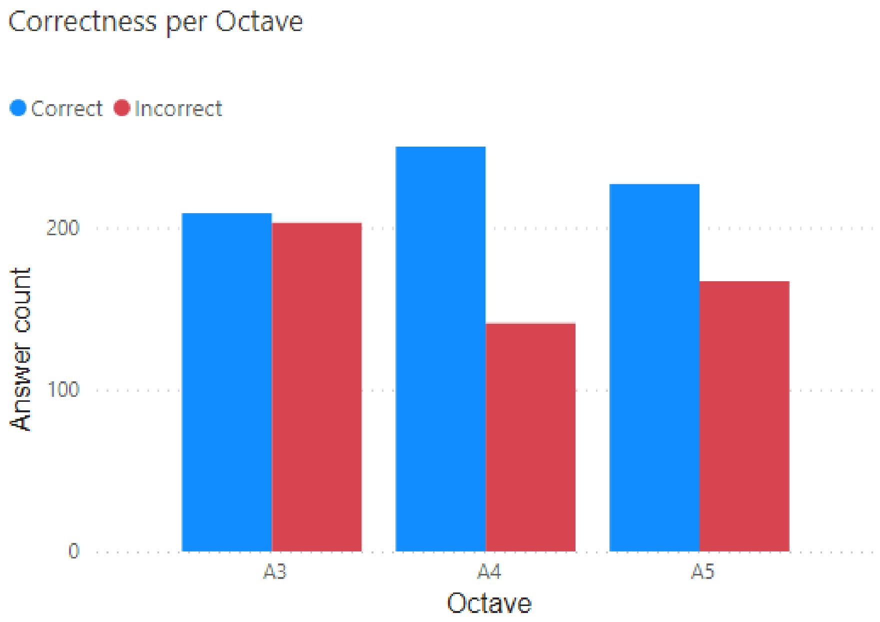


Fig. D.8: Results of the MCI test shown as correct and incorrect answers by octave. The plot shows that the A4 octave has the largest proportion of correctly identified answers and A3 has the lowest.

D.5.1 Analysis of Results

The distribution of correct answers shown in Figure D.11 shows that participants are able to achieve an accuracy that is equal to or higher than chance level. Plotting the distribution of correct and incorrect answers by the different types of contours shows some of the key areas where participants are more consistent in correctly identifying the contours. Figure D.7 shows that participants are most accurate with flat contours with an accuracy of 80.8%. The flat contour was also the most commonly given answer for a contour as shown in Figure D.10. This implies that participants struggled to discerning the difference in pitch between notes. As shown in Figure D.9, participants performed best with zero semitone interval (flat contours), and for the remaining semitone intervals, there is an improvement in accuracy with each increase in semitone interval. The interval of two semitones is the only interval which has less correctly identified contours than incorrect. There is a statistically significant median increase in accurate answers between two (median of 9) and five (median of 11) semitones, $p = .002$. This could be explained by the difference in pitch being the smallest for the three interval types causing participants to interpret Rising and Falling as Flat. The percentage of answers given as *"flat contour"* decreases as the semitone interval

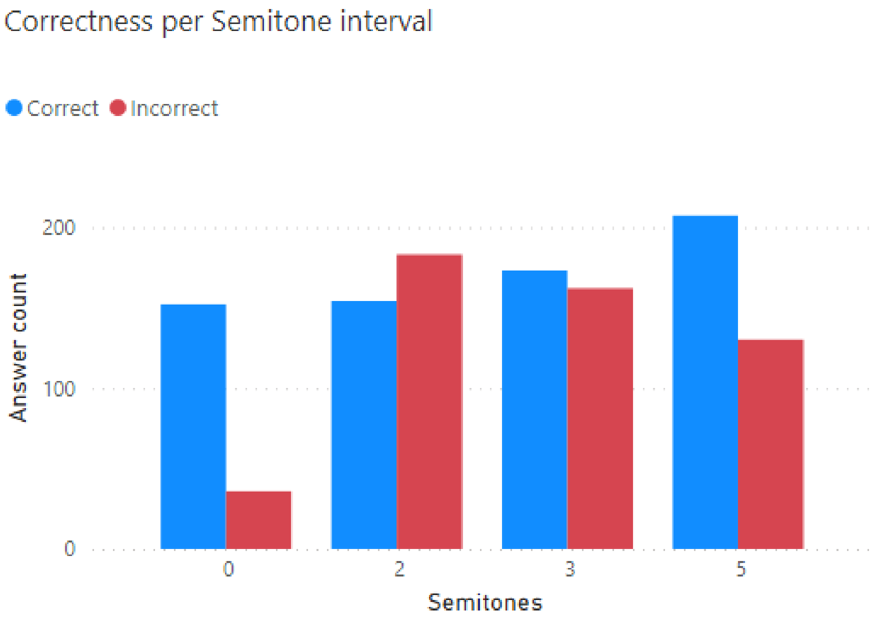


Fig. D.9: Results of the MCI test shown as correct and incorrect answers by Semitone interval between each note in the contours. The interval of 0 corresponds to all of the flat contours. The 0 semitone difference shows the largest difference with most of answers being correct. From 2 to 5 semitones the proportion of correct answers increase with only the 2 semitone difference showing less correctly identified contours than correctly identified.

increases, as shown in Table D.3, implying that it becomes easier to discern the difference in pitch as the interval between notes increase.

Of the three different octaves, A3 had the lowest accuracy. The frequency of the root note A3 is 220Hz which puts it outside of the cutoff frequency for the initial band pass filter used. This causes all of the notes in a falling and flat contour using A3 as root to be filtered by the band pass filter, with only the rising contour having notes that are not filtered. By filtering the results by the A3 octave, only the Rising and Flat contours show a positive difference in correctly identified contours. For the falling contour in A3, 110 out of 167 were not correctly identified, this makes up 45% of all incorrect answers for falling contours. Given that the notes for falling and flat both fall outside the lower limit of the frequency range (250Hz), this may increase the difficulty discerning the the difference between the two contours. This also explains why rising contours has a higher percentage of correct answers compared to falling in A3.

The Octave with the highest count of correct answers is A4, with a positive difference in the amount of correct answers compared to incorrect across all contours. The frequency of the root A4 is 440Hz, putting it 10Hz below

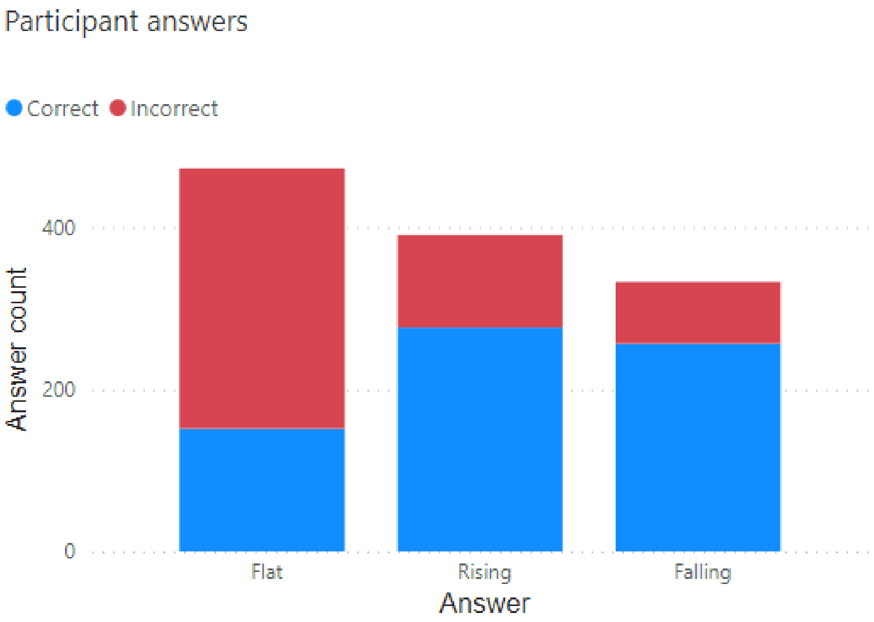


Fig. D.10: Distribution of answers by the participants for all contours. The most common selected contour was Flat. Falling was the least selected contour.

the upper cutoff frequency of the first channel band, making rising contours span up to 3 channels depending on semitone interval. For falling contours, semitone intervals of three and five will go below the limit of the frequency range. Results for falling and rising contours are similar, which could suggest that the initial notes of the contours are the most important for determining the type of contour.

Looking at the octaves themselves, the p-values shown in Table D.2 does not suggest that the octaves on their own influence the amount of correct answers. The only pair observed that produces a p-value that approaches 0.05 is A3 and A4 ($p = .07$), but as discussed earlier, this is likely due to other factors.

D.5.2 Simulated vs. Actual

The results of the MCI test show a mean result similar to the one performed in the reference paper [10], in which nine CI subjects performed a MCI test. Nevertheless, they show a much wider standard deviation in results due to participants using their clinically assigned CI, which includes several models as well as users having different backgrounds with their CI's. The detailed results show a similar issue in determining contours with small intervals

	%	A3	A4	A5	semi-2	semi-3	semi-5
Falling	51.20	34.13	62.87	56.55	42.86	52.38	58.43
Flat	80.85	75.00	76.36	92.98	-	-	-
Rising	54.63	56.21	60.94	46.75	48.52	50.90	64.33
Mean	57.31	50.00	65.79	56.89	45.03	53.56	60.52
SD	12.80	13.72	23.29	18.81	13.47	18.49	20.11

Table D.1: Correct answers in percentage of the contour combinations.

X	Y	p	p ¡ 0.008
Octave A3	Octave A4	0.07	No
Octave A3	Octave A5	0.5	No
Octave A4	Octave A5	0.3	No
Semitone 2	Semitone 3	0.2	No
Semitone 2	Semitone 5	0.002	Yes
Semitone 3	Semitone 5	0.02	No

Table D.2: Wilcoxon signed rank test for zero median with paired samples within their own category of contour test cases. The test pairs that has a statistically significant median difference is the contour pairs Semitone 2 and 5.

between notes, with the accuracy increasing with increases in semitone intervals. The results with octaves in [10] vary from user to user with some users performing better in some octaves that others. The results of the test performed on this simulator favors the A4 octave, but configured differently the results might be more in line with that of actual CI models. The CI users tested in [10] showed a similar tendency as the participants of this test, with the most frequent answer being *flat (contour)* and the least one being *falling (contour)*. Comparing the configuration of this simulator to the characteristics of the CI models used in [10], the one described in this article uses a shorter frequency range than that of the CI models. The lower frequency limit go as low as 120Hz and the upper frequency limit as high as 10853Hz. Configuring the simulator to have a similar range could improve performance as it would solve some of the previously mentioned issues with notes falling outside the band-pass filter. Results of both tests support that many contours are misinterpreted as flat contours. This is a result of the short intervals between notes, that when processed, becomes difficult to discern due to the limited spectral resolution.

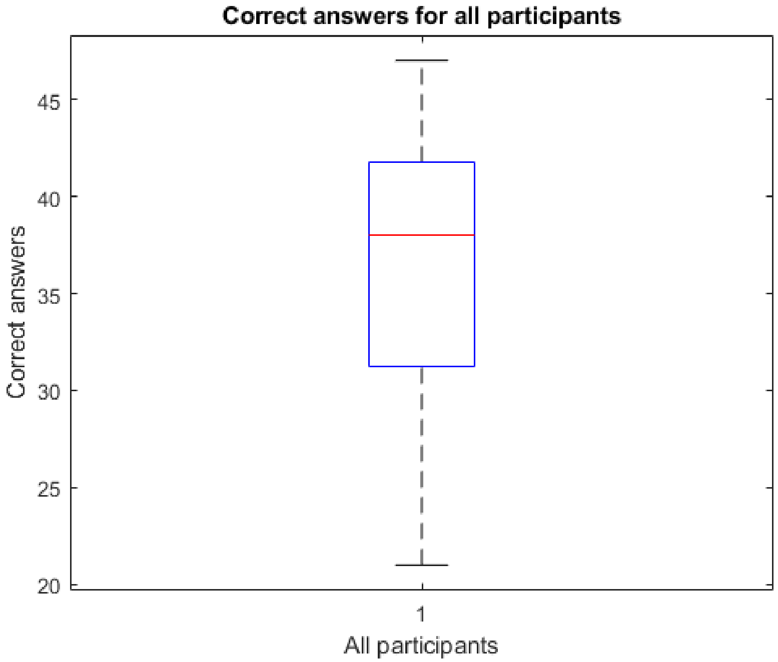


Fig. D.11: Box plot of the amount of correct answers per participant. The median of thirty-eight (38) is marked by the red line and the 25th percentile and the 75th percentile marked by horizontal blue lines showing 31.25 and 41.75 respectively. The lowest scoring participant had a score of twenty-one (21) and highest scoring participant had a score of forty-seven (47).

D.5.3 Research Limitations

The limited number of participants for the test reduces the amount of confidence in some of the results. The results produced from the test show a significant deviation in results between participants in certain categories. The limited amount of participants also makes determining outliers difficult. The test would benefit from a larger set of participants, making results more reliable.

D.5.4 Future Research Suggestions

Galvin et al. [10] showed that accuracy in determining contours of its participants increased with training. Rerunning the test with the simulator with a fixed set of NH subjects could investigate if this is also the case for a simulated environment. Rerunning the test with different sets of parameters could provide insight into which parameters impact the ability to perceive musical contours. Furthermore, in [16] it is implied that the PSHC carrier was judged more similar than a SINE or NOISE carrier for vocoder-based simulation of

%	Semitone 2	Semitone 3	Semitone 5
Flat	40.94	30.01	24.33

Table D.3: Percentage of answers given as Flat per semitone interval

CI by SSD-CI subjects. It could be interesting to evaluate if mixing different set of carriers was judged more or less similar to their CI ear.

Lastly, this study did not utilize the real-time feature of the CI simulator. Many variables must be taken into consideration when conducting an experiment in real time in terms of stimuli requiring a separate study investigating the interaction between different parameter configurations. Nevertheless, allowing participants to select their desired configurations could open up for interesting and individual results. Lastly, the experiment described in this article aimed to be as comparable as possible to the one presented by [16], in order to understand how our simulator performs relative to the reference one, therefore the real-time feature was not necessary for the MCI evaluation.

D.6 Conclusion

The results of the MCI test show that NH listeners show similar tendencies to CI listeners when identifying melodic contours that has been simulated. In general it becomes more difficult to discern differences in pitch in melodic contours, with test participants often false classifying contours as flat. This becomes increasingly clear as the interval between notes decrease. The test performed could benefit from having more participants, as some of the results gathered show a high standard deviation in accuracy between participants. In summary, the parameters used with the PSHC processor can mimic the reduced spectral resolution of a CI in a musical context. This could be used to further investigate which parameters could improve the music listening experience for CI users.

References

[1] Y. Adel, S. Nagel, T. Weissgerber, U. Baumann, and O. Macherey, "Pitch matching in cochlear implant users with single-sided deafness: Effects of electrode position and acoustic stimulus type," *Frontiers in Neuroscience*, vol. 13, 2019. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2019.01119

[2] S. A. Ausili, B. Backus, M. J. Agterberg, A. J. van Opstal, and M. M. van Wanrooij, “Sound localization in real-time vocoded cochlear-implant simulations with normal-hearing listeners,” *Trends in Hearing*, vol. 23, 2019.

[3] M. Bernard, “libgammatone: a generic c++ library for gammatone filterbank,” 2016. [Online]. Available: https://github.com/mmmaat/libgammatone

[4] E. D. Casserly, “Effects of real-time cochlear implant simulation on speech production,” *The Journal of the Acoustical Society of America*, vol. 137, 2015.

[5] D. Cavdir, F. Ganis, R. Paisa, P. Williams, and S. Serafin, “Multisensory Integration Design in Music for Cochlear Implant Users,” 06 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6798118

[6] M. F. Dorman, S. C. Natale, A. M. Butts, D. M. Zeitler, and M. L. Carlson, “The sound quality of cochlear implants: Studies with single-sided deaf patients.” *Otology and Neurotology*, vol. 38, pp. e268–e273, 2017. [Online]. Available: https://doi.org/10.1097/MAO.0000000000001449

[7] C. Fuller, D. Başkent, and R. Free, “Early deafened, late implanted cochlear implant users appreciate music more than and identify music as well as postlingual users,” *Frontiers in Neuroscience*, vol. 13, 2019. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2019.01050

[8] J. J. Galvin, Q. J. Fu, and R. V. Shannon, “Melodic contour identification and music perception by cochlear implant users,” vol. 1169, 2009.

[9] J. Galvin, E. Eskridge, S. Oba, and Q. J. Fu, “Melodic contour identification training in cochlear implant users with and without a competing instrument,” *Seminars in Hearing*, vol. 33, 2012.

