Tune In

Exploring Music and Sound Training
for Children with Hearing Challenges

PhD Thesis 2025

Francesco Ganis

Supervisors: Professor Stefania Serafin
Rigshospitalet - Dept. of Otorhinolaryngology, Head and Neck Surgery and Audiology, Denmark

Associate Professor Ali Adjorlu,
Aalborg University, Denmark

Professor Lone Marianne Percy-Smith,
Rigshospitalet - Dept. of Otorhinolaryngology, Head and Neck Surgery and Audiology, Denmark

Assessment: Associate Professor Timothy Robert Merritt (chair)
Aalborg University, Denmark

Associate Professor Melanie Ferguson
Curtin University, Australia

Affiliate Associate Professor Hamish Innes-Brown
Technical University of Denmark, DTU, Denmark

Department: Dept. of Architecture, Design & Media Technology

© Copyright: Francesco Ganis

In this PhD thesis, AI technologies (LLM, large language models) have been used to assist in the refinement of grammar and syntax corrections.

Curriculum Vitae

Francesco Ganis

I earned a bachelor's degree from the Conservatory of Music Cesare Pollini in Padua, Italy, specializing in Electronic Music and Sound Technology. This choice was driven by my interest in sound engineering, which combines my two main passions: music and technology. I worked as a sound technician in both classical and pop/rock contexts and acquired skills in electronic music production and live mixing at the Conservatory. During my studies, I developed an interest in Sound and Music Computing (SMC) thanks to Professor Nicola Bernardini, which led me to pursue a master's degree in SMC at Aalborg University in Copenhagen, Denmark, where I gained a strong theoretical and practical foundation. An internship at Oticon Medical and my master's thesis sparked my fascination with sound technology for individuals with hearing loss and prompted me to consider how I could further contribute to enhancing the musical experience for this population. Motivated by this, I began a PhD in September 2022 under the supervision of Professor Stefania Serafin, in collaboration with the Copenhagen Hearing and Balance Center at Rigshospitalet, Copenhagen. This thesis presents the outcome of that research.

Abstract

Hearing loss affects millions of people worldwide, and its prevalence is projected to rise due to factors such as population aging and increased exposure to loud noise. Although hearing aids and cochlear implants have significantly improved speech perception, individuals with hearing loss continue to face challenges in other domains, including music appreciation and speech understanding in noisy environments. Newer generations of children with hearing loss benefit from early detection and intervention programs, outperforming previous generations in language development and overall quality of life. Nevertheless, many still experience substantial difficulties with complex listening tasks, including spatial hearing and vocal emotion recognition, and those with special conditions often exhibit poor outcomes despite early intervention and auditory training.

This thesis investigates the potential of novel training tools to support the musical and auditory development of children with hearing loss. Grounded in user-centered design principles, the research involves clinical professionals, children, and their parents throughout the design and evaluation processes, in collaboration with the Copenhagen Center for Hearing and Balance at Rigshospitalet.

Part I provides an Extended Summary that includes a brief introduction (Chapter 1), synthesizes the background literature (Chapter 2), research questions (Chapter 3), a summary of the papers (Chapter 4), and findings (Chapter 5). The main contribution comprises six core research papers reported in Part II, including a state-of-the-art analysis (Paper A) and the development of five prototypes: three vibrotactile systems—Tickle Tuner (Paper B), designed for individuals with hearing loss; Vibrotactile Memory (Paper C); and Vibrotactile Teddy Bear (Paper D), both tailored for children—and two additional tools, SoundCubes (Paper E), designed to enhance spatial hearing, and EmotiCubes (Paper F), developed to improve vocal emotion recognition, both intended for pediatric use. These studies examine the design, implementation, and evaluation of the prototypes, providing insights into their usability and effectiveness, and identifying challenges and future research directions. Additionally, two further studies presented in Part III explore neurophysiological mechanisms of audio-tactile integration (Paper G) and participatory design approaches for multisensory music experiences with cochlear implant users (Paper H). Collectively, this work advances understanding of multisensory strategies for enhancing auditory experiences in children with hearing loss and contributes novel design solutions for inclusive auditory training.

Resumé

Høretab påvirker millioner af mennesker verden over, og forekomsten forventes at stige på grund af faktorer som befolkningens aldring og øget udsættelse for høj støj. Selvom høreapparater og cochleaimplantater har markant forbedret taleopfattelsen, møder personer med høretab fortsat udfordringer på andre områder, herunder musikoplevelse og taleforståelse i støjende omgivelser. Nyere generationer af børn med høretab drager fordel af tidlig opsporing og interventionsprogrammer og klarer sig bedre end tidligere generationer, hvad angår sprogudvikling og overordnet livskvalitet. Til tods for dette, oplever mange fortsat betydelige vanskeligheder ved komplekse lytteopgaver, herunder rumlig hørelse og genkendelse af følelser i stemmen, og dem med særlige tilstande udviser ofte utilfredsstillende resultater trods tidlig indsats og auditiv træning.

Denne afhandling undersøger potentialet i nye træningsværktøjer til at understøtte den musikalske og auditive udvikling hos børn med høretab. Forankret i brugercentrerede designprincipper involverer forskningen kliniske fagpersoner, børn og deres forældre gennem hele design- og evalueringsprocesserne i samarbejde med Copenhagen Center for Hearing and Balance på Rigshospitalet.

Del I giver en Udvidet Sammenfatning, der indeholder en kort introduktion (Kapitel 1), syntetiserer baggrundslitteraturen (Kapitel 2), forskningsspørgsmålene (Kapitel 3), en sammenfatning af artiklerne (Kapitel 4) og resultaterne (Kapitel 5). Hovedbidraget omfatter seks kerneforskningsartik-

ler beskrevet i Del II, herunder en state-of-the-art-analyse (Paper A) og udviklingen af fem prototyper: tre vibrotaktile systemer—Tickle Tuner (Paper B), designet til personer med høretab; Vibrotactile Memory (Paper C); og Vibrotactile Teddy Bear (Paper D), begge målrettet børn—samt to yderligere værktøjer, SoundCubes (Paper E), designet til at forbedre rumlig hørelse, og EmotiCubes (Paper F), udviklet til at forbedre genkendelsen af følelser i stemmen, begge tiltænkt pædiatrisk brug. Disse studier undersøger design, implementering og evaluering af prototyperne, giver indblik i deres anvendelighed og effektivitet og identificerer udfordringer og fremtidige forskningsretninger. Derudover præsenteres to yderligere studier i Del III, som undersøger neurofysiologiske mekanismer ved audio-taktil integration (Paper G) og deltagende design-tilgange til multisensoriske musikoplevelser med brugere af cochleaimplantater (Paper H). Samlet set fremmer dette arbejde forståelsen af multisensoriske strategier til at forbedre auditive oplevelser hos børn med høretab og bidrager med nye designløsninger til inklusiv auditiv træning.

Contents

Preface

Preface

Five years ago, when I started my master's degree in Sound and Music Computing, I could never have imagined that I would end up embarking on a PhD. The journey has been fun, challenging, rewarding, and at times frustrating, but overall, it has been an incredible learning experience that has shaped me both professionally and personally. I feel privileged to have had the opportunity to continue working on topics I am truly passionate about, especially music. Anyone who knows me, even a little, understands how much music means to me and how deeply I have been involved in its creation and enjoyment. Throughout these years, I have met amazing people who share the same passion, each in their own way, which makes it even more special. I could go on talking about music for days, but I will stop here and focus on those who contributed to making this PhD journey possible, enjoyable, and successful.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor, Stefania Serafin. She guided me through my Master's and PhD journey with patience, support, and energy. Regardless of the situation, she always found the strength to keep me motivated and moving forward, even when I doubted myself. Her passion for research and impressive networking skills

make her an exceptional mentor, and I feel incredibly lucky to have shared this experience with her. I am also thankful to my supervisor Ali Adjorlu, who has been a great source of support, Unity expertise, and a fantastic brainstorming companion, as well as a peer in the lab's social life. This PhD would not have been the same without Lone Marianne Percy-Smith, my supervisor from CHBC, whose clinical expertise, insights, and cheerful assistance have been truly invaluable. Her perspective opened my eyes and ears to new ways of thinking and approaching research, and I am grateful for the opportunity to work alongside her.

My well-known faulty memory will hopefully not let me forget anyone, but if so, I apologize in advance. Many people contributed to this PhD, and each shaped it in a unique way.

First of all, I want to thank Razvan Paisa, my PhD mate, who shared his knowledge and helped me immensely throughout this journey. Even better, he became a great friend, and I am grateful to have him by my side. Thanks to Jesper Andersen for being a wonderful colleague and friend, for all the great music chats, the listening sessions, and for introducing me to his awesome DKDM world. I would like to thank all the colleagues and friends from the ME Lab and Aalborg University, past and present, for creating such a stimulating and friendly environment. Special thanks to Anders Riddersholm Bargum, Doga Cavdir, Mark Crispin Sandholt Dourado, Sofia Dahl, Kewei Du, Cumhur Erkut, Emil Rosenlund Høeg, Mels Jagt, Prithvi Ravi Kantan, Marius Onofrei, Dan Overholt, Simon Rostami, Thomas Albert Rushton, and Silvin Willemsen for being part of this journey. Thanks also to Jesper Greve and Peter Williams for their technical support with lab equipment and the long problem-solving chats. Thanks to Lene Rasmussen for her excellent administrative support and for always being so kind and helpful.

A special thank you goes to Federico Fontana from the University of Udine for the insightful collaboration and for sharing his expertise, as well as for the nice time during his stay in Copenhagen. I am also grateful to Andrea Gulli from the University of Udine for sharing the challenge of writing a systematic review together. Finally, thanks to Romain Michon from INRIA for for the lovely time spent together during his visits to Copenhagen, for the great chats about life, and for organizing the best SMC conference ever.

I would also like to express my sincere gratitude to Abigail Anne Kressner from DTU for her collaboration, insights, professionalism, and precious work in the SoundCubes project.

A heartfelt thank you goes to Deniz Başkent for hosting me at db SPL in Groningen during my research stay, for her generous support, and for being an inspiring researcher who found time to guide me and share her knowledge. I am equally grateful to Laura Rachman for her guidance throughout my research experience with the EmotiCubes and for sharing her expertise, enthusiasm, and dedication.

I am deeply thankful to all the clinical professionals I had the pleasure to meet and work with at CHBC, Rigshospitalet, including Lærke Hammer, Emil Sønderskov Hansen, Lone Jantzen, Nille Elise Kepp, Nete Rudbeck Kamper, and Signe Wischmann. A special thank you goes to Cecilia Fernandez Samar for her support in my research activities, and to Karen Lise Roslyng for her amazing clinical help, carried out with kindness, commitment, and professionalism. Thanks to Christa Thomsen from Børnecenter København and Lotte Rømer from Høreforeningen for their support in recruiting participants for my studies.