[10] J. J. Galvin-III, Q.-J. Fu, and G. Nogaki, “Melodic contour identification by cochlear implant listeners,” 2007. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3627492/

[11] T. Goehring, A. Archer-Boyd, J. Arenberg *et al.*, “The effect of increased channel interaction on speech perception with cochlear implants.” 2021. [Online]. Available: https://doi.org/10.1038/s41598-021-89932-8

[12] D. D. Greenwood, “A cochlear frequency-position function for several species–29 years later,” 2017. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/2373794/

[13] G. Hilkhuysen and O. Macherey, "Optimizing pulse-spreading harmonic complexes to minimize intrinsic modulations after auditory filtering," *The Journal of the Acoustical Society of America*, vol. 136, no. 3, pp. 1281–1294, 2014. [Online]. Available: https://doi.org/10.1121/1.4890642

[14] L. K. Holden, J. B. Firszt, R. M. Reeder, R. M. Uchanski, N. Y. Dwyer, and T. A. Holden, "Factors affecting outcomes in cochlear implant recipients implanted with a perimodiolar electrode array located in scala tympani," *Otology and Neurotology*, vol. 37, pp. 1662–1668, 2016. [Online]. Available: https://doi.org/10.1097/MAO.0000000000001241

[15] S. Jain and P. G. V. Ghosh, "Acoustic simulation of cochlear implant hearing: Effect of manipulating various acoustic parameters on intelligibility of speech," *Cochlear Implants International*, vol. 19, 2018.

[16] C. Karoui, C. James, P. Barone, D. Bakhos, M. Marx, and O. Macherey, "Searching for the sound of a cochlear implant: Evaluation of different vocoder parameters by cochlear implant users with single-sided deafness," *Trends in Hearing*, vol. 23, p. 2331216519866029, 2019, pMID: 31533581.

[17] A. Lehmann, C. J. Limb, and J. Marozeau, "Editorial: Music and cochlear implants: Recent developments and continued challenges," *Frontiers in Neuroscience*, vol. 15, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2021.736772

[18] J. Marozeau, N. Simon, and H. Innes-Brown, "Cochlear implants can talk but cannot sing in tune," *Acoustics Australia*, vol. 42, no. 2, pp. 131–135, 2014.

[19] H. J. McDermott, "Music perception with cochlear implants: A review," *Trends in Amplification*, vol. 8, no. 2, pp. 49–82, 2004, pMID: 15497033. [Online]. Available: https://doi.org/10.1177/108471380400800203

[20] A. Mehta and A. Oxenham, "Vocoder simulations explain complex pitch perception limitations experienced by cochlear implant users," 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5688042/

[21] Q. Mesnildrey, G. Hilkhuysen, and O. Macherey, "Pulse-spreading harmonic complex as an alternative carrier for vocoder simulations of cochlear implants," *The Journal of the Acoustical Society of America*, vol. 139, no. 2, pp. 986–991, 2016. [Online]. Available: https://doi.org/10.1121/1.4941451

[22] Q. Mesnildrey and O. Macherey, "Simulating the dual-peak excitation pattern produced by bipolar stimulation of a cochlear implant: Effects

on speech intelligibility," *Hearing Research*, vol. 319, pp. 32–47, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378595514001841

[23] L. Migirov, J. Kronenberg, and Y. Henkin, "Self-reported listening habits and enjoyment of music among adult cochlear implant recipients," *Annals of Otology, Rhinology & Laryngology*, vol. 118, no. 5, pp. 350–355, 2009, pMID: 19548384.

[24] S. Paquette, G. Ahmed, M. Goffi-Gomez, A. Hoshino, I. Peretz, and A. Lehmann, "Musical and vocal emotion perception for cochlear implants users," *Hearing Research*, vol. 370, pp. 272–282, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378595518300388

[25] J. P. M. Peters, A. W. Wendrich, R. H. M. van Eijl, K. S. Rhebergen, H. Versnel, and W. Grolman, "The sound of a cochlear implant investigated in patients with single-sided deafness and a cochlear implant." *Otology and Neurotology*, vol. 39, pp. 707–714, 2018. [Online]. Available: https://doi.org/10.1097/MAO.0000000000001821

[26] S. F. Poissant, N. A. Whitmal, and R. L. Freyman, "Effects of reverberation and masking on speech intelligibility in cochlear implant simulations," *The Journal of the Acoustical Society of America*, vol. 119, 2006.

[27] R. V. Shannon, F.-G. Zeng, V. Kamath, J. Wygonski, and M. Ekelid, "Speech recognition with primarily temporal cues," *Science*, vol. 270, no. 5234, pp. 303–304, 1995. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.270.5234.303

[28] N. A. Whitmal, S. F. Poissant, R. L. Freyman, and K. S. Helfer, "Speech intelligibility in cochlear implant simulations: Effects of carrier type, interfering noise, and subject experience," *The Journal of the Acoustical Society of America*, vol. 122, 2007.

[29] B. S. Wilson and M. F. Dorman, "Cochlear implants: A remarkable past and a brilliant future," *Hearing Research*, vol. 242, 2008.

Paper E

Multisensory Integration Design in Music for Cochlear Implant Users

D. Cavdir, F. Ganis, R. Paisa, P. Williams and S. Serafin

The paper has been published in
The 19th Sound and Music Computing Conference Conference Proceedings,
2022.

The layout has been revised.

Abstract

Cochlear implant (CI) users experience several challenges when listening to music. However, their hearing abilities are greatly diverse and their musical experiences may significantly vary from each other. In this research, we investigate this diversity in CI users' musical experience, preferences, and practices. We integrate multisensory feedback into their listening experiences to support the perception of specific musical features and elements. Three installations are implemented, each exploring different sensory modalities assisting or supporting CI users' listening experience. We study these installations throughout semi-structured and exploratory workshops with participants. We report the results of our process-oriented assessment of CI users' experience with music. Because the CI community is a minority participant group in music, musical instrument design frameworks and practices vary from those of hearing cultures. We share guidelines for designing multisensory integration that derived from our studies with individual CI users and specifically aimed to enrich their experiences.

E.1 Introduction

While cochlear implants (CI) have achieved a high level of complexity in terms of hardware and ergonomics, training and rehabilitation programs for cochlear implant users are still lacking. This technology is quite advanced for facilitating speech perception, but music appreciation and rendering prove to be underwhelming. Specifically, most CI users report the inability to properly recognize the timbre and pitch of musical instruments, or have issues of sound localization [7, 8]. Additionally, they state to struggle with segregating the individual instruments in multi-instrument mixing [12]. In this paper, we describe a participatory design approach to designing novel technologies to help hearing impaired users' experience and appreciate music.

The hearing abilities, profiles, and perceptions significantly vary among people experiencing hearing impairments. This diversity is even wider among cochlear implant (CI) users due to "age, cognitive processing residual hearing, hearing aid use, and musical training" [13]. When approaching musical experience design for CI users, the assessment and evaluation might need to be process-oriented and individual-specific. Over the course of explorative workshops with CI users, we developed design practices and guidelines for integrating multisensory modalities to enrich their experiences with music. Our motivation derives from providing them with tools to better understand and enjoy musical features that many participant report difficulties.

E.2 Related Work

E.2.1 Accessible and Inclusive Design in Music

When designing accessible digital musical instruments (ADMIs) or accessible music technologies (AMTs), researchers differently approach this design and collaboration / participation process, ranging from participatory approaches to performance and improvisation.

Schroeder and Lucas discuss the process and evaluation of bespoke design approach to accessible music technologies [19]. The authors describe how bespoke designs are vital to provide access for disabled artist to music making. Lucas et al. investigate the evaluation methods for bespoke designs for music and provide their observation on assessing these designs for future ADMI designs [20]. Samuels and Schroeder study improvisation possibilities among performers of different background and abilities for increased inclusion [29] and emphasize the performance aspect in accessible and inclusive design for music.

Dickens et al. practice participatory methods to investigate real life musical interactions for people with complex disabilities and to explore potentials of embodied interactions with gesture-based technology [6]. Another participatory approach by Marti and Recupero [21] focuses on design of smart jewels beyond functionality for Deaf and Hard of Hearing (D/HoH) people with hearing aids. Similar participatory practices and their implications for rehabilitation are explored by accessibility researchers [27]; however their application to musical experience design, specifically for Deaf and Hard of Hearing participants are significantly limited. Like participatory design with D/HoH, community-engaged research with focus on music and hearing impairments is even more limited in this field. Gosine et al. discuss the importance of community building through inclusive music making and its benefits to disabled people through music theraphy [14]. They created collaboration possibilities among persons with physical disabilities and local community musicians following a workshop format.

Frid highlights that the majority of ADMIs focus on addressing users' complex needs in terms of physical and cognitive disabilities, rather than users' experience of music who live with vision and hearing impairments [11]. By 2019, only 6% of ADMIs focused on hearing impairments, even less studied specific cases of cochlear implant use and music.

E.2.2 Cochlear Implant Use and Music

Cochlear implants have witnessed an impressive evolution in the last 30 years, restoring hearing to more than half a million profoundly deaf people. Their success is usually measured through speech recognition tests. Common implant systems achieve 50% - 60% accuracy after 24 month of use when tested on monosyllabic words, and close to 100% on sentences [34]. Some patients achieve spectacularly high results providing proof of what is possible with a neuroimplant in an otherwise totally deaf cochlea. Variability is high though, with standard deviations ranging from about 10% to 30%, for various studies, but results are improving, especially in patients using bilateral implants [34].

The CIs available today still have significant limitations, offering a severely impaired pitch and timbre perception. Another known limitation is the difficulty users have when presented competing sounds; CI users struggle to discriminate musical events when multiple instruments are playing, or long reverberations are present [5, 9]. Furthermore, there is a general weak representation of the fundamental frequencies (F0) for complex sounds, with difference limens ten times lower than hearing without no impairments, even when signals are below that of the CI pitch saturation limit (300Hz) [34]. As a result of these cumulative factors, the evaluation of music experience is not included as a measurement of success for the implants, as the general music experience for CI users is poor.

E.2.3 Multisensory Integration in Music

At the core of this project lies the principle of multisensory integration that explains how humans form coherent experiences by merging information from multiple senses [31]. For this integration to occur, the only requirement is that the stimuli are temporally overlapping; this will produce a perceptual enhancement that is strongest for the stimuli which are least effective [31].

In the specific case of auditory-tactile stimuli, recent studies demonstrate that multisensory integration can in fact occur at very early stages of cognition, resulting in supra-additive integration of touch and hearing [1, 10, 18]. This is especially useful for CI users that are shown to be better multisensory integrators [28]. Furthermore, research within auditory-tactile interactions has shown that tactile stimulus can influence auditory stimulus perception when presented in unison [24, 35].

Multisensory integration has been exploited extensively in previous research focusing on tactile augmentation of music; in 2009 Karam et. al. drew inspiration from previous sensory substitution vocoders and aimed to increase the audio-tactile resolution through the skin [17]. Their project resulted in a chair that provided 4 pairs of voice coil actuators arranged in an array along the back rest, following the cochlea metaphor - lower frequencies are reproduced lower than the higher ones. Each one of the actuators could reproduce one octave of the piano, from 27.5Hz to 4186 Hz [17]. They evaluated their design with respect to emotional reaction and concluded that participants enjoy the two proposed techniques more than the audio signal alone. Further upgrades to the chair resulted in a a wide spectrum of feedback, mostly positive [2].

Another chair installation was designed by Nanayakkara et al. with the help of the hearing impaired community [23]. Initially, their haptic chair had two contact speakers as haptic transducers placed under the armrest that was upgraded later with actuators directed at the lower back area and a footrest, providing a *whole body stimulation* [22]. The actuators were reproducing an amplified version of the auditory stimuli, and was always used in conjunction with sound. The chair was used successfully in long term studies (12-24 weeks) to enhance the music listening experience, as well as speech therapy for deaf children, and underlying the importance of training when users are expected to adapt a novel haptic system [22].

In 2015 a collaboration between the Deaf arts charity organization *Incloodu* and Queen Mary University resulted in an installation in the shape of an armchair and a sofa [15]. The devices used voice coil actuators placed in the backrests and armrests, and a subPac [1] under the seat. The auditory signals were spatialized from low to high areas of the backrest, and a noisy component correlated to timber was reproduced through the armrests. The structure was designed by a profoundly deaf architect, specialized in developing interiors for hard-of-hearing customers [15]. Their evaluation shows that the type of music has a great impact on the experience, with highly rhythmic music eliciting more positive reactions than music where harmonic motion was most important [15]. When music with less transients was presented, users seemed to observe the therapeutic value of vibrations. This emphasizes an important aspect of vibrotactile musical devices: they should be designed in manner that places the musical context in the spotlight.

E.3 Research Approach

The goal of this study was to (1) invite CI users into the early stages of designing novel audio-tactile displays by introducing several multisensory installations and (2) to understand the limitations of presented configurations. We performed an exploratory study, collected by a triangulation of methods: think aloud protocol, observations, and enter and exit interviews [30].

E.3.1 Workshop Format

Each meeting followed a predefined structure and lasted 60 - 120 minutes; for the entire duration there was one of the authors taking notes and recording the conversations. Before the meeting, the participants were requested to fill an online survey, focusing on demographics and their past and current music listening habits. The answers from this survey formed the foundation for an semi-structured interview that was conducted before any installations were introduced; the focus was on exploring further the music engagement habits. Subsequently, the participants were guided to explore and experiment different installations described in section E.4, and concluded with a shorter exit interview, summing up their feedback. Throughout the whole meeting, the participants were in contact with at least one of the authors, and were encouraged to *think aloud*.

---

[1] https://subpac.com/

E.3.2 Participants

Three participants voluntarily participated in the study, invited via open invitation on the national CI user's Facebook[2] group or via email.

Participant 1 (P1) is 52F and started losing her hearing at the age of 3, currently with no residual hearing. In 2017, she got bi-implanted with Kanso CI, experiencing a positive transition from hearing aid to cochlear implants. She likes *Fleetwood Mac*, *Dolly Parton* or *The Beatles*, but dislikes techno, classical music and heavy metal. She has background in piano and dancing (in African and Danish dances). She sings in a choir but is challenged in distinguishing and synchronizing with accompaniment, misidentifying when to start singing. She reported using a water bottle or glass in her hands to feel the vibrations in concerts.

Participant 2 (P2) is 69M with genetic hearing disability, uses a cochlear implant in his left ear, and a hearing aid in his right ear. He has experience from a musician family, in singing in a church choir, and performing competitive dancing. He likes opera, waltzes, church and classical music, and dislikes rock. More recently, he rarely listens to music. When listening to familiar music, he expresses: *"[...]my memory was another [...] I have this sort of feeling of something is in another way."*

Participant 3 (P3) is 41M. He uses a Nucleus Cochlear implant in the right ear, and near deaf in the left ear, with hearing threshold at +95dB. He has been using hearing aids since the age of 3, frequently upgrading them to higher amplification ones. When listening, he can identify when music is playing, the sex of the singer, and the instrument if the music is performed live on stage. He regularly attends to festivals, mostly for the social reasons. Lately, he enjoys listening to music for short periods of time (5 minutes) since after about 10 minutes it becomes exhausting. He mostly likes rock, especially the band *Dizzy Mizz Lizzy*.

E.4 Design and Implementations

E.4.1 Installation 1

CI users experience significant difficulties in identifying individual instruments in a musical piece [25]. In this installation, we addressed this issue by creating a multi-channel listening experience. The installation tested CI users' instrument segregation process through reproducing multi-channel recordings in a four channel speaker setup. We encourage the listeners to freely move around the room and hear individual sound sources to compare and contrast the single and multi-instrument mixings.

---

[2] Facebook CI Group

Setup

The experiment was conducted on campus at Aalborg University Copenhagen, in an anechoic room in order to prevent room reverberation altering or reducing loudspeaker directionality. The setup consisted of four *Dynaudio BM5 MKIII* loudspeakers connected to a laptop through an *Steinberg UR44C* audio interface. Each loudspeaker was fed with a dedicated output from the audio interface with only one instrument. We played multi-track recordings using Reaper - a Digital Audio Workstation (DAW) to route the instruments to independent speakers: (1) drums, (2) bass, (3) vocals, (4) keyboard or guitar alternating.

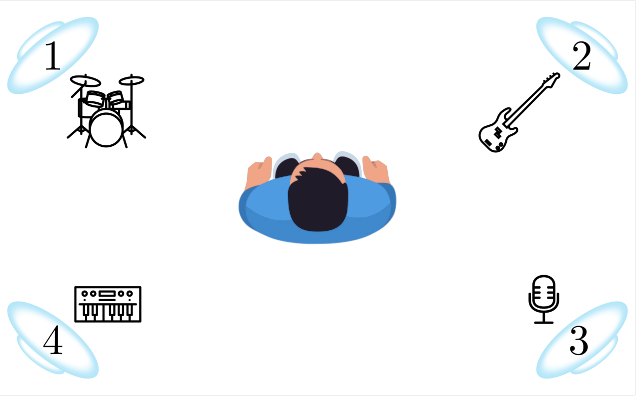


Fig. E.1: Scheme of installation 1.

For all the three sessions with the participants, no routing changes were applied to maintain consistency between the experiences. The dB level of each channel was set to obtain a balanced mix that allowed a hearing person to perceive all the instruments with perceived equal loudness in the center of the room by the authors. The single recordings were played without any effect such as reverberation or compression to avoid any possible confusion in the listener.

Experience

Once entered the room, we explained briefly what the experience was about and we let the test subject choose which music they preferred between three famous Rock, Soul and Reggae songs. Later, we proceeded setting a proper loudness level that was agreed together with the user. For all the test we set all channels to a *conversation level*.