A huge thank you goes to all the children and their families who participated in my studies and let that Italian guy give them a prototype to try out. Without them, this PhD would not have been possible. Their enthusiasm, curiosity, and joy have been a driving force throughout this journey.

Outside the academic world, I would like to thank my friends for being there for me, especially those in Italy whom I do not see often but who are always close to my heart. Thanks to Nicola Bernardini, who believed in me from the very beginning and encouraged me to pursue this path. Your friendship and support mean a lot to me.

I kept the dearest thank you for last: to my parents, for their unconditional support, love, and understanding that made all my achievements possible. You always believed in me and let me pursue my interests and dreams, even when they took me far away from home. Thanks to my lovely big brother, who also lives far away and has been a great source of support and inspiration. Vi voglio bene! Finally, I want to thank my girlfriend for believing in and supporting me throughout this journey, for listening to my rants, and for tirelessly checking all the drafts. I am grateful to have you by my side. I love you!

Francesco
Copenhagen, December 17, 2025

Glossary

List of Acronyms

Notation	Description	Page List
ADC	analog-to-digital converter	16
ADMI	accessible digital musical instruments	232, 233
AM	amplitude modulation	52, 53, 149–152
AMT	accessible musical technologies	232
AN	auditory nerve	12, 54, 55
ANSD	auditory neuropathy spectrum disorder	14, 15, 54, 55, 70, 160, 161, 163, 172, 173, 179, 180
AT	auditory training	20, 38, 39, 41–43, 49, 55, 57, 59
AVT	auditory-verbal therapy	19, 21, 38–40, 42, 55–57, 59, 69, 70, 72, 179, 180, 183, 187–190
CHBC	Copenhagen Hearing and Balance Center	4, 6, 54–57, 68, 72, 73, 183

Notation	Description	Page List
CI	cochlear implant	3, 4, 6, 14–21, 23–26, 34–36, 38, 39, 42, 43, 52–57, 60, 63–66, 69, 96, 107, 108, 110, 126, 143, 144, 148–153, 159, 160, 173, 179, 180, 183, 186–188, 217, 218, 231, 233–236, 238, 240, 246–248
CND	cochlear nerve deficiency	4, 14, 15, 55–57, 70, 72, 179, 180, 187, 188, 190
D/HOH	Deaf and Hard of Hearing	33, 232, 241
d/hoh	deaf and hard of hearing	20, 21, 57
DAC	digital-to-analog converter	52, 143, 145, 146
DAW	digital audio workstation	165
DBSC	double-sideband suppressed carrier	34, 149
DSP	digital signal processing	16, 33, 127, 129, 144
ERM	eccentric rotating mass	32, 119, 128
F_0	fundamental frequency	25, 26, 62, 63, 68, 207, 210, 211, 214, 215, 217, 218, 222, 233
FFR	frequency following response	62, 64, 202, 203, 206–219, 221, 222
FFT	fast Fourier transform	63, 210
HA	hearing aid	4, 15, 16, 18, 23–26, 35, 38, 54, 55, 57, 127, 159, 179, 231, 232, 236, 243, 246
HC	hearing challenges	3–6, 9, 27, 29, 38–40, 47, 50, 67, 72, 74–76, 179
HF	high frequency	210, 214

Notation	Description	Page List
HL	hearing loss	6, 9, 13–16, 18–28, 32, 35–43, 47–51, 53, 55, 57–61, 68–70, 72, 73, 75, 76, 96, 106, 159, 160
ICC	interclass correlation coefficient	58
IHC	inner hair cell	12, 17, 18, 22
ILD	interaural level difference	36, 57
ITD	interaural time difference	15, 57
JNDF	just noticeable difference in frequency	30
LRA	linear resonant actuator	32
MCI	melodic contour identification	148, 150
ME Lab	Multisensory Experience Lab	3–5, 62
NH	normal hearing	96, 105–109, 111, 112, 160
OHC	outer hair cell	12–14, 22
PTA	pure tone audiometry	16, 18
QOL	quality of life	9, 13–15, 19, 23, 38, 39, 55, 57, 60
RMS	root mean square	212
RMSE	root mean square error	58
SMC	Sound and Music Computing	3, 7
SNHL	sensorineural hearing loss	14, 15
SNR	signal-to-noise ratio	212, 213, 218
SUS	System Usability Scale	58, 60
TH	typical hearing	18, 19, 23, 26, 34, 37, 40, 43, 58, 69

Notation	Description	Page List
TPT	Timbre Perception Test	159, 162, 167, 169, 172
TUI	tangible user interface	60
VR	virtual reality	44

List of Papers

Included in Thesis

The thesis is based on the following papers:

A **F. Ganis**, A. Gulli, F. Fontana, and S. Serafin, “The Role of Haptics in Training and Games for Hearing-Impaired Individuals: A Systematic Review,” in *Multimodal Technologies and Interaction (MTI)*, 2024.

B **F. Ganis**, M. Vatti, and S. Serafin, “Tickle Tuner - Haptic Smartphone Cover for Cochlear Implant Users’ Musical Training,” in *Haptic and Audio Interaction Design, in Lecture Notes in Computer Science*, 2022.

C **F. Ganis**, A. Adjorlu, L. Percy-Smith, C. F. Samar and S. Serafin, “Vibrotactile Memory: a Case Study of Timbre Perception Training in Children with Cochlear Implants Using a Video Game”, in *Proceedings of the 21st Sound and Music Computing Conference*, 2024.

D **F. Ganis**, A. Adjorlu, K. L. Roslyng, L. M. Percy-Smith and S. Serafin, “Vibrotactile Teddy Bear: Enhancing Musical Experiences for Children with Cochlear Nerve Deficiency through a Vibrotactile Soft Toy”, in *Proceedings of the 22nd Sound and Music Computing*

Conference, 2025.

E **F. Ganis**, A. A. Kressner, L. Percy-Smith, A. Adjorlu, and S. Serafin, ”SoundCubes - Preliminary Evaluation of a Spatial Hearing Training Tool and a Sound Localization Test for Children with Hearing Loss” - Submitted to *International Journal of Human-Computer Interaction*, 2025.

F **F. Ganis**, L. Rachman, D. Başkent, and S. Serafin, ”EmotiCubes - A Training Tool for Vocal Emotion Recognition in Children with Hearing Loss” - Submitted to *Journal of Clinical Medicine*, 2025.

Additional Papers

The following additional papers resulted from collaborations during the PhD but are not included in the thesis for assessment purposes:

G M. Jagt, **F. Ganis**, and S. Serafin, “Enhanced neural phase locking through audio-tactile stimulation,” Frontiers in Neurosciences, 2024.

H D. Cavdir, **F. Ganis**, R. Paisa, P. Williams, and S. Serafin, ”Multisensory Integration Design in Music for Cochlear Implant Users,” in *Proceedings of the 19th Sound and Music Computing Conference*, 2022.

Part I

Extended Summary

CHAPTER 1

Introduction

1.1 Prelude

Music has been an integral part of human life for millennia, and my own experience is no exception: in fact, it had a major role in my upbringing, and shaping my identity, study, and career path. I began my musical education at the age of three and started playing drums at five. I got involved in several music projects, many of my friendships and connections have developed through music. These and many more reasons made me think that making experiencing music for children with hearing challenges (HC) more accessible and enjoyable would be a perfect fit for my research, since I personally cannot imagine life without sound and music. Indeed, music plays a vital role in the lives of many people, encompassing social interaction, entertainment, and even therapy. Being a passionate musician or avid music consumer is not what everyone is aiming for, but I believe that everyone should have the opportunity to choose to be one.

During my Sound and Music Computing (SMC) master's internship, I had the opportunity to work on a project focused on making music more accessible for people with HC, specifically with cochlear implant (CI) through vibrotactile feedback. This collaboration between the Multisensory Experience Lab (ME Lab) and Oticon Medical, supervised by Professor Stefania Serafin and Marianna Vatti, became the basis for my master's thesis, and revealed how I could combine my passion for music and technology with accessibility. With the support of my supervisors, I continued this journey in my PhD, which has culminated in this thesis. Recognizing that the thesis is a milestone rather than an endpoint, I am excited to further contribute to music accessibility for people with hearing challenges.

1.2 Definition of the Project

This PhD project aims to explore how training can be used to enhance the music experience for children with hearing challenges. The project has its roots in the ME Lab at Aalborg University, where research on human-computer interaction, the development of novel technologies, and solutions for specific target groups is a key focus. Furthermore, this project shares common foundations with research conducted by my esteemed colleagues: Dr. Razvan Paisa, who has investigated vibrotactile feedback to enhance music enjoyment for adults with CI and hearing aid (HA), and Jesper Andersen, who has focused on music training and concert experiences for individuals with HC.

As mentioned in the Prelude 1.1, this project evolved from a master's thesis and initially aimed to explore how vibrotactile feedback could enhance the music experience for individuals with CI. Through the beginning of a fruitful collaboration between the ME Lab and the Copenhagen Hearing and Balance Center (CHBC), we realized that children could also benefit from training supported by vibrations. However, as the project progressed and insights were gained through interactions with both children and clinicians, it became evident that vibrotactile feedback was not always the most suitable solution. Consequently, the focus broadened toward music training, which emerged as a crucial element in enhancing the music experience for children with HC, with particular attention to individuals with specific etiologies such as cochlear nerve deficiency (CND) (see Section 2.1).

The main objectives of this project are to investigate how training can enhance the music experience for children with HC and to develop novel solutions applicable in both clinical and domestic settings. The PhD research examines various aspects of training from an engineering and technological perspective, including prototype development, technology, and methodology, and how these factors influence the music experience for children with HC.

Each phase of the project has involved a specifically designed prototype and a case study, with one or more children and their families participating in testing and evaluation.

1.3 Research Methodology

A central tenet of this PhD project has been the active involvement of children with HC and their families throughout the research process. The methodology is grounded in user-centered design with a participatory approach, ensuring that end users and clinicians contribute meaningfully to both the development and evaluation of prototypes and training programs. By integrating the lived experiences of children with the practical expertise of clinicians, the project aims to create solutions that are both relevant and effective. Iterative feedback loops enabled continuous refinement, with each cycle informed by direct input from participants. This approach aligns with the established practices of the ME Lab, which emphasizes multidisciplinary, user-driven research in technology development for specialized populations. To address the unique challenges of this work, iterative design processes and close collaboration among clinicians, children, and their families were employed—including recruitment of young participants, tailoring training to individual needs, measuring the impact of training on musical experience, and ensuring activities remain engaging.