For the first part of the experiment, we asked the subject to stand in the middle of the room and try to identify which instruments were played and from which loudspeaker they were coming from. After collecting the answers, we asked the user to walk around the room moving close to each loudspeaker to confirm or correct his/her statement about which and where instruments were played. For the second and last part, we let the subject find a sweet-spot in the room where the music sounded best for him/her. During the whole experiment the test subject was free to comment or explain at any moment their thoughts and perception of the experience.

E.4.2 Installation 2

A design process was undertaken to explore if and how audio-tactile feedback might be integrated into a seating installation to enhance CI users' music listening experience. We focused on providing low-frequency enhancement since CI users experience poor auditory resolution in this range.

We tested two mock-ups with 3 CI users and 3 hearing participants (including the designers). Each mock-up consisted of three components, a seat, a footrest and a hand held device, used both independently and simultaneously. All users accessed to the gain control for each actuator, through a headphone splitter used to feed the same signal to the amplifier for each transducer. Only the first user chose to manipulate the gain balance herself, while the last two provided verbal instructions to the researchers. The audio was played through either a pair of *B&W 800D* speakers for the first user, and a pair of *Mackie SRM450* + *Mackie SRM1550* for the second and third participant. The users had access to the master volume knob that controlled the audio level, as well as the signal feeding the headphone amplifier (used here as a multi-channel signal splitter), thus coupling the auditory and the tactile volume.

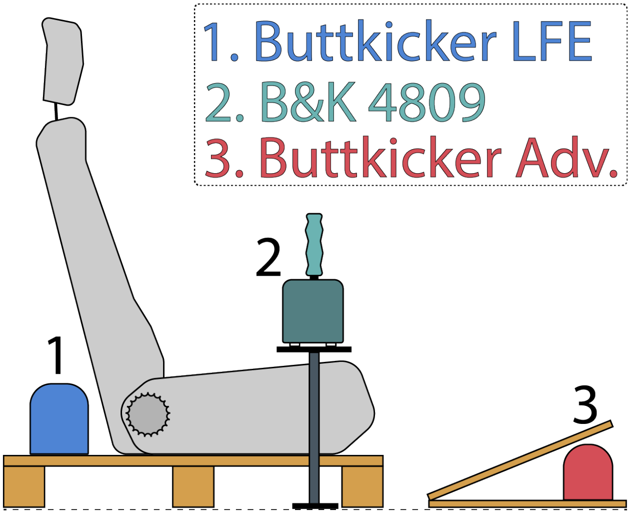


Fig. E.2: One configuration experienced by all participants

Hardware

Three types of seated installations provided different experiences: The first installation was a tactile car seat actuated by a *Buttkicker*[3] LFE that was initially powered by a *Buttkicker BKA1000* and later StageLine ST600 in bridge mode. The *BKA1000* amplifier was found to limit the higher frequencies. Both the chair and the actuator were bolted onto a wooden EUR-pallet platform, with the actuator behind the seat(see Figure E.2). The actuator provided strong enough tactile feedback throughout the entire body, including the headrest. P1 and 3 hearing participants reported that it could easily felt overwhelming with higher gain.

The second and third type of seated experience shifted from a low seating position to a more upright one through a bar stool instead of the car seat, based on P1's feedback that rated the first design overwhelming. We chose the bar stool design since it affords control over the amount of weight the user applies onto, linking to the amount of feedback received. A *Buttkicker Advanced* powered by a *Buttkicker BKA300* actuated this seating. As a much smaller actuator compared to the one from the car seat, this setup required less power. The authors noticed a substantial difference between the frequency responses of the two, with a high frequency emphasis for the setup with *Buttkicker Advanced*.

P2 and P3 experienced different configurations; for P2, the actuator was bolted perpendicular to the seating area, while for P3, the actuator was fixed parallel to the ground on the side of the seating area. The side actuator configuration aimed to conduct more tactile stimuli, as P2 commented on the low intensity of the bar stool (possibly in comparison to the car seat). We observed an unexpected phenomenon in P3's setup that loud transients laterally shook the bar stool, feeling like a small "kick in the back of the chair", potentially due to the loose joints.

The footrest was designed according to H. Dreyfuss measurement recommendations and featured an inclined plane at 22° [32]. In order to have clearance for the actuator underneath the inclined plane, the footrest measured 45cm length and 60cm width. The same *Buttkicker Advanced + BKA300* combination was used, as with the bar stool. The transducer was bolted underneath the footrest, perpendicular to the ground (see Figure E.2). All participants experienced the same setup.

Two handheld devices were used. A cylindrical handheld grip measuring 204mm in length and 110mm in diameter was fabricated by stacking 51 laser cut slices of 4mm HDF, following design recommendations from H. Dryeyfuss [32]. This grip was was attached to a *Brüel & Kjær(B&K) Type 4809 portable vibration exciter* (see Figure E.2). The second interface, VAM (Vibrotactile Actuator for Music), was built around the Tactuator BM1C[4] [26] with an ovoid shape measuring 84mm in width, 58mm in height and 89mm depth. P1 indi-

---

[3] https://thebuttkicker.com/

[4] http://tactilelabs.com/

vidually tested the cylindrical grip and the VAM in combination with the seat and footrest, but the latter was deemed ″not adding much″ and abandoned for P2 and P3.

Audio Stimuli

The first audio stimulus, *Peggy Lee*'s *Fever* was presented in every Installation 2 configuration due to its clear instrument separation and the prominence of the female vocal track, matching CI users' appreciation [3]. Firstly, two different signals were sent to the handheld grips in consecutive renditions: the first identical to other actuators' signal and the second filtered to isolate the female vocal range and pitch shifted to one octave lower to skin's sensitivity range [16, 33], only applied to P1's experiment. P2 and P3 heard solo bass improvisation of *Fever* on the double bass or ukulele bass. The performance presented different playing styles (pizzicato, slapping, staccato, etc.), using the full range of the instrument, and experienced in bar stool and car seat setups. The drums accompanied to P3's experiment.

Participants selected extra audio material for their preferred setup. All three preferred the setup in Figure E.2. P1 listened to *Fleetwood Mac - Dreams*, P2 *Vienna Philharmonic - An der schönen, blauen Donau* (excerpts), and P3 *Dizzy Mizz Lizzy - Silverflame*.

E.4.3 Installation 3

We also studied participants' experience with embodied interactions using movement-based performance and in-air haptics. To simulate this experience, we discussed excerpts from a previous inclusive performance study, Felt Sound, designed for both D/HoH and hearing audience members [4]. Originally, Felt Sound, consisting a digital musical instrument, a performance setting, and a user study, was performed in-person with an 8-subwoofer speaker setup where the participants sat close to or touched the speakers. Due to time and space restrictions and COVID-19 precautions, performance excerpts[5][6] were individually shared with the participants with two subwoofers enclosing their sitting area and facing the participant. The participants were still encouraged to interact with the speakers and feel the vibrations through touch.

We briefly described Felt Sound's motivation, concept, and performance practice. After providing the participants with its context, we presented its excerpts. Following the performance, we discussed their experience both with Felt Sound and with their own movement and music practices. Presenting a new movement-based musical concept led participants to share their own associations and experiences with movement practice and music.

---

[5] https://tinyurl.com/2p8axhwp

[6] https://tinyurl.com/yck63zbz

E.5 Experiences and Results

We audio recorded the discussions with each participant and transcribed them after the study. This chapter will present a summary of these discussion sessions, focusing on their appreciation of the installations and their overall experiences.

E.5.1 First Participant

P1 listened using 4 vibrotactile devices: car seat, footrest, hand grip, the VAM, in 2 cases (processed and unprocessed signals) as detailed in Section E.4.2. The audio volume was tied to the overall actuator amplitude. The researchers initially set the individual tactile amplitudes to "perceptually equal" and the listening volume to "comfortably loud", slightly over conversation level.

In the first case (listening to the processed audio), she reported how it was *"fun to feel the vibrations in the entire body"*, re-iterating her experience with the water bottle during concerts (see Section E.3.2). She did not understand the mapping of the vocals to the haptic feedback, stating that she could already hear the voice through the speakers, and would not need extra stimuli representing the vocals. Additionally, she only adjusted the volume of the hand grip up several times.

When presented with the second case (listening to the unprocessed audio), she seemed more engaged in the song, grooving with the rhythm and moving to music. Similar to the first case, she experimented with slightly turning up the hand grip, footrest, and seat. When the song was over, she stated that she preferred this listening method over the first case because she can feel the melody in the footrest. She also expressed that listening to the vibrations through the chair setup could sometimes feel overwhelming.

Her perception changed over the course of the experiment. She reported that she could feel the vocals through the hand grip and the bass line (initially she assumes it was a keyboard) through the foot pad and the seat, expressing that it was fun. Although all actuators reproduced the same signal, different haptic experiences where perceived at different locations on the body, that amplified their perception of pitch and instrument type. She answered to whether she would use such a device at a concert as *"I would like to have some help from vibrations"* and explained how she sits very close to the speaker at concerts to get the haptic feedback. Furthermore, she said she would use them if *"it is trusty"*. She less emphasized her experience using the VAM compared to the other haptic listening tools, stating that it was not strong enough. We interpreted her articulations about the VAM as "not strong, relative to other actuators".

After listening to Installation 3, she discussed her experience with music and movement. This installation led her to articulate her movement practice and more embodied experiences with music such as singing. She reported that when she sings in a choir, she experiences the difficulty of identifying the onsets, specifically knowing when to start singing only by listening to the piano. She stated that she would be interested in incorporating gestures to her singing practice to assist her and to support her conductor's assistance for her. Additionally, she expressed that seeing a gesture-based performance was supporting her understanding and enjoyment of music.

E.5.2 Second Participant

P2 listened to *Ain't No Mountain High Enough* by *Marvin Gaye & Tammi Terrell* with Installation 1. When listening to the piece in the middle, he correctly identified the left and right channels of the instrument sources. However, he guessed the incorrect instruments at each channel. After we asked him to move closer to each speaker, he correctly identified all the instruments, including the male and female voice alternating, not being able to distinguish the lyrics. Similar to the voices, he was able to identify that the guitar and keyboards were playing together in the same channel. He was very unsure of his answers, stating that *"it's always about guessing"*. He always directed his non-implanted ear towards the speakers, making use of his hearing aid.

We lastly asked him to freely select a spot where the music sounds the best for him. He chose a spot in the middle of the 1-2 3-4 speaker pair, closer to the 1-2 speakers, and said *"... I think this must be the ideal (spot) for this kind of music that all of it is, is possible to hear."* After being exposed to all instruments individually he said that they became clearer once he separately heard and identified them. Similarly, when identifying the lyrics, he could follow them once he was told what the chorus lyrics were.

The second installation consisted of the car seat, the bar stool (with vertical actuator) the footrest and the hand grip powered by the *B&K* actuator, with the same volume settings as initially set for P1. The setup was split in two: (1) bar stool with footrest and had hand grip and (2) car seat with footrest and hand grip. We played the same music without any processing for the actuators. After approximately 90 seconds of listening through the first setup, we paused the listening for intermediate discussion and the participant described where he most significantly felt the vibrations: in the thigh, ankles, and up to the elbow. He provided verbose feedback regarding the locations and intensities of perceived vibrations, but limited in terms of perceptual qualities of the stimuli. He could identify the female voice and the deep bass. He also stated he could easily identify the melody.

When we asked him how the music made him feel, he said: *"It was more like a little bit sad music."* and stated how there should be more happiness in it for him to appreciate it. Furthermore, when one of the researchers played the double bass solo, P2 appreciated the live music aspect but he stated that he does not like the bass (as an instrument). He further reported that the installation was more involving but influenced by the choice of music since the music piece was not a style of music he enjoys; thus, becoming and enhancement of something he does not prefer. He requested listening to *An der schönen, blauen Donau* composed by J. Strauss. From the very first chord, the participant said *"... yeah this is much better, much better, yeah and I can feel it supports the music. So, if you like the music, this gives extra power"*. The second setup was experience only with the waltz playing, but the discussion diverted towards commercial value of musical experiences, and no feedback on the second setup was noted. He mentioned that he would not use such system (setup 2) in an concert environment, stating that *"[he] is rather conservative, and he'd prefer a regular chair, unless explicitly invited to try on in a concert hall"*.

His experience with Installation 3 varied from P1's. He less enjoyed the low frequency content of the music. He reported that he could feel the vibrations on his body but this form of listening did not enhance his experience of music. He finally stated that the gestural performance aspect of the music was effective.

E.5.3 Third Participant

P3 selected to listen to *Don't stop be now* by *Queen* in Installation 1. By standing in the center, he correctly identified the voice and mentioned that there was a lower volume coming from the speaker that was playing the bass line. After getting closer to each speaker, he quickly identified the voice correctly, and mislabeled the piano as guitar. When he approached the speaker playing the bass line, he experienced difficulty in identifying the instrument, asking if it was a tuba. He correctly distinguished the drums.

When we asked him to choose a favorite spot in the room he walked for several minutes, moving between speakers and overall listening area. The chosen spot spot was equally distant from speakers 1 and 2, and much further from speaker 3 and 4 that he was facing. At this spot, he stated that he could hear "a bit of everything", but only mentioning the drums, bass, and vocals. During the post-experiment discussions, we observed that he enjoyed listening to instruments separately since he could make sense of them on his own terms. He further shared his discussions with other people about the sound of bass (at concerts) that *"[he] could never distinguish [the individual instruments] because everything sounds like "mush", but it was a bit easier in this case, after hearing each instrument separately"*.

The second installation followed a similar structure to experiment with P2, only difference being the orientation of the actuator on the bar stool as described in Section E.4.2. After about 90 seconds (before the second verse), the music was stopped and the participant rapidly mentioned that he mostly felt the hand and the bar stool did not add anything to the experience. When asked, he could not identify the valence of the song. Before resuming the music, all actuators were turned down and we slowly increased their amplitude one by one while we instructed the participant to focus on preference over actuated areas. The results were the same; he preferred the hand grip and the footrest (especially when it was turned up more). He mentioned that it's difficult to identify the mood of the song claiming that on one side it's *"slow and heavy, but the singing (voice) sounds happy"*. For the live ukulele bass performance all actuators were set to initial amplitudes; for feedback, P2 said that he preferred the lower frequencies from the footrest, but when the frequency gets higher, it's better through the hand handle. Additionally, when short and fast notes were played, he reported that it was easier to *"feel what happens"* through the hand grip. Similar to the first case, the bar stool *"did not have much to offer"* in this experience.

Moving to the second setup, the participants mentioned that *"this is much better to have it in the back, this way"* further mentioning that setup 1 felt a bit distant. During this experience, the actuators' volume was manipulated by a researcher leading to the conclusion that it's best when all 3 actuators are perceivable, and that if feels "empty", when the seat is not actuated. After the live bass performance (same as for setup 1), P3 claimed that it's fun to use the setup, but still feels like he is *"missing something"* and that he *"just misses actually being able to enjoy music"*, a fact that was not changed by using the presented setup. Nevertheless, he could *"feel"* the voice more through the hand grip, just as with setup 1. When asked whether he preferred the live performance, or the recorded one, he said that the latter one is nicer because there's more instruments, "more different sounds". This led us to and impromptu drum and bass duo performance with two of the authors, briefly jamming over the bass line from *Fever*. The participant claimed that he always thought the bass sound is coming from the drums (in live shows), but now he understands how to separate the two.

His experience with Installation 3 reflected P1's comments on the gestural performance. He reported that he never experienced a music performance where music was played by the gestures and felt on the body.

E.6 Discussion and Future Directions

E.6.1 Process-oriented Assessment on CI and Music

Due to the variance in CI users' perception, experience, and understanding of everyday sounds, speech, and music, we believe that the experience designs should be personalized to the individual CI users and offer customization. Although CI users might share common difficulties in experiencing music such as pitch identification, source localization, and instrument segregation (auditory streaming), their priorities in addressing these challenges significantly vary from individual to individual. For example, P1 experienced hearing the nuances in pitch variances of singing however due to her music practice, she prioritize practicing onset detection and phrasing to support her singing in choir. Similarly, P2 preferred limiting his experience to the music styles he enjoys and enhancing these specific styles rather than practicing for the gaps in his music perception. Researchers and designers should consider such interpersonal differences not only in hearing profiles but also musical appreciation, engagement, and preferences. The factors such as age, hearing aid use, musical training among many others have significant influence in such design considerations when working with CI users.

Similarly, for many CI users, experiencing music is new and requires constant practice and learning. An ongoing musical engagement where users can practice where they experience difficulty in understanding music becomes crucial. Our assessment approach reflects this process of exploring and understanding CI users' hearing and engaging with technology in ways to both support their hearing development and music appreciation. Their participation in ideation and leading the design directions was crucial to the research process.

Because their reference of music is more subjective when they articulate their music perception and experience, we frequently referred to the current literature on assessing CI hearing and informed our experience design research. We believe that a more holistic approach to supporting CI users' music engagement offers more embodied approaches to listening and music-making. Developing new musical interaction experiences leads an integrated and a participatory research process rather than distinctly dividing design, assessment, and evaluation processes. Additionally, we observed that this process-oriented assessment facilitates designers to find more collaboration opportunities with CI users since finding participants in the CI community still remains one of the biggest challenges. We believe that creating a more formal organization around cochlear implant use and music can support their participation in design and research, enhancing their musical experiences.

E.6.2 Guidelines for Designing Multisensory AMTs

Designers who develop tactile displays for CI users can benefit from creating devices that are flexible and that can account for different musical tastes, hearing abilities, and musical engagement levels. While our sample size limits us from generalizing overall CI users' experience in the broader community, the very different requirements from each participant only underlines the need for flexibility and customization in design. Furthermore, special attention should be taken towards not creating unpleasant experiences, as it was briefly the case for P1 (tactile stimulation too powerful) and P2 (unpleasant music choice). Prior knowledge of target groups can help with the preparation, but a certain step towards this pre-study is ensuring that displays have basic controls for tactile and auditory stimuli levels and in the case of multi-actuator devices, setups have independent control for each transducer in paramount. Another helpful approach is to consider flexible or modular hardware that can be easily reconfigured according to user's needs. Through participatory action, research can explore individual requirements. Lastly, whenever possible, we suggest the integration of visual feedback in forms of gestural or movement-based performance or visualization that can support the gaps in perception from either the tactile or the auditory channel.

E.7 Conclusions

In this paper, we study cochlear implant (CI) users' engagement in music and ways to support their musical experiences both in listening and participating. We conduct exploratory workshops with three participants who all use cochlear implants with different hearing profiles. Based on our discussions, we addressed their individual musical needs and tested their experience in listening music through three different installation setups. Each installation investigated a different musical aspect that CI users experience difficulty perceiving. The motivation behind the installations extends beyond informing CI users about musical content but also to enrich their listening experience and musical appreciation. We discuss key findings, results, our observations on their interaction with these three listening modalities. We detail our process-oriented assessment and provide guidelines for designing multisensory integration to creating musical interaction and experiences, with specific focus on CI users. Our efforts address the lack of available resources for CI users' music perception, understanding, and enjoyment.

Music listening needs to be approached as a multifaceted experience which can be challenging and effortful for the hearing impaired individuals. Moving forward, we hope to utilize our interaction tools and listening experiences for CI users in offering them new rehabilitation and practice frameworks while supporting their musical enjoyment. We further plan to address one of the prominent research challenge and limitation we faced during our workshop series: accessing the cochlear implant users and Deaf communities. We hope to continue our work on music for hearing impairments through building communities and meaningful collaborations between CI users, musicians, designers, and researchers, as there seems to be genuine enthusiasm and interest in using hearing assistive devices s for music, from CI users.

References

[1] *An Exploration on Whole-body and Foot-based Vibrotactile Sensitivity to Melodic Consonance*. Zenodo, 09 2016. [Online]. Available: https://doi.org/10.5281/zenodo.851213

[2] A. Baijal, J. Kim, C. Branje, F. A. Russo, and D. I. Fels, "Composing vibrotactile music: A multisensory experience with the emoti-chair," *CoRR*, vol. abs/1509.05452, 2015.

[3] W. Buyens, B. V. Dijk, M. Moonen, and J. Wouters, "Music mixing preferences of cochlear implant recipients: A pilot study," *International Journal of Audiology*, vol. 53, 2014.

[4] D. Cavdir and G. Wang, "Felt sound: A shared musical experience for the deaf and hard of hearing," in *Proceedings of the 20th international conference on new interfaces for musical expression (nime-20)*, 2020.

[5] M. V. Certo, G. D. Kohlberg, D. A. Chari, D. M. Mancuso, and A. K. Lalwani, "Reverberation time influences musical enjoyment with cochlear implants," *Otology and Neurotology*, vol. 36, 2015.

[6] A. Dickens, C. Greenhalgh, and B. Koleva, "Facilitating accessibility in performance: participatory design for digital musical instruments," *Journal of the Audio Engineering Society*, vol. 66, no. 4, pp. 211–219, 2018.

[7] M. F. Dorman, L. Loiselle, J. Stohl, W. A. Yost, A. Spahr, C. Brown, and S. Cook, "Interaural level differences and sound source localization for bilateral cochlear implant patients," *Ear and hearing*, vol. 35, no. 6, p. 633, 2014.

[8] M. F. Dorman, L. H. Loiselle, S. J. Cook, W. A. Yost, and R. H. Gifford, "Sound source localization by normal-hearing listeners, hearing-impaired listeners and cochlear implant listeners," *Audiology and Neurotology*, vol. 21, no. 3, pp. 127–131, 2016.

[9] M. D. Fletcher, N. Thini, and S. W. Perry, "Enhanced pitch discrimination for cochlear implant users with a new haptic neuroprosthetic," *Scientific Reports*, vol. 10, 2020.

[10] J. Foxe, G. Wylie, A. Martinez, C. Schroeder, D. Javitt, D. Guilfoyle, W. Ritter, and M. Murray, "Auditory-somatosensory multisensory processing in auditory association cortex: An fmri study," *Journal of neurophysiology*, vol. 88, pp. 540–3, 08 2002.

[11] E. Frid, "Accessible digital musical instruments—a review of musical interfaces in inclusive music practice," *Multimodal Technologies and Interaction*, vol. 3, no. 3, p. 57, 2019.

[12] J. J. Galvin III, E. Eskridge, S. Oba, and Q.-J. Fu, "Melodic contour identification training in cochlear implant users with and without a competing instrument," in *Seminars in hearing*, vol. 33. Thieme Medical Publishers, 2012, pp. 399–409.

[13] K. Gfeller, V. Driscoll, and A. Schwalje, "Adult cochlear implant recipients' perspectives on experiences with music in everyday life: A multifaceted and dynamic phenomenon," *Frontiers in neuroscience*, p. 1229, 2019.

[14] J. Gosine, D. Hawksley, and S. L. Quinn, "Community building through inclusive music-making," in *Voices: A World Forum for Music Therapy*, vol. 17, 2017.

[15] R. Jack, A. Mcpherson, and T. Stockman, "Designing tactile musical devices with and for deaf users: a case study," *Proc. of the International Conference on the Multimedia Experience of Music*, 2015.

[16] L. A. Jones and N. B. Sarter, "Tactile displays: Guidance for their design and application," *Human Factors*, vol. 50, 2008.

[17] M. Karam, F. Russo, and D. Fels, "Designing the model human cochlea: An ambient crossmodal audio-tactile display," *IEEE Transactions on Haptics*, vol. 2, pp. 160–169, 2009.

[18] C. Kayser, C. I. Petkov, M. Augath, and N. K. Logothetis, "Integration of touch and sound in auditory cortex," *Neuron*, vol. 48, 2005.

[19] A. Lucas, M. Ortiz, and F. Schroeder, "The longevity of bespoke, accessible music technology: a case for community," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2020, pp. 243–248.

[20] A. M. Lucas, M. Ortiz, and F. Schroeder, "Bespoke design for inclusive music: The challenges of evaluation." in *NIME*, 2019, pp. 105–109.

[21] P. Marti and A. Recupero, "Is deafness a disability? designing hearing aids beyond functionality," in *Proceedings of the 2019 on Creativity and Cognition*, 2019, pp. 133–143.

[22] S. Nanayakkara, L. Wyse, and E. Taylor, "The haptic chair as a speech training aid for the deaf," *Proceedings of the 24th Australian Computer-Human Interaction Conference, OzCHI 2012*, pp. 405–410, 2012.

[23] S. Nanayakkara, E. Taylor, L. Wyse, and S. Ong, "An enhanced musical experience for the deaf: Design and evaluation of a music display and a haptic chair," 04 2009, pp. 337–346.

[24] R. Okazaki, H. Kajimoto, and V. Hayward, "Vibrotactile stimulation can affect auditory loudness: A pilot study," 06 2012, pp. 103–108.

[25] A. J. Oxenham, "Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants," *Trends in amplification*, vol. 12, no. 4, pp. 316–331, 2008.

[26] R. Paisa, J. Andersen., N. C. Nilsson, and S. Serafin, "A comparison of audio-to-tactile conversion algorithms for melody recognition," in *Baltic Nordic-Acoustic Meetings*, 2022.

[27] C. Quintero, "A review: accessible technology through participatory design," *Disability and Rehabilitation: Assistive Technology*, pp. 1–7, 2020.

[28] J. Rouger, S. Lagleyre, B. Fraysse, S. Deneve, O. Deguine, and P. Barone, "Evidence that cochlear-implanted deaf patients are better multisensory integrators," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, 2007.

[29] K. Samuels and F. Schroeder, "Performance without barriers: improvising with inclusive and accessible digital musical instruments," *Contemporary Music Review*, vol. 38, no. 5, pp. 476–489, 2019.

[30] M. Someren, Y. Barnard, and J. Sandberg, *The Think Aloud Method - A Practical Guide to Modelling Cognitive Processes*. London: Academic Press, 1994.

[31] B. Stein, P. Stein, and M. Meredith, *The Merging of the Senses*, ser. A Bradford book. MIT Press, 1993. [Online]. Available: https://books.google.se/books?id=uCV9QgAACAAJ

[32] A. Tilley, H. Dreyfuss, and H. D. Associates, *The Measure of Man and Woman: Human Factors in Design*. Whitney Library of Design, 1993. [Online]. Available: https://books.google.dk/books?id=5qzHQgAACAAJ

[33] A. Wilska, "On the vibrational sensitivity in different regions of the body surface," *Acta Physiologica Scandinavica*, vol. 31, 1954.

[34] B. S. Wilson and M. F. Dorman, "Cochlear implants: A remarkable past and a brilliant future," *Hearing Research*, vol. 242, 2008.

[35] G. Young, D. Murphy, and J. Weeter, "Haptics in music: The effects of vibrotactile stimulus in low frequency auditory difference detection tasks," *IEEE Transactions on Haptics*, vol. PP, pp. 1–1, 12 2016.

Paper F

Design and Evaluation of a Multisensory Concert for Cochlear Implant Users

R. Paisa, D. Cavdir, F. Ganis, P. Williams, L. M. Percy-Smith and S. Serafin

The paper has been published in
Arts, Feeling the Future — Haptic Audio 2023

Abstract

This article describes the design, implementation, and evaluation of vibrotactile concert furniture, aiming to improve the live music experience of people with hearing loss using hearing technology such as cochlear implants (CI). The system was the result of a series of participatory design sessions involving CI users with different hearing assistive setups (bi-implant, bimodal, and monoimplant), and it was evaluated in a concert scenario (drums, bass, and female vocals) at the Royal Danish Academy of Music. The project aimed to improve the music appreciation for CI users by providing a multisensory concert designed with CI challenges in mind, but not excluding normal-hearing individuals or individuals with other forms of hearing aids from participating in the event. The evaluation was based on (video-recorded) observations and post-experience semi-structured interviews; the data were analyzed using event analysis and meaning condensation. The results indicate that tactile augmentation provides a pleasant experience for CI users. However, concertgoers with residual hearing reported being overwhelmed if the tactile stimulation amplitude exceeds a certain threshold. Furthermore, devices that highlight instrument segregation are preferred over ones that present a tactile mixdown of multiple auditory streams.

F.1 Introduction

Cochlear implants (CI) are neuroprosthetic devices that partially restore auditory sensations for people with severe to profound hearing loss. The implant directly stimulates the auditory nerve in the cochlea, completely bypassing the acoustic mechanisms of the ear [53]. The electrical stimulation is derived from the auditory input received by an external microphone array, usually located around the implanted ear. The journey of restoring hearing generally starts with the surgical implantation and continues with a long rehabilitation process that requires users to (re)learn how to respond to the new sense of hearing. Multiple factors may influence the final auditory performance of CI users, including device and surgical properties, duration of hearing loss or whether it occurred prelingually or postlingually, the quality of recovery from surgery, therapy and rehabilitation strategies, and many more [26].

After receiving a CI, individuals do not appreciate music any longer, as the device aims to restore speech capabilities, often disregarding musical percepts [36]. Therefore, music appreciation and rendering fall short. Specifically, the majority of CI users report difficulties with sound localization or correctly identifying the timbre and pitch of musical instruments ([13] [13], [14]). These limitations translate into a difficult listening experience of multi-instrument mixes, as a result of poor instrument separation [22]. Nevertheless, the hardware and ergonomics of CI have advanced to a high level of sophistication, and there is clear evidence that training and rehabilitation schemes for CI users result in better musical perception and, as a byproduct, better speech performance. Unfortunately, these programs are few and far between and usually not available to the general public [21].

Academics and the media have focused on music's nonmusical cognitive and academic benefits. While some anecdotal benefits have been repeatedly disproven (e.g., the link between classical music and academic performance in children), several studies have shown the benefits of music listening [47, 48, 57]. In intensive care units, music helps relieve tension, distract from pain, and promote spatial awareness [3, 11, 38]. Other research on the psychological, emotional, and social benefits of music listening suggests that it gives a platform for multifaceted self-related thoughts on feeling and sentiments, escape, coping, consolation, and purpose of life [48]. When it comes to CI users, there is evidence that they benefit from music listening in similar ways as normal-hearing individuals. However, the challenges that CI users experience in perceiving and appreciating music limit them from participating in musical activities as often as they would like, and this has a severe effect on their physical, psychological, and social health [16].

In this article, we extend our efforts to improve the music listening experience of CI users by using vibrotactile devices in a concert scenario, initially described in [8]. We start with a short introduction to the underlying principles and related work, and continue with a presentation of a participatory design workshop aiming to integrate multisensory feedback into listening experiences. Our observations from this workshop resulted in design guidelines for vibrotactile concert furniture that supports the perception of specific musical features and elements. We organized a concert for CI users to evaluate these devices, and the whole process is described in Section F.3.2. Lastly, we summarize our findings and present a discussion and conclusion about our attempt to improve the musical appreciation of concert performance through vibrotactile augmentation.

F.2 Related Work

F.2.1 Accessible and Inclusive Design in Music

Researchers propose several approaches to the design and collaboration/participation process while creating accessible digital musical instruments (ADMIs) or accessible music technology (AMTs), ranging from participatory approaches to performance and improvisation.

Schroeder and Lucas [33] discuss the custom design process and evaluation for accessible music technologies. The researchers analyze the importance of bespoke designs in enabling impaired musicians to create music. According to Lucas et al. investigation's of the criteria for judging custom musical instrument designs, these designs should be evaluated for the next ADMI designs [34]. Samuels and Schroeder [46] focus on the performance element in accessible and inclusive music design and investigate improvisation possibilities among performers of all backgrounds and abilities for improved inclusion.

Dickens and colleagues [12] use participatory methods to study how people with complex disabilities interact with music in real life and to look into the possibilities of embodied interactions with gesture-based technology. Marti and Recupero [35] have proposed a participatory method that focuses on designing smart jewels that are more than just functional for deaf and hard of hearing (D/HoH) people who wear hearing aids. Other researchers interested in accessibility examine comparable participatory practices and what they mean for rehabilitation [44]. However, their use in designing musical experiences, especially for people who are deaf or hard of hearing, is very limited. Like participatory design with D/HoH, research that involves the community and focuses on music and hearing problems is even less common in this field. Gosine et al. [24] talk about how important it is to build community through inclusive music-making and how it can help disabled people through music therapy. Using a workshop format, they made it possible for people with physical disabilities and local community musicians to work together.

Frid [19] points out that most ADMIs are more concerned with meeting the complex needs of users with physical and cognitive disabilities than with how people with vision and hearing problems experience music. By 2019, only 6% of ADMIs were focused on hearing problems, and even fewer looked at how CIs are used with music.

F.2.2 CI Music

People with hearing impairments have vastly different hearing profiles, abilities, and perceptions. This is due to factors including age, cognitive processing residual hearing, hearing aid use, and musical training [23]; this variability is considerably greater among cochlear implant (CI) users.

In the past 30 years, CIs have seen a remarkable transformation, giving more than 500,000 profoundly deaf partially restored hearing [37]. In order to gauge the effects of the implants, speech recognition tests are predominately used, at least in the development stages. The *Nijmegen Cochlear Implant Questionnaire* [25] is frequently used on a global scale to assess the benefits of cochlear implant surgery, but only 3 out of 60 items are related to music. The *Cochlear Implant Quality of Life* [37] survey tries to address this disparity by dedicating one category out of five to auditory entertainment evaluation. This does not mean that surveys focusing on music have not been developed, as the *Music-Related Quality of Life Measure* has been created to guide music rehabilitation for CI users [15]. This measure has first been recommended to be used in conjunction with the Nijmegen one by [20]. Nevertheless, when evaluated on monosyllabic words, common implant systems obtain 50–60% accuracy after 24 months of usage and nearly 100% on phrases [53]. Some individuals have astonishingly good outcomes, demonstrating what may be done with a neuroimplant in a cochlea that is otherwise completely malfunctioning. The results are improving, especially in individuals utilizing bilateral implants, although variability is still substantial, with standard deviations for different trials ranging from roughly 10% to 30% [53].