Given the diverse needs of children with HC, as well as the rarity of specific conditions, this PhD project largely adopted a case study design involving single or small groups of participants. These case studies enabled in-depth exploration of individual experiences and provided valuable insights into the effectiveness of the developed solutions; when possible, a test–retest design was incorporated to assess the reliability of findings over time. An in-depth discussion of these methodological aspects is provided in Chapter 2 and in the included papers.

1.4 Parts Involved

Over the course of this PhD project, several partners have contributed in various capacities. Some participated directly in the studies, while others provided support, expertise, and/or access to participants. The project benefited from the collaboration of several institutions and companies, each bringing specific expertise relevant to the research objectives.

The main partner is Copenhagen Hearing and Balance Center, a specialized center within Rigshospitalet, Denmark's leading hospital. The CHBC provides comprehensive care for children and adults with hearing and balance disorders. It offers advanced diagnostics and treatment, including CIs and bone-anchored hearing systems, and emphasizes a multidisciplinary, family-centered approach. It also houses some research facilities, including a spatial hearing lab and neuroimaging tools, enabling research in auditory processing and rehabilitation. CHBC was instrumental in providing access to children with HC and their families, as well as to clinicians who contributed to the design and evaluation of the prototypes. Personnel from CHBC also offered valuable insights into the challenges faced by children with HC, which informed the design of the prototypes and training programs.

Oticon Medical, a company that operates globally in implantable hearing solutions and part of the Demant Group, was involved in project for the first prototype featuring vibrotactile feedback, through the supervision of Marianna Vatti.

Professor Federico Fontana from the University of Udine, whose research focuses on sonic interaction design, haptic feedback, and computational modeling of sound, contributed to the design aspects of some of the studies. Andrea Gulli, also from the University of Udine, collaborated on the systematic review presented in Paper A, bringing his knowledge at hand about assistive technologies and hearing loss (HL).

Dr. Laura Rachmann and Professor Deniz Başkent from db SPL (Deniz Başkent's Speech Perception Lab) provided audiological expertise and contributed to the development of the study on training for emotion recognition. The db SPL, based at the University Medical Center Groningen, focuses on auditory perception, speech processing, and cochlear implant research. The lab investigates a wide range of topics, including vocal emotion perception, music training, and communication in children and adults with hearing impairments. During my PhD, I had the pleasure to undertake a short research stay in Groningen in June 2024, where I worked closely with Laura and Deniz, laying the foundation for the project on training vocal emotion recognition presented in Paper F.

Christa Thomsen from Børnecenter København, a municipal center specializing in child services, supported the recruitment of children with HC.

The center's speech and hearing consultants work closely with children with complex communication needs, providing specialized pedagogical support and expertise in augmentative and alternative communication.

This PhD project was also supported by the Nordic SMC Network (NordForsk fund), which co-funded the project alongside Aalborg University, and the GN Foundation. The NordicSMC brings together leading researchers from all five Nordic countries—Denmark, Finland, Iceland, Norway, and Sweden—and promotes interdisciplinary research in SMC. The network spans the arts, humanities, sciences, and engineering, and supports collaborative projects, conferences, and short-term scientific missions. Its goal is to advance the understanding and development of sound and music technologies through computational methods.

The contribution of each partner has been essential to the development of this PhD project, supporting research progress and enriching my personal knowledge with diverse perspectives and highly valuable expertise.

CHAPTER 2

Background

This chapter aims to provide a comprehensive overview of the current state of the art in the topics that this thesis touched upon and are necessary to have a good understanding of the context in which this PhD project has been developed. Following an introductory section about HL (2.1), it examines the technologies and approaches used to address these hurdles (2.1.3), with a particular emphasis on children (2.1.4). The section also discusses the Danish approach to managing hearing impairment (2.1.5), explores the role of music in the context of HC (2.1.6), and concludes with an overview of the consequences of HL on emotion recognition abilities (2.1.7), a topic not strictly related to music but sharing common perceptual aspects and playing a major role in quality of life (QOL), and relevant for Paper F. The subsequent section presents the multisensory integration process and its relevance to hearing-related challenges and solutions (2.2), which is essential for understanding the rationale behind the use of vibrotactile stimulation in Papers B, C, and D. This is followed by a focus on the training of hearing capabilities, a central aspect of the thesis (2.3) and common to all the prototypes presented. For the same reason, the chapter concludes with a section on gamification and serious games (2.4).

2.1 Hearing and its Challenges

Hearing is the process by which the auditory system detects, transmits, and interprets sounds–fluctuations in air pressure–from the surrounding environment. It is one of the senses that enables us to perceive and make sense of the world around us. Being immersed in a sound-rich environment, makes hearing an effective source of information - ranging from the voice of a loved ones, the rustling of wind through trees, to the intricate soundscape of a symphonic orchestra piece, such as Claude Debussy's *Prélude à l'Après-midi d'un Faune*. Hearing allows us to interpret our surroundings, providing insight into spatial dimensions and the materials that compose them. It helps us estimate the direction and distance of a sound source, and whether it is moving or stationary [93]. Through sound, we communicate, express emotions, share thoughts and ideas, and evoke memories. Although often taken for granted especially in favor of vision (e.g., the McGurk effect [90]), hearing is a complex, delicate, and powerful sense that plays a crucial role in our daily lives and is relied upon continuously.