Unfortunately, the pitch and timbre perception offered by the current generation of CIs is drastically compromised. Another notable drawback is that users have trouble distinguishing between competing sounds when numerous instruments are playing or when there are extended reverberations present [10, 17]. Furthermore, even when signals are below the CI pitch saturation limit (300Hz), there is a weak representation of the fundamental frequencies (F0) for complex sounds, with difference limens ten times lower than normal hearing. Lastly, the dynamic range available for electrical stimulation in CI users is only about an eighth of what is available for listeners with normal hearing [55, 56], further compromising musical listening abilities, and by extension, musical appreciation. Since the general musical experience of CI users is subpar, these cumulative circumstances prevent the evaluation of music experience from being used as a gauge of the success of the implants.

F.2.3 Multisensory Integration

The foundation of the research presented in this paper is the multisensory integration principle, which describes how people combine information from several senses to create coherent experiences [50]. The prerequisite for this integration to be most prominent is that the stimuli overlap in time; this will result in a perceptual enhancement that is strongest for the stimuli that are least effective [50].

Recent research shows that multisensory integration can in fact occur at very early stages of cognition, leading to supra-additive integration of touch and hearing in the specific scenario of auditory–tactile stimuli [1, 18, 32]. This is particularly helpful for CI users who have been found to be stronger multimodal integrators, as documented by [45]. Additionally, research on auditory–tactile interactions has demonstrated that the presentation of a tactile signal can alter cross-modal perception and vice-versa [30, 41]. A popular example demonstrating the possibilities is listening to loud concerts that feel powerful partially because of the vibrations produced by the massive speakers, but listening equally loudly through headphones does not feel as exciting.

F.2.4 Vibrotactile Augmentation of Music

Previous research on the tactile augmentation of music has made substantial use of multisensory integration; in 2009, Karam et al. increased the audio-tactile resolution of the skin, drawing inspiration from earlier sensory substitution vocoders [31]. In accordance with the cochlea metaphor, which describes that lower frequencies are reproduced through lower body areas than higher frequency ones, their project produced a chair with four pairs of voice coil actuators set in an array along the backrest. According to them, each actuator was capable of reproducing an octave of the piano, from 27.5 Hz to 4186 Hz. They assessed their design in terms of emotional response and concluded that participants preferred audio-tactile stimulation strategies above the auditory signal alone [31]. A wide range of opinions, largely favorable, were expressed in response to further chair improvements [4].

With the assistance of the hearing-impaired community, Nanayakkara et al. created another chair installation [40]. Their haptic chair was designed to provide whole-body stimulation at first, with two contact speakers acting as haptic transducers placed beneath the armrest. For further iterations, actuators aimed at the lower back were added, alongside a vibrotactile footrest [39]. The tactile stimulation was always presented in conjunction with sound, reproducing an amplified version of the auditory input. Their chair has been successfully employed in longitudinal research (12–24 weeks) to improve music listening and speech therapy for deaf children, underscoring the need for training when users are expected to adjust to a novel haptic system [39].

A 2015 collaboration between Queen Mary University and the Deaf arts charity group *Incloodu* led to the creation of an installation in the form of a sofa and armchair [28]. The devices utilized a subPac[1] under the seat and voice coil actuators mounted in the backrests and armrests. The armrests recreated a noisy component associated with timber, while the spatial auditory information was distributed from low to high frequencies corresponding to sections of the backrest, similar to the cochlea metaphor described in [31]. A severely deaf architect who specialized in creating accessible furniture was employed to design the furniture. Their analysis demonstrates that the musical style has a significant influence on the experience, with highly rhythmic music evoking more favorable reactions than music where harmonic motion was most essential [28].

F.3 Materials and Methods

This section describes and discusses the participatory design process, as well as the multisensory concert and its evaluation.

F.3.1 Multisensory Integration Design Workshop

By introducing several multisensory installations, this study aimed to (1) involve CI users in the early stages of building novel audio-tactile displays and (2) analyze the constraints of the configurations that were shown. We conducted an exploratory study, gathering data using a triangulation of techniques, including observations, pre- and post-workshop interviews, and the think-aloud protocol [49].

Workshop Format

Each meeting had a set agenda and lasted between 60 and 120 min; one of the authors was present the entire time, taking notes and recording the conversations. Before the meeting, attendees were asked to complete an online survey that asked about their demographics and their past and present musical preferences. Before any installations were installed, a semi-structured interview was held with the goal of learning more about people's musical engagement patterns, and the survey results served as the basis for that interview. The participants were then led to investigate and experiment with the various installations mentioned in Section F.3.1. The session ended with a brief exit interview that summarized the participants' feedback. The attendees were encouraged to *think aloud* and were in constant communication with one or more of the authors during the entire conference.

Participants

Three people voluntarily took part in the study after being contacted by email or an open invitation posted to the Facebook group for the Danish Cochlear Implant users[2]. Participant 1 (P1) is a 52-year-old female who started losing her hearing at the age of 3, currently with no residual hearing. In 2017, she got bi-implanted with Kanso CI, experiencing a positive transition from hearing aids to cochlear implants. She likes *Fleetwood Mac*, *Dolly Parton*, and *The Beatles*, but dislikes techno, classical music, and heavy metal. She has a background in piano and dancing (in African and Danish dances). She sings in a choir but is challenged in distinguishing and synchronizing with accompaniment, misidentifying when to start singing. She reported using a water bottle or glass in her hands to feel the vibrations in the few occasions she attends concerts.

Participant 2 (P2) is a 69-year-old male with a genetic hearing disability who uses a cochlear implant in his left ear and a hearing aid in his right ear. He has experience from a musical family, singing in a church choir, and performing competitive dancing. He likes opera, waltzes, church, and classical music, but dislikes rock. More recently, he rarely listens to music. When listening to familiar music, he expresses: *"[...] my memory was another [...] I have this sort of feeling of something is in another way."*

Participant 3 (P3) is a 41-year-old male. He uses a Nucleus Cochlear implant in the right ear and is near deaf in the left ear, with a hearing threshold of +95dB. He has been using hearing aids since the age of 3, frequently upgrading them to higher amplification ones. When listening, he can identify when music is playing, the sex of the singer, and the instrument if the music is performed live on stage. He regularly attends festivals, mostly for social reasons. Lately, he enjoys listening to music for short periods of time (5 min), because after about 10 min it becomes exhausting. He mostly likes rock, especially the Danish band *Dizzy Mizz Lizzy*.

Workshop Experiences

Three different setups were presented to the participants, each focusing on unique approaches to music listening and augmentation.

Installation 1 Since CI users find it extremely difficult to distinguish specific musical instruments in a mixdown [42], we set up a multichannel listening environment. The installation used a four-channel speaker configuration to reproduce multichannel recordings in order to test CI users' instrument segregation processes; a diagram of the setup can be seen in Figure F.1. To compare and contrast the single and multi-instrument mixings, we invited listeners to freely walk about the space and hear various sound sources.

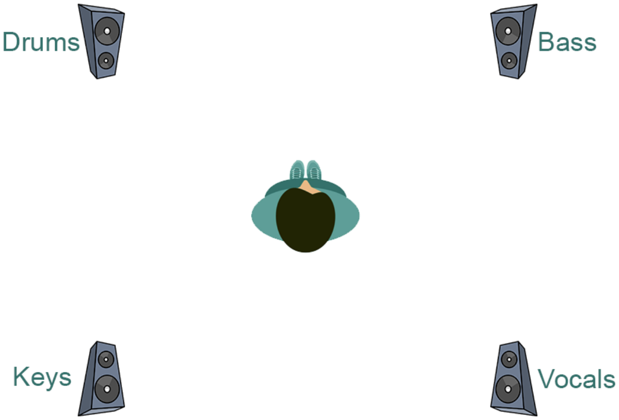


Fig. F.1: Diagram of installation 1.

To avoid having the loudspeaker directionality of the experiment be altered or diminished by room reverberation, the experiment was carried out in an anechoic room on the campus of Aalborg University, Copenhagen. Four *Dynaudio BM5 MKIII* loudspeakers were used in the configuration, which also included a *Steinberg UR44C* audio interface connecting the laptop to the speakers. Only one instrument per channel was used as a signal for each loudspeaker. Drums, bass, vocals, alternating piano or guitar, and multitrack recordings were reproduced using the Digital Audio Workstation *Reaper* to route the instruments to separate speakers. For all three sessions with the participants, no routing changes were applied to maintain consistency between the experiences. The amplitude level of each channel was set by the authors to obtain a balanced mix that allowed a normal-hearing person to perceive all the instruments with equal loudness in the center of the room. The single recordings were played without any effect, such as reverberation or compression, to avoid any possible confusion for the listener.

Upon entering the room, a brief explanation of the experience was given before allowing the participants to select their favorite song from three well-known rock, soul, and reggae tunes. Then, after reaching an agreement on the appropriate loudness level with the user, we set all channels to a *discussion level* throughout the whole test.

For the first portion of the experiment, the subject was instructed to attempt to identify the instruments being played and the speakers they were coming from while standing in the middle of the room. After making their choices, we requested the user to circle the room, moving in close proximity to each loudspeaker, to confirm or retract his/her statement regarding the types of instruments played and the locations. In the second and final section, we let the subject choose a location in the space where the music sounded the best to them. The participant was free to comment or explain at any time during the entire experiment what they were thinking and how they felt about the listening situation and were encouraged to do so with guiding questions from the authors.

Installation 2 In order to improve the music-listening experience for CI users, we started a design process to investigate whether and how audio-tactile feedback might be included in a sitting installation. We concentrated on delivering low-frequency augmentation due to the poor auditory resolution in the range that CI users experience [5].

With three CI users and three normal-hearing participants, we tested two prototypes that included three parts: a seat, a footrest, and a hand-held gadget. These components could be used both separately and simultaneously. Through a headphone splitter that fed the same signal to the amplifier for each transducer, all users had access to the gain control for each actuator. Only the first user made the decision to change the gain balance herself; the other two gave the researchers verbal directions. For the first user, the music was played over a pair of *B&W 800D* speakers, and for the second and third users, a pair of *Mackie SRM450 + SRM1550* speakers. Users had access to the master volume knob, which coupled the tactile and auditory volumes, as well as the signal feeding the headphone amplifier, which served as a multi-channel signal splitter in this case.

Three different sitting installation types offered three distinct experiences. The first installation was a tactile vehicle seat that was activated by a *Buttkicker LFE*[3] and powered initially by a *Buttkicker BKA1000* and then by a *Stage-Line ST600* in bridge mode, since we discovered that higher frequencies were limited by the *BKA1000* amplifier. The actuator was placed behind the car seat, which was fastened to a wooden EUR-pallet platform using bolts, as shown in Figure F.2. The headrest, as well as the backrest, received adequate tactile feedback from the actuator, even though it was not in direct contact with it. Participants P1 and P3 noted that high tactile amplitude might easily become overwhelming in this setup.

Based on P1 and P3, which assessed the first design as potentially overwhelming, the second and third types of sitting experiences changed from a

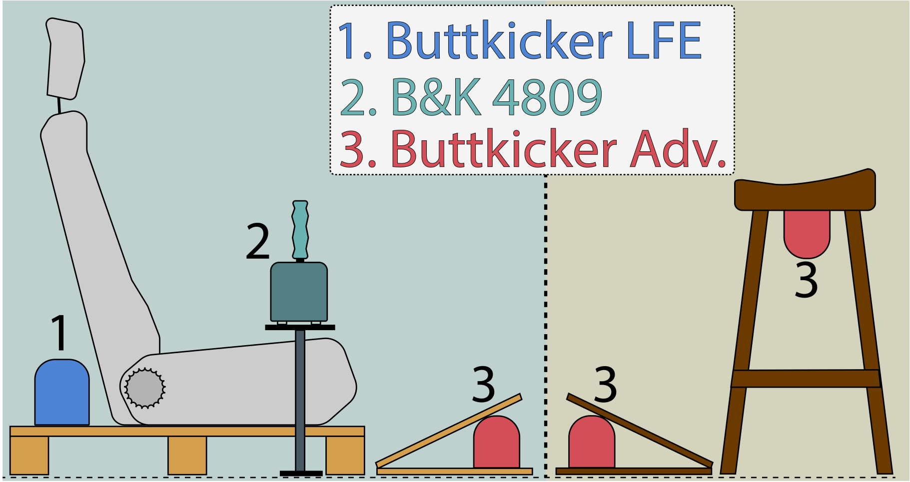


Fig. F.2: Example configurations experienced during the workshop.

low seating position to a more upright one through a bar stool instead of the car seat. We chose the bar stool design because it gives users a choice over how much body weight they apply, which is coupled with the tactile sensation received. The bar stool was activated by a *Buttkicker Advanced* driven by a *Buttkicker BKA300*. This arrangement used less power because the actuator was significantly smaller than the one from the vehicle seat but still provided enough tactile stimulation. The frequency responses of the two setups were significantly different, with a high-frequency focus for the arrangement with *Buttkicker Advanced*, as noted by the authors.

Different configurations were present for P2 and P3, with P2's actuator bolted perpendicular to the seating area and P3's actuator fixed parallel to the ground on the side of the seating. In response to P2's observation about the bar stool's lack of intensity (possibly in comparison with the car seat), the side actuator arrangement sought to conduct more tactile stimulation. We saw an unusual behavior in P3's setup, where strong transients caused the bar stool to shake laterally, maybe as a result of the loose joints, and it seemed like a slight *kick in the back of the chair.*

The footrest has a 22° inclined plane that was created in accordance with H. Dreyfuss' measuring standards and guidelines [51]. The footrest's dimensions of 45 cm in length and 60 cm in width were necessary to allow room for the actuator underneath the inclined plane and to accommodate two adult feet on it. The same *Buttkicker Advanced* + *BKA300* combination as for the bar stool was employed. In accordance with the seat setup shown in Figure F.2, the transducer was bolted beneath the footrest. The setting was the same for each participant.

There were two handheld devices. By stacking 51 laser-cut slices of 4 mm

HDF, a cylindrical handheld grip measuring 2040 mm in length and 110 mm in diameter was created. These measurements fall within the guidelines for design provided by H. Dryeyfuss [51], which are between sizes 3 and 4 on a tennis racket, representing the middle sizes for adults. As shown in Figure F.2, this grip was fastened to a *Brüel & Kjaer Type 4809* portable vibration exciter. The second interface, called VAM (Vibrotactile Actuator for Music), was created around the *Tactuator BM1C*[4], an ovoid device with dimensions of 84 mm in width, 58 mm in height, and 89 mm in depth, described thoroughly in [43]. P1 tested the cylindrical grip and the VAM in conjunction with the seat and footrest on their own, but P2 and P3 abandoned the VAM because they felt it *"didn't add much."*

When it comes to music, the initial auditory stimulus was *Peggy Lee's Fever*, due to its distinct instrumentation and the importance of the female vocal track, matching the preferences of CI users as presented in [7]. This opening track was played in each configuration of Installation 2. Firstly, two distinct signals were transmitted in succession to the handheld grips; the first was congruent to the one offered through other actuators, but the second was band-pass-filtered to isolate the female vocal range and pitched one octave lower to correspond to the skin's sensitivity range [29, 52]. This only applied to P1's experiment, as she claimed that the pitch-shifting process decouples the voice from the tactile sensation. P2 and P3 were exposed to a solo bass improvisation on the double bass or ukulele bass of the same song. The performance featured a variety of playing techniques (pizzicato, slapping, staccato, etc.), making use of the entire instrument's range and using bar stools and vehicle seats as props. P3's experience featured a section with live drums as well.

Participants chose additional audio content for their ideal arrangement. All three of them favored the configuration shown in Figure F.2. P1 listened to *Fleetwood Mac—Dreams*, P2 *Vienna Philharmonic—An der schönen, blauen Donau* (excerpts), and P3 *Dizzy Mizz Lizzy—Silverflame*.

Installation 3 With the aid of movement-based performance and in-air haptics, we also looked at the participants' experiences with embodied interactions. We reviewed passages from the inclusive performance research *Felt Sound* [9], which was created for both hearing and D/HoH audience members. Initially, *Felt Sound* was performed live with an 8-subwoofer speaker configuration while the participants sat near to or touched the speakers. It consisted of a digital musical instrument and a performance environment. Due to constraints on time and space, as well as COVID-19 safety measures, performance snippets were presented to each participant separately, with two subwoofers positioned such that they were facing them and surrounding their seating area. Participants were still urged to engage with the speakers and use their hands to physically feel the vibrations.

We gave a succinct overview of the inspiration, idea, and performance style of *Felt Sound*. We explained its context to the participants before presenting its extracts. After the performance, we talked about their interactions with *Felt Sound*, as well as their own musical and dance routines. Participants shared their own associations and experiences with movement practice and music as a result of the presentation of a novel movement-based musical concept.

Experiences and Results

Each participant's conversation with us was audio-recorded, and following the study, we transcribed them. An overview of these discussions, emphasizing their enjoyment of the installations and their overall experiences, is provided in this chapter.

First Participant As described in the previous section, P1 experienced four vibrotactile displays, including a car seat, footrest, hand grip, and the VAM in two situations (processed and unprocessed signals). We initially adjusted the listening volume to "comfortably loud", just above conversation level, and the separate tactile amplitudes to "perceptually equal"; the tactile amplitude was coupled to the audio loudness through the *"master out volume"* knob on the sound card.

In the first instance (listening to the processed music), she stated that it was *"fun to feel the vibrations throughout the entire body"*, reiterating her experience with the water bottle during concerts (see Section F.3.1). She did not comprehend how the vocals were mapped to the haptic feedback because she could already hear the voice through the speakers and did not require additional stimulus representing the voices. Moreover, she barely raised the volume of the hand grip a few times.

She appeared more engrossed in the song and moved to the music when she was given the second case (listening to the raw audio). Similar to the first instance, she tried slightly turning up the seat, footrest, and hand grip. When the music was finished, she said she favored this listening method over the first case, because she could feel the melody in the footrest. She added that it could occasionally be overpowering to listen to the vibrations through the chair in this configuration.

Over the course of the experiment, her perspective evolved. She said it was enjoyable and that she could feel the voices through the hand grip and the bass line—which she initially thought was coming from a keyboard—through the foot pad and the seat. The same signal was replicated by all actuators, but different haptic sensations were felt at various points on the body, amplifying the notion of pitch and instrument type. In response to the question of whether she would use a textit-type gadget at a concert, she said, "I would want to have some help from vibrations", and she went on to explain how she usually sits quite near to the speaker at live shows in order to receive the haptic feedback. Additionally, she claimed that if *"it is trustworthy"* she would utilize them. In comparison with the other haptic listening instruments, she placed less emphasis on her experience using the VAM, claiming that it was insufficient. We took her comments concerning the VAM to mean that it was *"not strong, comparing to other actuators"*.

She recounted her experience with music and movement after hearing Installation 3. She was inspired by this installation to describe her physical routine and more embodied musical experiences, including singing. She stated that she finds it challenging to recognize the beginnings when she sings in a choir, specifically being able to tell when to begin singing solely by listening to the piano. She expressed interest in using gestures during her vocal practice to help her and reinforce her conductor's support for her. She also mentioned how watching a gesture-based performance helped her to understand and appreciate music.

Second Participant P2 experienced Installation 1 with the song *Ain't No Mountain High Enough by Marvin Gaye and Tammi Terrell*. He properly recognized the left and right channels and the instrument they played while listening from the center of the room. However, he misidentified the instruments at each channel. When we invited him to get closer to each speaker, he was able to correctly identify all the instruments, including the alternating male and female voice, but was unable to tell the words apart. He recognized that the guitar and keyboards were playing simultaneously on the same channel, just like the voices. He said that *"it's always about guessing,"* which showed how uncertain he was of his responses. He always used his hearing aid mostly and pointed his nonimplanted ear in the direction of the speakers.

Finally, we let him freely choose a location where the music suited him the most. He positioned himself closer to the 1–2 speakers in the center of the 1–2 3–4 speaker pair, saying, *"I think this must be the perfect (place) for this kind of music because it is feasible to hear all of it."* He claimed that after being introduced to each instrument independently, they all started to make more sense to him. Similar to that, after being told what the chorus lyrics were, he could follow them while identifying the words—confirming the pop-up effect.

With the same volume controls as those first established for P1, the second installation included the car seat, the bar stool (with vertically mounted actuator), the footrest, and the hand grip designed around the *B&K* actuator. A bar stool with a footrest and a hand grip was one part of the setup, and the vehicle seat with a footrest and a hand grip was another. Without using any processing for the actuators, we played the same music as for P1. We interrupted the listening for intermediate conversation after about 90 s of listening through the initial setup, and the subject stated where he felt the vibrations most strongly: in the thigh, ankles, and up to the elbow. Although he was limited in his comments regarding the perceptual quality of the stimuli, he was verbose about the locations and intensities of the reported vibrations. He recognized the deep bass and the female voice. Lastly, he added that he could recognize the melody with ease.

When we asked him how the music made him feel, he replied: *"It was more like a little bit sad music."* He then went on to say that the music needed to be happier for him to like it. Additionally, P2 acknowledged that he dislikes the bass as an instrument but loved the live music part when one of the researchers played a double bass solo. He added that the installation was more engaging but that the music selection had an impact on his experience, because it was not in his preferred musical genre, enhancing something he did not like. He asked to hear *J. Strauss'* composition *"An der schönen, blauen Donau."* The participant commented, *"Yeah, this is much better, much better, yeah, and I can feel it (the vibrotactile system) supports the song,"* after the very first chord and continued *[...] if you enjoy the music, this gives extra power.* The second configuration was limited to playing waltz music; however, the conversation shifted to the economic worth of musical experiences and no comments on the second setup were made. He continued to describe himself as fairly conservative, and he would prefer a traditional chair unless specifically encouraged to try one in a concert hall, but otherwise, he would not utilize such a system (setup 2) in a concert context.

Regarding Installation 3, he had a different reaction compared with P1; the music's low frequency content did not appeal to him as much. Although he said that he could physically feel the vibrations, this style of listening did not improve his enjoyment of the music, but he concluded by saying that the music's gestural performance element worked well.

Third Participant In Installation 1, P3 chose to hear *"Don't stop me now"* by *"Queen"*. He recognized the voice by positioning himself in the center of the room and noted that the speaker playing the bass line had a much lower volume than the others. He soon recognized the voice after moving closer to each speaker, but he mistook the piano for the guitar. He had trouble recognizing the instrument from the speaker playing the bass line and inquired as to whether it was a tuba. He identified the drums accurately.

When we asked him to pick a favorite spot in the room, he spent several minutes walking between the speakers and the general listening area. The location he chose was far from speakers 1 and 2 but much closer to speakers 3 and 4, which he was facing—these speakers were playing voice and guitar, respectively, as seen in Figure F.1. He mentioned that he could hear *"a little bit of everything"* at this point, focusing solely on the vocals, bass, and drums. We saw that he preferred listening to the instruments on their own during the postexperiment talks, since he could interpret them in his own way. Discussing the bass sound at concerts further, he said, *"[he] could never discern [the individual instruments] since everything sounds like "mush"*, but in this case, after hearing each instrument independently, *"[...] it was a bit simpler"*.

The only change between P2 and P3 regarding the second installation was how the actuator was positioned on the bar stool, as explained in Section F.3.1. Once the music had stopped after about 90 s (before the second verse), the participant quickly remarked that the hand grip and the bar stool did not offer much to the experience. When pressed, he was unable to specify the song's valence. Following this, all actuators were muted before the music started again, and we gradually increased each one's amplitude, while instructing the participant to place their attention on preferred locations rather than activated ones; the outcomes were the same: he preferred the hand grip and the footrest (especially when it was turned up more). He claimed that while on one side of the song it is *"slow and heavy"*, the singing (voice) sounds cheerful, thus it is impossible to pin down the song's atmosphere.

All actuators were set to their initial amplitudes for the live ukulele bass performance. P3 provided comments, stating that he favored the footrest for lower frequencies but preferred the hand grip for higher frequencies. Additionally, he noted that it was simpler to *"feel what happens"* through the hand grip when short, quick notes were played. Similar to the first inquiry, the bar stool *"did not have anything to give"*.

Moving on to the setup with the actuated car chair, the participants remarked that setup 1 felt a little distant by comparison and that *"[...] this is much better to have it in the back, this way"*. The volume of the actuators was altered during this experience, and we came to the conclusion that it works best when all three actuators are perceivable and that it seems *"empty"* when the car seat is not being actuated. After the live bass performance (same as for setup 1), P3 said that while utilizing the setup is enjoyable, he still feels like he is *"missing something"* and *"just misses genuinely being able to enjoy music"*, which is a truth that was unaffected by the setup that was demonstrated. However, similar to setup 1, he could *"feel"* the voice more through the hand grip.

When asked which he liked more, the recorded or the live performance, he responded that the latter is better since it had more instruments and "more varied sounds." As a result, two of the researchers spontaneously performed a drum and bass duo set while grooving out to the bass line from *Fever*. The participant mentioned that before experiencing this, he did not know how to tell the difference between the bass sound and the drum sound (in live performances), but through this multisensory experience he could identify both.

His encounter with Installation 3 was consistent with P1's observations regarding the gestural performance. He claimed that he had never witnessed a musical performance in which the audience participated by moving their bodies to the music.

Workshop Discussion

Although CI users may share common difficulties in experiencing music, such as pitch identification, source localization, and instrument segregation (auditory streaming), their priorities in addressing these obstacles vary greatly from person to person. P1 was able to hear the nuances of pitch variations in singing, but due to her musical training, she prioritizes practicing onset detection and phrasing to support her choir singing. Similarly, P2 preferred limiting his experience to the musical genres he enjoys and enhancing these specific genres over practicing to fill the gaps in his musical perception. Not only should researchers and designers consider individual differences in hearing profiles but also in musical appreciation, engagement, and preferences. Age, the use of hearing aids, and musical training, among many others, have a substantial impact on these design considerations when working with CI users. Therefore, due to the variation in CI users' perception, experience, and comprehension of everyday sounds, speech, and music, on top of the ones mentioned above, we believe that researchers designing multisensory experiences should focus on personalization and customization as core design requirements.

Similarly, for many CI users, the experience of music is novel and requires ongoing practice and education. A continuous musical activity in which users can practice becomes essential for those who struggle to comprehend music. Our assessment methodology reflects this process of investigating and comprehending CI users' hearing and engaging with technology in ways that support both their hearing development and appreciation of music. Their participation in generating ideas and directing design directions was essential to the research procedure.

We routinely consulted the most recent research on evaluating CI hearing and used it to guide our experience design research, because the references CI users use for music are more subjective when they describe their perception and experience of music. We feel that a more comprehensive approach to facilitating the music engagement of CI users allows more embodied listening and music-making practices. Rather than separating design, assessment, and evaluation processes, developing new musical interaction experiences leads to an integrated and participatory research approach. In addition, we discovered that this process-oriented evaluation enables designers to find more options for engagement with CI users, as participant recruitment in the CI community remains one of the greatest obstacles. We believe that by establishing a more formal organization around cochlear implant use and music, their engagement in design and research can be encouraged, thereby enriching their musical experiences.

It is advantageous for designers of tactile displays for CI users to create devices that are adaptable and can accommodate varying musical preferences, hearing ability, and musical engagement levels. Although the small size of our sample prevents us from generalizing the experience of CI users in the larger community, the vastly varying needs of each participant emphasize the necessity for design flexibility and adaptability. In addition, great care should be made to avoid creating unpleasant experiences, as was temporarily the case for P1 (intense tactile stimulation) and P2 (unpleasant music choice). Preparation can be aided by prior knowledge of the target audience, but a crucial step towards this prestudy is ensuring that displays have basic controls for tactile and aural stimulus levels and, in the case of multiactuator devices, that each transducer has independent control. Consider flexible or modular hardware that is easily reconfigurable based on the demands of the user. Participatory design practices allow researchers to investigate individual needs. Lastly, we recommend, whenever possible, the incorporation of visual feedback in the form of gestural or movement-based performance or visualization, which can supplement gaps in perception from the tactile or auditory channel. Listening to music must be viewed as a multifaceted experience that might be difficult and laborious for the hearing impaired, and great care should be taken to insure the comfort of the listener.

F.3.2 Concert

This section describes the design, implementation, and initial evaluation of vibrotactile concert furniture, aiming to improve the live music experience of cochlear implant (CI) users. The system created was a direct result of the workshop described in Section F.3.1 and was evaluated in a concert scenario (drums, bass, and female vocals) at the Royal Danish Academy of Music. The project aimed to create a better live music experience for CI users by providing a multisensory concert designed with CI limitations in mind but not excluding normal-hearing individuals from participating in the event.

Tactile Displays

For the concert, the hand grip was used, and a new tactile display was designed following the guidelines presented in Section F.3.1, which were distilled into the following design objectives:

- Enhance the concert experience by providing congruent vibrotactile feedback.
- Afford multiple interaction modes and postures to accommodate the variate needs of CI users.
- Present as furniture rather than medical apparatus.
- Encourage a social experience.
- Usable by CI users and normal-hearing participants alike.

Vibrotactile Furniture The main tactile display was derived from the bar stool/car chair and footrest setups presented in the workshop Section F.3.1 but with several bespoke elements. Firstly, in order to accommodate individual vibrotactile preferences, we designed a double-slanted system formed by a leaning bench and a footrest, as shown in Figure F.3. This configuration allows users to adjust the amount of stimulation they desire by altering their weight distribution, affording control analog to tactile mixing between the feet, hands, and buttocks. Another intention of the double-slanted design was to discourage users from resting in a fixed position, in order to avoid overwhelming experiences expressed when sitting in the car chair, as discussed in Section F.3.1. Neither the leaning bench nor the footrest is ergonomically viable unless used together, forcing the user to find an individual balance between the two, both in terms of comfort and vibrotactile preference.

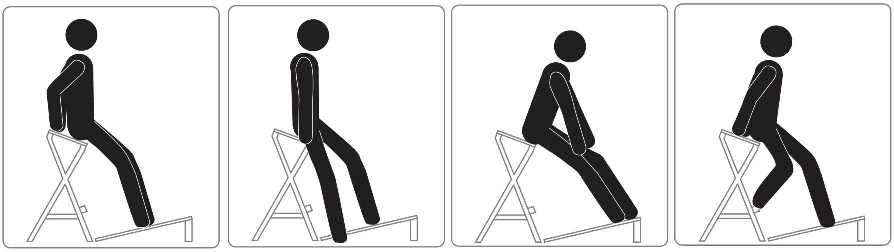


Fig. F.3: Different posture suggestions engraved on the footrests

The height and angle of the bench were designed with a single goal in mind: to ensure a comfortable posture for individuals with different statures. The oversized footrest plays an especially important role, as both short and tall individuals can find a position on it that matches users' individual preferences.

Two benches measuring 120 × 30 × 75 cm and four footrests of 38 × 73 × 18 cm were built, using common woodworking materials: beams of 38 × 76 cm for the bench legs, the beam connecting the footrests, and beams of 38 × 38 cm or 14 × 76 cm for structural reinforcements. The top of the benches were cut out of 17 mm high-density fiber board and were covered with a thin carpet-like material in order to avoid slipping. The footrests were cut out of 15 mm-thick pine wood, and each one was engraved with Figure F.3 in order to encourage exploration of different postures. Figure F.4 shows the leaning benches and footrests as they were used at the concert. All contact points with the ground had rubber feet of 2 × 2 × 2 cm in order to decouple the system from the floor and to avoid slipping—a behavior noticed especially in the case of the footrests. Each of the benches was actuated by a *ButtKicker Mini Concert* bolted in the middle, underneath the sitting area, while the footrests had a single *ButtKicker LFE* attached; this meant that the outer ones produced slightly less vibrotactile stimulation.

Audio and Tactile Signals Similar to the setup used in the workshop, the tactile signals were split between the leaning benches and the footrests. The benches were reproducing the signal captured by the vocalist's microphone, and the footrest played the signal from the double bass, as seen in Figure F.5. The hand grip mixed both the bass and vocal signals. The drum sounds were conveyed only acoustically, as it was assumed that the timekeeping

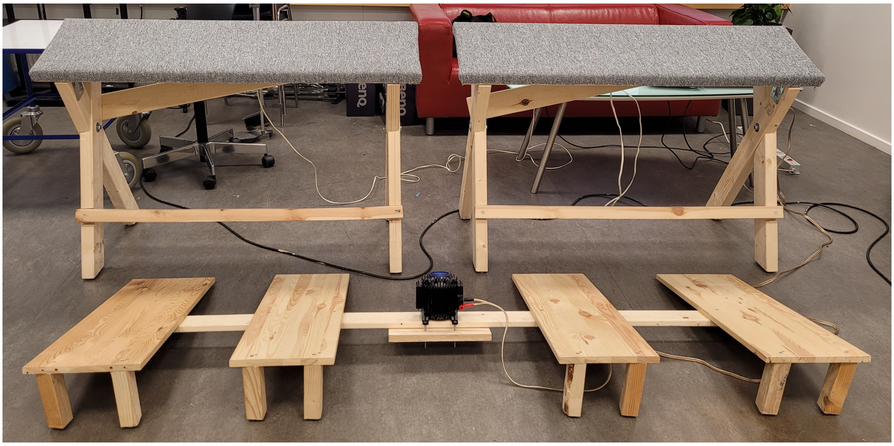


Fig. F.4: Leaning benches and footrests.

and rhythm aspects of drum-playing would be sufficiently received through this channel alone [36]. The levels for the tactile displays were calibrated during sound check, based on the discussions from the workshop in Section F.3.1, with help from a professional Tonmeister, with the aim to provide a comfortable balance between acoustic and vibrotactile stimulation.

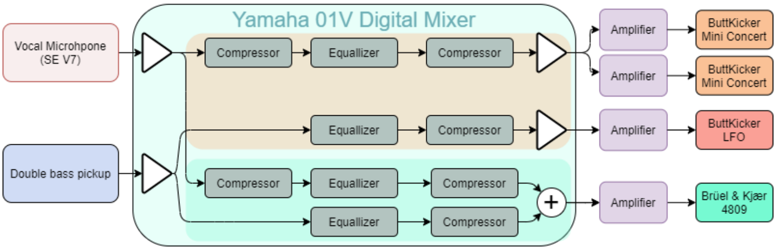


Fig. F.5: Tactile stimulation signal chain.