2.1.1 The Hearing System

The hearing system is a articulated network of structures that work together to detect, transmit, and process sound, and its main components are highlighted in Figure 2.1. It can be divided into three main parts: the outer ear, the middle ear, and the inner ear. Sound waves enter the auditory system through the outer ear, which comprises the pinna and the external auditory canal. The pinna, with its complex arrangement of ridges and grooves, helps collect and amplify high frequency sounds, particularly around 5000 Hz, thereby aiding sound localization. These sounds are then directed into the ear canal, which functions as a resonator due to its tube-like structure. This resonance typically enhances frequencies between 3000 and 4000 Hz in adults, contributing to the natural quality of sound perception [97].

The middle ear, an air-filled cavity within the temporal bone, receives these acoustic signals at the tympanic membrane. Vibrations of the tympanic membrane are conveyed through the ossicular chain–three small bones called malleus, incus, and stapes. This chain amplifies sound energy through three mechanisms: the difference in surface area between the tympanic membrane and the stapes footplate, the lever action of the ossicles, and the buckling

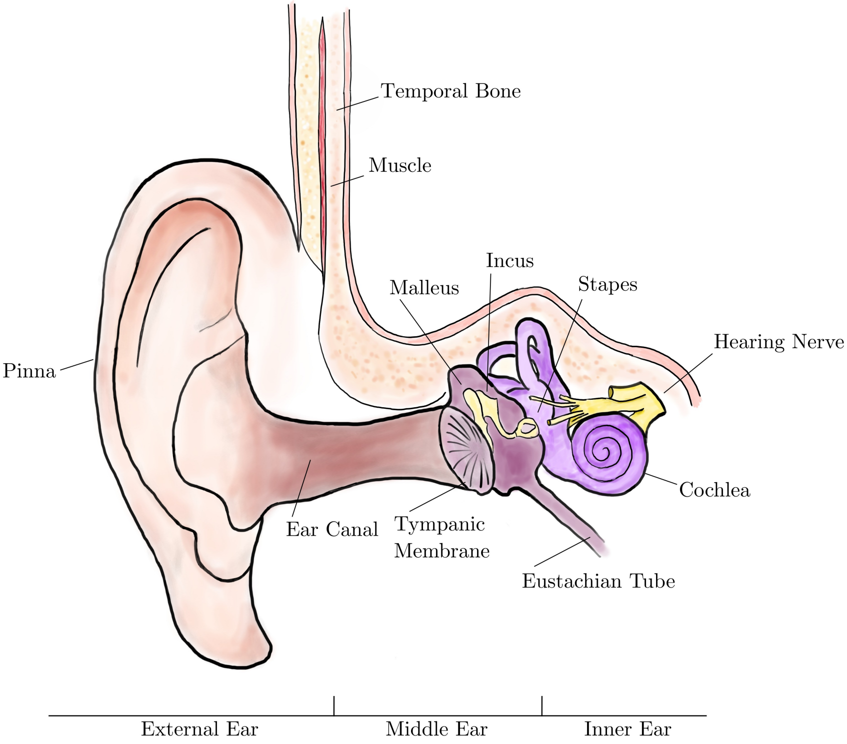


Fig. 2.1: Overview of the human hearing system.

motion of the tympanic membrane. These mechanisms help overcome the impedance mismatch between air in the outer ear and fluid in the inner ear. Additional components of the middle ear include the Eustachian tube, which regulates air pressure, and two muscles–the tensor tympani and stapedius–that contract in response to loud sounds, initiating the acoustic reflex. Ligaments and a branch of the facial nerve also provide structural support to the ossicular chain [97].

Ultimately, the amplified mechanical energy is transmitted to the cochlea in the inner ear, where it is converted into neural signals that travel through the auditory pathways to the brain for processing. The cochlea, plays a pivotal role in transducing mechanical sound energy into neural signals interpretable by the brain. Anatomically, the cochlea is a spiral-shaped, fluid-filled structure embedded in the petrous portion of the temporal bone. Inside the cochlea, the organ of Corti is situated atop the basilar membrane; this organ contains two types of hair cells: inner hair cell (IHC) and outer hair cell (OHC). Inner hair cells are primarily responsible for transmitting auditory information to the brain, whereas OHCs contribute to the amplification and fine-tuning of sound through their electromotile properties. The stereocilia of these hair cells are deflected by shearing forces generated by basilar membrane motion, starting a process which activates the auditory nerve fibers [97].

In a healthy ear, the basilar membrane inside the cochlea acts as a sophisticated frequency analyzer. Its stiffness gradient creates a tonotopic map, where high frequencies peak near the basal end and low frequencies near the apex. Frequency discrimination is achieved through two complementary mechanisms: temporal coding for frequencies < 5 kHz and spatial coding for frequencies ≥ 5 kHz [95]. Outer hair cells further enhance this process by actively amplifying basilar membrane motion and sharpening tuning curves, especially at low sound intensities. This active mechanism improves sensitivity by up to 50 dB and introduces nonlinearity, allowing for amplitude compression and dynamic adaptation across a wide range of sound intensities [95].

The auditory nerve (AN) consists of approximately 30'000 fibers and conveys the electrical impulses generated in the cochlea to the brainstem. Functionally, the AN encodes sound frequency through both place and temporal mechanisms. The characteristic frequency of each fiber corresponds to the frequency at which it responds most sensitively. Intensity is encoded by the firing rate and the number of active fibers [97]. Together, the inner ear and auditory nerve form a highly specialized and finely tuned system for the initial stages of auditory processing, converting acoustic stimuli into precise neural codes for further interpretation by central auditory structures.