The main goal for the signal processing was to present stimulation when the vocal/bass signal was present acoustically, as well as to prevent the tactile stimulation from becoming overwhelming; this required a substantial amount of compression/limiting and equalization, all carried out inside the Yamaha 01V96i digital mixer. The settings for the effects were preset during one of the rehearsals, with minor adjustments during the soundcheck.

Concert Setup

The concert was organized in collaboration with the *Tonmeister department* from *The Royal Danish Academy* and the *Center for Hering and Balance* from

Rigshospitalet—Copenhagen, as the main act of a workshop showcasing assistive hearing technologies called *CoolHear Workshop*[5]. Participants were recruited from the entire spectrum of hearing characteristics, as well as professional personnel working with CI and hearing-impaired individuals such as audiologists, logopedists, technicians, and artists. The event was free, was promoted locally through social networks and individual invitations, and gathered more than 50 attendees, of which at least 6 were CI users. The main act took place in Studio 3 of *The Royal Danish Academy*. During the concert introduction, the concept was described to the audience, and the multisensory setup was introduced. The tilted benches were placed among the regular seats in order to avoid the feeling of exclusion; Figure F.6 shows the tilted benches in use during the concert. Participants were asked to prioritize individuals with CIs and suggested that they take turns interacting with the vibrotactile furniture.

Fig. F.6: Vibrotactile furniture in use during the concert

Music Three professional musicians were employed to perform on the stage, playing drums, double bass (3/4 size), and female vocals (mezzo-soprano). The entire band was part of the concert design from the beginning, and two of them are among the co-authors who participated in the design workshop as well. This early relationship resulted in great communication regarding the needs and challenges manifested by CI users, and it was especially beneficial that the same bass player and drummer were employed who were present throughout the participatory design workshop described in Section F.3.1.

The setlist for the concert was collectively decided upon based on factors that surfaced throughout the first workshop Section F.3.1 and some requirements imposed by the multisensory hardware setup, as well as the instruments available:

- The music should be popular in order to increase chances of cognoscibility;
- The music should include lyrics and should be focused on the vocal line;
- The music should revolve around simple, repetitive riffs;
- The melody should be easily represented by a single instrument;
- The music should be easily transposable to the vocalist's lowest register;
- The setlist should contain songs from multiple genres.

Based on the items above, several songs were trialed during rehearsal sessions, and the setlist consisted of two sets of two songs each. The decision to split the concert into two sections was taken in order to avoid auditory fatigue, usually associated with the CI music listening experience. The final setlist was based around the following songs, adapted for bass, drums, and vocals:

- Peggy Lee—Fever;
- The Civil Wars—Billie Jean;
- Louis Armstrong—What a Wonderful World;
- Postmodern Jukebox—Seven Nation Army;
- J. Knight, D. Farrant, and H. Sanderson—Don't Give Up On Love (improvisation);
- Ruth Brown—5-10-15 Hours (improvisation);
- Etta James—Something's got a hold on me (backup song—not used).

During a casual conversation in the break in between the two sections of the concert, some of the CI users that tried the multisensory furniture expressed their appreciation for the experience and requested more songs played in the second part. The musicians agreed to play an encore after receiving a standing ovation at the end of the second set, resulting in an impromptu adaptation of the last two songs on the list presented above. The interpretation for these songs was more improvised, promoting solo passages for drums and bass, respectively. The cumulated play time added up to a little over 25 min.

F.3.3 Concert Experience Evaluation

The goal for the evaluation was to understand the perceived quality of the multisensory concert experience and to assess the usability of the vibrotactile systems. Throughout the concert, more than 20 participants tried multisensory furniture, and a similar number of people experienced the hand grip. Of these, five participants were cochlear implanted, and only those are included in the results.

Methodology The methodology implied a triangulation of qualitative methods: video-based observations, postexperience semistructured interviews, and a later-written structured interview [6]. Due to the challenges presented by CIs when it comes to speech in noisy environments, the semistructured interview transformed into a focus group, where one participant relayed the opinions of others while communicating through sign language with their peers. The fundamental method to synthesize results from the different data-gathering methods was *Meaning Condensation*, as described in [6].

Concert Results

Focus Group Once the concert was completed, the CI users were invited to answer questions regarding their experience and their preferences. The interview was recorded and transcribed, but due to the aforementioned issue with noise in the room that resulted in participants resorting to using sign language, only a limited amount of data were obtained through this method and are treated as coming from a collective participant.

The participant group mentioned that the concert was *"very exciting and fun"*, that they could *"almost hear the voice"*, and jokingly continued that they would have liked a whole hour concert instead of the limited time allocated. Furthermore, they collectively expressed their preference for the vibrotactile furniture over the hand grip, arguing that the latter's signal is too complex, and it is much easier to understand distinct source streams, as was the case for the tilted benches and footrests. This setup allowed participants to witness an experience of *"loud bass"* without saturating the compressor in the CI, and participants mentioned that they usually have problems in this regard when attending regular concerts. Nevertheless, one CI user complained that they could not feel the drums, and when explained that the drums were only acoustical, they expressed their wish for this signal to be represented as vibrotactile as well. Lastly, the same participant mentioned that once they felt the tactile stimulation generated by the singer's laughter (in between songs), they could *"tune in"* and better understand the vibrotactile sensation.

Towards the end of the focus group, the participants mentioned that the balance between the instruments' volume, both acoustic and tactile, was pleasant, and it made sense to them, but complained that sometimes they could feel the stimulation from the double bass in the benches, which was supposed to process only the vocal signal. This was an unfortunate effect due to the stage being small, so the bass sounds were picked up by the vocalist's microphone. On top of this bleed, the generous amount of compression and makeup gain only exacerbated the unwanted signal in the tilted benches. Lastly, the participant group collectively mentioned that, towards the end (after the encore), the tilted benches became slightly overwhelming. This was the result of an error we made when bringing back the tactile volume after the encore; while the person in charge of levels was confident that the mixing faders were at the same level, the participants reported otherwise, further emphasizing the importance of a good balance and mix between all types of signals.

Video-Based Observations The entire concert was recorded using a Zoom Q8 video camera[6] mounted on a tripod and aimed at the two installations; a screenshot from the recording can be seen in Figure F.7. The observation analysis focused on reporting the interaction with the vibrotactile furniture and the hand grip during the concert.

Once the music started, the four CI participants using vibrotactile furniture at the time signified a good reception of rhythm by starting to tap their feet or nod their heads, while experimenting frequently with posture adjustments, rotation of positions between them, and making room for others to use the tilted benches. Throughout the event, the CI users used sign language to communicate with one another, resulting in a great deal of interaction and laughter. The hand grip was explored similarly but did not elicit the same head bobbing, or foot tapping, and the interaction with it was shorter than the tilted benches; in the beginning, it was several seconds long, but as the concert evolved, some CI participants spent minutes at a time interacting with it. Nevertheless, while the tilted benches were in use 100% of the time, the hand grip was not.

During the second song, participants slowed down the frequent adjustments and seemed to focus more on the music, probably because they had already found some comfortable postures, but they continued to communicate through sign language. One participant was particularly interested in the entire system—exploring different postures and interaction methods, including grabbing both benches, the hand grip, and a bench concomitantly, touching the bench and the footrest at the same time (as seen in Figure F.7, and even covering their ears. Two participants even experimented with the vibrotactile furniture barefooted.

As the concert progressed, the interaction with the hand grip slowed down even more, but the tilted benches were constantly in use, with more and more time being spent on them. This is probably because the novelty effect was slowly wearing off, and they could focus on the music. Towards the end, two participants in particular (that seemed to be friends) experienced the last three songs on the benches, without interruption. On the other hand, one participant sat right next to the benches, constantly touching them with

Fig. F.7: Exploratory interaction with the tilted benches and footrest.

her hand, while still engaging with her peers through sign language.

Online Interview In order to validate some of the data gathered through the interviews and video-based observations, an online written interview was sent to the CI participants via email, 10 days after the event. In addition to validation, these interviews aimed to verify the midterm impression of the concert, the attitude towards it, and to receive more feedback that could have been overlooked.

The interview consisted of six questions:

1. What was it like attending the concert with the vibrating furniture?
2. How was it feeling to sit on the furniture during the concert?
3. How did you experience the connection between music and vibrations?
4. Which instrument did you prefer and why?
5. Would you recommend this type of concert furniture to other CI users?
6. Do you have any other comments?

Two participants responded to the interview, reiterating that the experience was *fun* and *nothing like I have tried before*, adding that one could almost *hear better and sense it (the music)*. Furthermore, both participant respondents mentioned that they preferred the bass instrument, as it created a good sense of rhythm, and both mentioned that they would have liked to have the drums reproduced through the vibrotactile furniture. On the flip side, they both remembered the microphone bleed as a problem and concluded that the hand grip did not add anything to the experience. One respondent mentioned that the vibrotactile furniture became a bit overwhelming at some point, while the other reported that it was *strong, but not uncomfortable*. Lastly, a respondent said that the furniture might take up too much space but appreciated the fact that one can get a good experience by interacting with their hands as well. Both of them concluded that they would recommend this type of concert to others, and one extended the idea to include visual support that would indicate which instruments are being represented through the tactile modality.

Limitations By organizing an open and public event, we aimed to increase the validity by ensuring an ecological evaluation, but this came with its own challenges and limitations. First and foremost, it was hard to estimate the number of participants who would be using a CI or have a good understanding of each participant's hearing characteristics. We tried to promote the event through the relevant channels, targeting the CI community directly, but only about 10% of the participants were in our target group. In addition, it is safe to assume that the CI users that decided to participate in the event are those seeking musical experiences, resulting in some degree of positive bias towards the experience. As an additional limitation, the participants told us that other CI users from their network were interested in attending but a bit frightened by the fact that the event was running in English. This is an aspect to take into account for future iterations.

Another limitation regarding the data collection method was the assumption that we would be able to interview the CI users present in the same environment. As mentioned above, the noise floor in the room was high enough to discourage some CI users from talking, instead resorting to relaying their input through sign language via their peers. This limitation could have easily been avoided if a dedicated and silent room had been prepared a priori.

Lastly, because the participants were encouraged to move around when interacting with the vibrotactile systems, there was a considerable amount of occlusion in the video recordings, preventing us from observing all the details of their interaction. A multicamera setup could have solved this issue, but unfortunately, we did not prepare for that scenario.

F.4 Discussion

Combining the results from the three different data-gathering methods allows us to draw preliminary conclusions regarding the experience of the

CI users. First and foremost, it is important to highlight the fact that the multisensory concert was well-received, as can be deduced from analyzing all three data channels independently. However, this finding should not be generalized, as our participant pool was limited and, as mentioned in Section F.3.3, CI users who attended could be positively biased towards openly searching for musical experiences in general. Nevertheless, the data we collected strongly indicate that a multisensory concert experience is appreciated, especially for the CI users who avoid concerts due to the poor musical reception the device produces. One of the most important factors that contributed to this is the flexibility the vibrotactile furniture provided; as observed from the video recording, participants experimented with a multitude of postures in the beginning, but as the concert progressed, they settled into the seemingly most comfortable one, using their multiple points of contact. Out of these, it seems that the hands are preferred for sensing vibrotactile stimulation, especially when dealing with mid–high frequency content such as voice. Another appreciated factor in the multisensory concert is the ability to experience loud music, generally associated with concerts, without fatigue. As mentioned by [54], the dynamic range for electrical stimulation is frequency-dependent but always lies between 10–30dB, as opposed to approx. 120dB for a normal-hearing individual. This reduced dynamic range can be perceptually extended by augmenting the auditory signal with vibrotactile stimulation [2], thus allowing us to create experiences that feel loud. Since bass frequencies usually carry the most energy, participants associated the increased loudness with this frequency range without being overwhelming. Lastly, it was positive to acknowledge that participants could associate and integrate the vibrotactile stimuli generated by the vocalist with the electric stimulation provided by the CI. We expect this integration to enhance with a longer exposure time, resulting in better multisensory concert experiences.

An important factor worth discussing is the importance of multisensory processing and mixing. It is paramount to use limiting amplifiers in order to avoid tactile sensations from becoming overwhelming but at the same time high enough to be perceived. As mentioned by [27], the dynamic range for tactile stimulation reception is fairly limited, between 36 and 47 dB, depending on the area, further emphasizing the importance of strict dynamic processing and mixing. Our observations fall in line with [27] regarding the difference in body-area-dependent comfort levels at the top end of vibrotactile amplitude, as observed when one participant decided to sit next to vibrotactile furniture while still holding her hand on the benches. Furthermore, both in the postexperience interview and in the online interview, it was mentioned that vibrotactile stimulation did become overwhelming. This was pointed out as happening in the encore stages of the concert after the mixing engineer had reset the tactile amplitude to supposedly the same level as during the sound check, indicating that the range between pleasant and uncomfortable is very small.

Since the vibrotactile furniture was used considerably more than the hand grip, even though the participants tried extensive hand interaction with the benches, it is safe to conclude that the problem with the latter one is with the signals it reproduced. We are confident to suggest that mixing auditory streams into one tactile display is not an efficient technique, especially since the hand grip was the preferred display in the participatory design workshop Section F.3.1 when it was reproducing vocal stimulation alone. The importance of this one-to-one mapping is also expressed throughout the two interviews, when participants complained that they could *"feel"* the double bass signal through the benches, which were supposed to reproduce only the vocal signal. It was also interesting to observe one participant that tried to feel the two vibrotactile actuators, hand grip and chair, at the same time to compare the vibrations.

As the concert was part of a larger workshop, many hearing-aid users and children with cochlear implants attended as well. While these two groups are not the focus of this study, it is important to note that hearing-aid users found the vibrotactile stimulation distracting and potentially annoying, resulting in a much shorter interaction with the systems, while children avoided them entirely. Young CI users, on the other hand, were more enthusiastic about virtual reality experiences and computer/tablet-based applications designed for the implanted population.

F.5 Conclusions

This article described how we applied a user-centric participatory design method to create a multisensory concert experience including audio, video, and vibrotactile stimulation. The vibrotactile furniture designed focused on flexibility and user choice, as opposed to most previous work that employed a fixed-posture experience. Through the design process, we have concluded that it is fundamental to acknowledge individual CI user needs in the design process. These needs could be related to their hearing characteristics and, just as importantly, to their attitude towards music listening in general. Furthermore, we observed that when designing multiactuator systems, it is important to consider the body area actuated, as different zones have different vibrotactile sensing characteristics.'

Two vibrotactile displays were premiered at the Danish Royal Academy of Music during a workshop created for people with hearing impairment. More than 50 participants participated in the event and 5 of them were CI users. The concert lasted 25 min and presented adaptations of 6 popular songs from different generations, interpreted by a female vocalist, with double bass and drums as accompaniment. The evaluation focused on the usability of the displays and the multisensory concert experience as a whole, and it indicated that such a concert concept is welcomed by the CI community as it provides several benefits, such as better stream separation. Although the concert experience was welcomed, our evaluation concluded that audio-vibrotactile mixing is crucial, and small adjustments can result in an overwhelming experience, as is to be expected from the limited dynamic range both the CI and the tactile receptors have. Lastly, when it comes to multiactuator displays, the participants had a strong preference towards a one-to-one mapping between musical instruments and body areas actuated. This idea has been explored and confirmed by many previous researchers working with a distribution of speaker based on *The Model Human Cochlea* [31], which suggests to employ a positive correlation between stimuli frequency and the vertical distribution of vibrotactile stimulation. It is also important to take into consideration their taste and preferences in terms of leisure activities. As a matter of fact, the CoolHear Workshop included both a demo session with different gamified experiences and a concert. We observed that younger audience members preferred to engage in the gamified experiences and skipped the concert, while older participants engaged in the concert but did not enjoy the gamified experiences, especially those including a virtual reality display.

The work presented in this article suggests that multisensory live music does have the potential to improve the concert experience for CI users, but the field is in its infancy. Future work should focus on further understanding what factors are responsible for such an improvement, as well as continuing to involve CI users int the vibrotactile display design process as early as possible.

References

[1] *An Exploration on Whole-body and Foot-based Vibrotactile Sensitivity to Melodic Consonance*. Zenodo, 09 2016. [Online]. Available: https://doi.org/10.5281/zenodo.851213

[2] S. C. Aker, H. Innes-Brown, K. F. Faulkner, M. Vatti, and J. Marozeau, "Effect of audio-tactile congruence on vibrotactile music enhancement," *The Journal of the Acoustical Society of America*, vol. 152, no. 6, pp. 3396–3409, 2022.

[3] D. Aydın and S. Yıldız, "Effect of classical music on stress among preterm infants in a neonatal intensive care unit," *HealthMED*, vol. 6, pp. 3162–3168, 01 2012.

[4] A. Baijal, J. Kim, C. Branje, F. A. Russo, and D. I. Fels, "Composing vibrotactile music: A multisensory experience with the emoti-chair," *CoRR*, vol. abs/1509.05452, 2015.

[5] D. Başkent and R. V. Shannon, "Interactions between cochlear implant electrode insertion depth and frequency-place mapping," *The Journal of the Acoustical Society of America*, vol. 117, no. 3, pp. 1405–1416, 2005.