In terms of auditory perception, several key aspects can be highlighted: a healthy and young auditory system can detect a wide frequency range, spanning from 20 Hz to approximately 20 kHz [92]. The dynamic range of human hearing is extensive–typically around 120 dB–allowing us to perceive sounds from the faintest whispers to loud explosions [145]. The non-linearity in sensitivity across different sound levels enables the perception of a wide range of sounds without distortion, thanks to the active mechanisms provided by the OHC. This phenomenon is well illustrated by the equal-loudness-level contour graphs (see Suzuki et al. [127] for an example). Frequency resolution, or the ability to distinguish between different frequencies, allows for the detection of frequency differences smaller than 1 Hz in the lower frequency range [92]. Frequency discrimination becomes significantly poorer at higher frequencies, with minimum frequency gap of 120 Hz at 1 kHz [110]. Around 5 kHz, we loose the ability of discerning pitch, and thus we do not recognize melodies above this threshold [110]. Another interesting characteristic is the auditory system's non-linear response to sound intensity, acting like a compressor. As sound level increases, response growth becomes markedly sublinear, with mid-to-high levels exhibiting shallow slopes; for example, a 10 dB increase in input results in only about a 2 dB increase in output. This property supports robust encoding across a wide dynamic range [110].

2.1.2 Hearing Loss

Hearing loss is a common condition that affects millions of people worldwide [139] and can significantly impact an individual's QOL. There are several etiologies of HL, which can be categorized as conductive, sensorineural, or mixed. Conductive HL prevents sound from being transmitted through the outer ear canal to the eardrum and the ossicles of the middle ear. Sensorineural HL involves damage to the inner ear or auditory nerve and can result in permanent impairment. Mixed HL is a combination of both conductive and sensorineural components. Hearing loss can also be classified as mild, moderate, severe, or profound, depending on the degree of impairment. It may be congenital (present at birth) or acquired later in life due to various factors such as aging, exposure to loud noises, infections, ototoxic medications, or genetic predisposition [99]. Over the years, advancements in screening techniques have enabled earlier detection and intervention [109]. The impact of HL can vary widely among individuals, depending on the severity and type of loss, the age of onset, and the individual's overall health.

Hearing loss has wide-ranging effects on individuals of all ages, impacting not only the ability to perceive sounds but also cognitive, emotional, and social well-being. In older adults, it is associated with an increased risk of depression, social isolation, and reduced QOL, as well as a higher likelihood of cognitive decline and dementia, making it one of the most significant modifiable risk factors for these conditions [23, 30, 80, 105]. Even mild HL can increase listening effort, strain memory, and disrupt spoken language comprehension by requiring the brain to recruit additional cognitive resources, potentially leading to fatigue and further cognitive challenges [30, 123]. Structural brain changes, including reductions in gray and white matter, have been observed in both congenital and acquired HL, suggesting that auditory deprivation can affect brain organization across the lifespan [78]. Balance and mobility may also be affected, particularly in complex environments or when other senses are compromised, potentially increasing the risk of falls [24, 98]. Noise-induced HL, a common form of acquired HL, can further contribute to stress, social withdrawal, and diminished QOL [81]. Unilateral HL, or one-sided, impairs spatial hearing, making it difficult to localize sounds and understand speech in noisy environments, which can negatively affect learning, work, and social interactions [14]. Bilateral implantation of CIs has been shown to be insufficient for restoring spatial localization abilities, and older age at testing as well as earlier onset of deafness are associated with greater average localization error [11]. Early detection, rehabilitation, and appropriate interventions are crucial for mitigating these effects and improving overall outcomes for individuals with HL [12, 24, 30].

During the development of this thesis project, sensorineural hearing loss (SNHL) associated with auditory neuropathy spectrum disorder (ANSD) and CND received particular attention. Auditory neuropathy spectrum disorder is characterized by desynchronized transmission of auditory signals from the inner ear to the brain, despite preserved OHC function. This condition is typically identified by normal otoacoustic emissions–indicating functional OHCs–combined with absent or severely abnormal auditory brainstem responses, reflecting impaired neural synchrony [37, 113, 146]. Such disruption can lead to difficulties in speech perception, particularly in noisy environments, and may result in communication and social challenges [113]. In addition, individuals with SNHL may present with a wide range of hearing capabilities on pure tone audiometry, from normal sensitivity to profound HL, without a consistent audiometric pattern. Exposure to high-intensity or prolonged tonal stimuli can lead to auditory fatigue, complicating the assessment of hearing thresholds. Progressive hearing decline is commonly reported across both pediatric and adult populations [37]. Temporal aspects of auditory processing are notably affected in ANSD. Deficits include reduced pitch perception at lower frequencies–suggesting impaired phase-locking mechanisms–and diminished sensitivity to brief acoustic events and silent intervals between sounds. Spatial hearing, particularly the ability to localize sound using interaural time difference (ITD), is also impaired. Additionally, the perception of loudness for continuous tones presented over extended periods (e.g., three minutes) is diminished in ANSD patients, with the extent of this reduction varying across different frequencies [37].