[6] T. Bjorner, *Qualitative Methods for Consumer Research: The Value of the Qualitative Approach in Theory and Practice*. Hans Reitzels Forlag, 2016. [Online]. Available: https://books.google.com/books?id=xAjtrQEACAAJ

[7] W. Buyens, B. V. Dijk, M. Moonen, and J. Wouters, "Music mixing preferences of cochlear implant recipients: A pilot study," *International Journal of Audiology*, vol. 53, 2014.

[8] D. Cavdir, F. Ganis, R. Paisa, P. Williams, and S. Serafin, "Multisensory Integration Design in Music for Cochlear Implant Users," 06 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6798118

[9] D. Cavdir and G. Wang, "Felt sound: A shared musical experience for the deaf and hard of hearing," in *Proceedings of the 20th international conference on new interfaces for musical expression (nime-20)*, 2020.

[10] M. V. Certo, G. D. Kohlberg, D. A. Chari, D. M. Mancuso, and A. K. Lalwani, "Reverberation time influences musical enjoyment with cochlear implants," *Otology and Neurotology*, vol. 36, 2015.

[11] D. D. Coffman, "Music and quality of life in older adults." *Psychomusicology: A Journal of Research in Music Cognition*, vol. 18, 2002.

[12] A. Dickens, C. Greenhalgh, and B. Koleva, "Facilitating accessibility in performance: participatory design for digital musical instruments," *Journal of the Audio Engineering Society*, vol. 66, no. 4, pp. 211–219, 2018.

[13] M. F. Dorman, L. Loiselle, J. Stohl, W. A. Yost, A. Spahr, C. Brown, and S. Cook, "Interaural level differences and sound source localization for bilateral cochlear implant patients," *Ear and hearing*, vol. 35, no. 6, p. 633, 2014.

[14] M. F. Dorman, L. H. Loiselle, S. J. Cook, W. A. Yost, and R. H. Gifford, "Sound source localization by normal-hearing listeners, hearing-impaired listeners and cochlear implant listeners," *Audiology and Neurotology*, vol. 21, no. 3, pp. 127–131, 2016.

[15] G. Dritsakis, R. M. van Besouw, P. Kitterick, and C. A. Verschuur, "A music-related quality of life measure to guide music rehabilitation for adult cochlear implant users," *American Journal of Audiology*, vol. 26, 2017.

[16] G. Dritsakis, R. M. van Besouw, and A. O. Meara, "Impact of music on the quality of life of cochlear implant users: a focus group study," *Cochlear Implants International*, vol. 18, 2017.

[17] M. D. Fletcher, N. Thini, and S. W. Perry, "Enhanced pitch discrimination for cochlear implant users with a new haptic neuroprosthetic," *Scientific Reports*, vol. 10, 2020.

[18] J. Foxe, G. Wylie, A. Martinez, C. Schroeder, D. Javitt, D. Guilfoyle, W. Ritter, and M. Murray, "Auditory-somatosensory multisensory processing in auditory association cortex: An fmri study," *Journal of neurophysiology*, vol. 88, pp. 540–3, 08 2002.

[19] E. Frid, "Accessible digital musical instruments—a review of musical interfaces in inclusive music practice," *Multimodal Technologies and Interaction*, vol. 3, no. 3, p. 57, 2019.

[20] A. Frosolini, G. Badin, F. Sorrentino, D. Brotto, N. Pessot, F. Fantin, F. Ceschin, A. Lovato, N. Coppola, A. Mancuso, L. Vedovelli, G. Marioni, and C. de Filippis, "Application of patient reported outcome measures in cochlear implant patients: Implications for the design of specific rehabilitation programs," *Sensors*, vol. 22, no. 22, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/22/8770

[21] C. D. Fuller, I. John J. Galvin, B. Maat, D. Başkent, and R. H. Free, "Comparison of two music training approaches on music and speech perception in cochlear implant users," *Trends in Hearing*, vol. 22, no. 9999, p. 2331216518765379, 2018, pMID: 29621947. [Online]. Available: https://doi.org/10.1177/2331216518765379

[22] J. J. Galvin III, E. Eskridge, S. Oba, and Q.-J. Fu, "Melodic contour identification training in cochlear implant users with and without a competing instrument," in *Seminars in hearing*, vol. 33. Thieme Medical Publishers, 2012, pp. 399–409.

[23] K. Gfeller, V. Driscoll, and A. Schwalje, "Adult cochlear implant recipients' perspectives on experiences with music in everyday life: A multifaceted and dynamic phenomenon," *Frontiers in neuroscience*, p. 1229, 2019.

[24] J. Gosine, D. Hawksley, and S. L. Quinn, "Community building through inclusive music-making," in *Voices: A World Forum for Music Therapy*, vol. 17, 2017.

[25] J. B. Hinderink, P. F. Krabbe, and P. V. D. Broek, "Development and application of a health-related quality-of-life instrument for adults with cochlear implants: The nijmegen cochlear implant questionnaire," *Otolaryngology - Head and Neck Surgery*, vol. 123, 2000.

[26] L. K. Holden, J. B. Firszt, R. M. Reeder, R. M. Uchanski, N. Y. Dwyer, and T. A. Holden, "Factors affecting outcomes in cochlear implant recipients implanted with a perimodiolar electrode array located in scala tympani," *Otology and Neurotology*, vol. 37, pp. 1662–1668, 2016. [Online]. Available: https://doi.org/10.1097/MAO.0000000000001241

[27] C. Hopkins, S. Maté-Cid, R. Fulford, G. Seiffert, and J. Ginsborg, "Vibrotactile presentation of musical notes to the glabrous skin for adults with normal hearing or a hearing impairment: Thresholds, dynamic range and high-frequency perception," *PLoS ONE*, vol. 11, 2016.

[28] R. Jack, A. Mcpherson, and T. Stockman, "Designing tactile musical devices with and for deaf users: a case study," *Proc. of the International Conference on the Multimedia Experience of Music*, 2015.

[29] L. A. Jones and N. B. Sarter, "Tactile displays: Guidance for their design and application," *Human Factors*, vol. 50, 2008.

[30] V. Jousmäki and R. Hari, "Parchment-skin illusion: sound-biased touch." *Current biology : CB*, vol. 8, 1998.

[31] M. Karam, F. Russo, and D. Fels, "Designing the model human cochlea: An ambient crossmodal audio-tactile display," *IEEE Transactions on Haptics*, vol. 2, pp. 160–169, 2009.

[32] C. Kayser, C. I. Petkov, M. Augath, and N. K. Logothetis, "Integration of touch and sound in auditory cortex," *Neuron*, vol. 48, 2005.

[33] A. Lucas, M. Ortiz, and F. Schroeder, "The longevity of bespoke, accessible music technology: a case for community," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2020, pp. 243–248.

[34] A. M. Lucas, M. Ortiz, and F. Schroeder, "Bespoke design for inclusive music: The challenges of evaluation." in *NIME*, 2019, pp. 105–109.

[35] P. Marti and A. Recupero, "Is deafness a disability? designing hearing aids beyond functionality," in *Proceedings of the 2019 on Creativity and Cognition*, 2019, pp. 133–143.

[36] H. J. McDermott, "Music perception with cochlear implants: A review," *Trends in Amplification*, vol. 8, no. 2, pp. 49–82, 2004, pMID: 15497033. [Online]. Available: https://doi.org/10.1177/108471380400800203

[37] T. R. McRackan, B. N. Hand, C. A. Velozo, J. R. Dubno, J. S. Golub, E. P. Wilkinson, D. Mills, J. P. Carey, N. Vorasubin, V. Brunk, M. L. Carlson, C. L. Driscoll, D. P. Sladen, E. L. Camposeo, M. A. Holcomb, P. R. Lambert, T. A. Meyer, C. Thomas, A. C. Moberly, N. H. Blevins, J. B. Larky, R. P. Herzano, M. E. Hoffer, S. M. Prentiss, R. N. Samy, S. P. Gubbels, J. Brant, J. B. Hunter, B. Isaacson, J. W. Kutz, R. K. Gurgel, D. M. Zeitler, C. A. Buchman, J. B. Firszt, R. H. Gifford, D. S. Haynes, and R. F. Labadie, "Cochlear implant quality of life (ciqol): Development of a profile instrument (ciqol-35 profile) and a global measure (ciqol-10 global)," *Journal of Speech, Language, and Hearing Research*, vol. 62, 2019.

[38] L. Mitchell, R. MacDonald, C. Knussen, and M. Serpell, “A survey investigation of the effects of music listening on chronic pain,” *Psychology of Music*, vol. 35, pp. 37–57, 01 2007.

[39] S. Nanayakkara, L. Wyse, and E. Taylor, “The haptic chair as a speech training aid for the deaf,” *Proceedings of the 24th Australian Computer-Human Interaction Conference, OzCHI 2012*, pp. 405–410, 2012.

[40] S. Nanayakkara, E. Taylor, L. Wyse, and S. Ong, “An enhanced musical experience for the deaf: Design and evaluation of a music display and a haptic chair,” 04 2009, pp. 337–346.

[41] R. Nordahl, S. Serafin, N. C. Nilsson, and L. Turchet, “Enhancing realism in virtual environments by simulating the audio-haptic sensation of walking on ground surfaces,” *Proceedings - IEEE Virtual Reality*, 2012.

[42] A. J. Oxenham, “Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants,” *Trends in amplification*, vol. 12, no. 4, pp. 316–331, 2008.

[43] R. Paisa, J. Andersen., N. C. Nilsson, and S. Serafin, “A comparison of audio-to-tactile conversion algorithms for melody recognition,” in *Baltic Nordic-Acoustic Meetings*, 2022.

[44] C. Quintero, “A review: accessible technology through participatory design,” *Disability and Rehabilitation: Assistive Technology*, pp. 1–7, 2020.

[45] J. Rouger, S. Lagleyre, B. Fraysse, S. Deneve, O. Deguine, and P. Barone, “Evidence that cochlear-implanted deaf patients are better multisensory integrators,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, 2007.

[46] K. Samuels and F. Schroeder, “Performance without barriers: improvising with inclusive and accessible digital musical instruments,” *Contemporary Music Review*, vol. 38, no. 5, pp. 476–489, 2019.

[47] E. Schellenberg, *Music and nonmusical abilities*, 01 2016, pp. 149–176.

[48] T. Schäfer, P. Sedlmeier, C. Städtler, and D. Huron, “The psychological functions of music listening,” *Frontiers in Psychology*, vol. 4, p. 511, 2013. [Online]. Available: https://www.frontiersin.org/article/10.3389/fpsyg.2013.00511

[49] M. Someren, Y. Barnard, and J. Sandberg, *The Think Aloud Method - A Practical Guide to Modelling Cognitive Processes*. London: Academic Press, 1994.

[50] B. Stein, P. Stein, and M. Meredith, *The Merging of the Senses*, ser. A Bradford book. MIT Press, 1993. [Online]. Available: https://books.google.se/books?id=uCV9QgAACAAJ

[51] A. Tilley, H. Dreyfuss, and H. D. Associates, *The Measure of Man and Woman: Human Factors in Design*. Whitney Library of Design, 1993. [Online]. Available: https://books.google.dk/books?id=5qzHQgAACAAJ

[52] A. Wilska, "On the vibrational sensitivity in different regions of the body surface," *Acta Physiologica Scandinavica*, vol. 31, 1954.

[53] B. S. Wilson and M. F. Dorman, "Cochlear implants: A remarkable past and a brilliant future," *Hearing Research*, vol. 242, 2008.

[54] F.-G. Zeng, *Compression and Cochlear Implants*. New York, NY: Springer New York, 2004, pp. 184–220. [Online]. Available: https://doi.org/10.1007/0-387-21530-1_6

[55] F. G. Zeng and J. J. Galvin, "Amplitude mapping and phoneme recognition in cochlear implant listeners," *Ear and Hearing*, vol. 20, 1999.

[56] F.-G. Zeng, G. Grant, J. Niparko, J. Galvin, R. Shannon, J. Opie, and P. Segel, "Speech dynamic range and its effect on cochlear implant performance," *The Journal of the Acoustical Society of America*, vol. 111, 2002.

[57] R. Črnčec, S. Wilson, and M. Prior, "The cognitive and academic benefits of music to children: Facts and fiction," *Educational Psychology - EDUC PSYCHOL-UK*, vol. 26, pp. 579–594, 08 2006.

Paper G

A Concert-Based Study on Melodic Contour Identification Among Varied Hearing Profiles

R. Paisa, J. Andersen, F. Ganis, L. M. Percy-Smith and S. Serafin

The paper has been submitted and is under revision

The layout has been revised.

Abstract

This study investigates the impact of hearing profiles on melodic contour identification in an ecologically valid concert setting. The research aimed to understand how different hearing profiles, including normal hearing, double-sided hearing aids, bimodal (hearing aid and cochlear implant), single-sided cochlear implant, and bilateral implants, influence the ability to perceive melodic contours played by multiple combinations of instruments. The stimuli was played piano or accordion, with and without an electric bass as masker and accompanied by a simple drum groove. In total, 44 participants were involved in the study and the data was analyzed by fitting Bayesian logistic mixed-effects models. The key findings revealed that introducing an electric bass as masker to a piano melody did not significantly impact MCI performance for any of the hearing groups, contrary to its effect on accordion melodies and what previous literature suggests. Furthermore, participants demonstrated greater difficulty in following melodies played on the accordion, especially when accompanied by an electric bass groove than those performed on the piano. Additionally, the study collected data on MCI performance for hearing aid users as well and shows that their score are similar to other hearing impaired profiles, challenging the initial hypothesis that they would outperform cochlear implant users. These results emphasize the importance of considering specific instrument combinations when studying music perception among individuals with diverse hearing profiles. Lastly, we have designed and published a cohort of short melodies to be use for contour identification tasks, that take inspiration from western music styles.

Paper H

Effects of Music Training on Concert Experiences and Melodic Contour Identification for Hearing Impaired Individuals

J. Andersen, R. Paisa and S. Serafin

The paper has been submitted and is under revision

The layout has been revised.

Abstract

This study presents and evaluates a music training methodology designed for hearing aid and cochlear implant users. The training consisted of a series of podcast-style lessons, totalling approximately four hours, focused on dissecting and discussing four popular songs over two months. The eight participants were offered identical loudspeakers and half of them had access to a custom vibrotactile display designed for music listening. Their music listening performance was evaluated in a live concert scenario utilising a comparative Melodic Contour Identification test, and the experience was investigated through semi-structured interviews. Results show that music training has a perceived positive effect on the experience of concert going for hearing impaired, and there was a noticeable effect of training over their hearing performance. Furthermore, it was observed that the vibrotactile element did not significantly enhance the training experience, and that the 15-minute lesson duration might be overly demanding, particularly for older participants, leading to incomplete training sessions.

Music listening is a quintessential aspect of human life; it has been a constant companion of humanity throughout the ages, for as long as we can find evidence of people. Nowadays, the scientific community agrees that music has numerous benefits for the individual well-being, mental and even physical health, by encouraging social relationships, emotional healing, memory recollection, as well as it's intrinsic value --- music is good. A big obstacle in accessing music is hearing impairment that many individuals must live with. In the case of restoring severe hearing loss, the solution is usually found in a cochlear implant - a neuroprosthetic device that restores hearing by stimulating the inner ear directly. While these implants show fantastic results for speech related activities, indicated by the fact that most cochlear implant users live their daily life as normal hearing individuals, the music perception is severely degraded. As this segment of the population will only increase in the next 25 years, it is imperative to research new and innovative solutions that will work in cooperation with the inevitable advancement in hearing assistive devices.

One such opportunity can be found in using multiple sensory channels while ensuring the rules for multisensory integration are upheld. This idea is as old as the cochlear implant itself, as the creator of the first commercially available cochlear implant designed a tactile device called the "Tickle Talker", that was supposed to aid in the perception of sound by the hearing impaired user. This is the legacy this project continues, as advancements in the tactile technology afford designing significantly better devices, the research needs to explore and propose the best usage of those novel tools.

This dissertation is an interdisciplinary project conducted in collaboration with two institutions that have been immensely supportive of the work - The Royal Danish Academy of Music as well as the Copenhagen Center for Hearing and Balance from Rigshospitalet. Through an applied research project, we explored the possibilities of using vibrotactile stimulation designed around the needs of cochlear implant users in order to improve their music hearing performance and experience.

The primary contribution of this doctoral research is in the eight papers presented in part II, as well as the discussion presented in part I. The first paper in this series presents a scoping review of vibrotactile devices applicable to music, underscoring recurring themes and gaps within existing literature, and emphasizing the lack of standardization in this field. The subsequent two papers delve into the study of musical perceptual features through the use of vibrotactile displays, while also addressing the constraints inherent in single-actuator devices. The remaining five articles explore different facets of music listening experiences for cochlear implant users in social settings, particularly in concert environments. Among these, some studies are dedicated to the development of tactile displays specifically designed for concert use, which I refer to as 'concert furniture' while others focus on assessing and training the auditory performance of CI users in live music scenarios. These studies collectively aim to provide a deeper understanding of how cochlear implant users interact with and perceive music in concert scenarios and how vibrotactile technology can enhance these experiences.