The most common form of ANSD is CND, accounting for approximately 45% of cases [79]. This condition is categorized as either nerve hypoplasia–where the nerve is present but underdeveloped–or nerve aplasia–where the cochlear nerve is absent. Cochlear nerve deficiency can significantly impair auditory processing and speech comprehension [7]. Both ANSD and CND can substantially affect an individual's ability to hear and process sounds, regardless of the use of CIs or HAs, leading to considerable challenges in daily life. The incidence of CND varies across studies but is generally reported in 25%–50% of children with unilateral SNHL and about 1% of those with bilateral SNHL, increasing to 10%–16% among children with severe or profound bilateral HL [26]. Although post-implantation outcomes in children with CND are often less favorable compared to those without the condition, several studies indicate that these patients can still achieve meaningful auditory and speech improvements following CI [26]. Prognosis appears to depend on the anatomical characteristics of the cochlear nerve; for example, cochlear nerve hypoplasia offers a better outlook for sound perception and speech recognition than complete nerve aplasia, as residual fibers may leverage neuroplasticity–the brain's ability to adapt and reorganize in response to sensory changes, injury, or learning [58]–to support auditory function [104].

2.1.3 Technologies and Approaches

In the past decades, significant advancements in hearing technology have led to the development of assistive devices capable of restoring auditory function in individuals with mild to severe HL. These devices have enabled affected individuals to regain access to the auditory world, thereby improving their QOL and facilitating greater participation in social, educational, and professional activities. The most common devices include HAs and CIs. The latter are particularly beneficial for individuals with severe to profound sensorineural

HL who do not gain sufficient benefit from HAs. With both typologies of devices, research shows that timely intervention provides the best hearing outcomes in both children and adults [17].

Hearing Aids

Hearing aids are electronic sound-amplifying devices worn inside or behind the ear. They include a microphone that captures sound, an amplification and digital signal processing (DSP) stage that processes and increases the sound's intensity, and a speaker that delivers the amplified sound to the ear. The input captured by the microphone is low-pass filtered at 9 kHz, and then band-limited signal is processed applying frequency-dependent amplification, amplitude compression and limiting, noise reduction, cancellation of acoustic feedback, and directional processing [111]. The dynamic range of HAs is generally up to 96 dB, which is constrained by the 16-bit precision of the analog-to-digital converter (ADC) integrated into the processor [111]. Hearing aids can be tailored to an individual's specific HL profile, enhancing speech understanding and overall auditory experience [111]. They are particularly effective for individuals with mild to moderate HL: the whole chain still rely on the ability of the ear to detect, convert and transmit the treated and amplified signal to the brain through the auditory nerve.

To be eligible for a HA in Denmark, individuals are evaluated based on their pure tone audiometry (PTA) thresholds at 1, 2, 3, and 4 kHz. If the average threshold of HL exceeds 45 dB, they are considered candidates. Guidelines also recommend including functional hearing assessments, which should account for more realistic listening environments. The frequency range of commercially available HAs typically spans from 200 Hz to 8 kHz and includes dynamic range compression features to better accommodate the specific needs of users [71].

Cochlear Implants

Cochlear implants are surgically implanted devices designed to bypass the outer and middle ear, directly stimulating the auditory nerve. They consist of an external processor that captures sound and converts it into electrical signals, which are transmitted via an antenna to an internal receiver implanted beneath the skin (see Figure 2.2). The internal component then decodes the received signal and generates the electrical stimuli that are sent to electrodes placed within the cochlea, enabling sound perception by stimulating the auditory nerve [87, 142]. Depending on the model, CI contain between 12 and 24 electrodes, which divide the audible spectrum into discrete bands that stimulate different regions of the cochlea [144]. This configuration represents a significant simplification compared to the natural auditory system, which comprises approximately 30,000 afferent nerve fibers–each connected to an IHC region tuned to a specific frequency [95]. This stark difference highlights the degree of information compression inherent in cochlear implant technology. A schematic of the main components of a CI is shown in Figure 2.2.

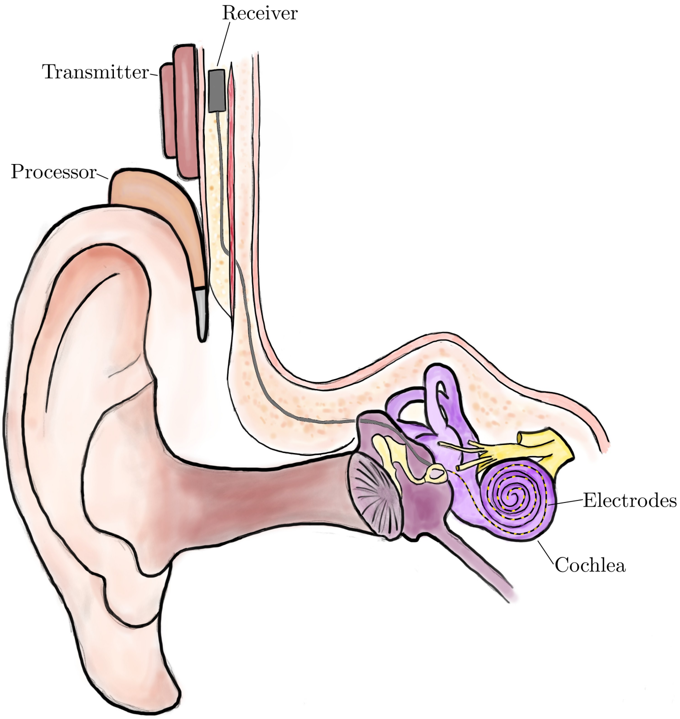


Fig. 2.2: Schematic representation of a cochlear implant and its main components: external processor, transmitter, internal receiver, and electrode array.

Cochlear implants leverage the ear's tonotopic map to stimulate specific frequency regions within the cochlea (coding by place); the lower the stimulus frequency, the closer the stimulation occurs to the apical (deeper) region of the cochlea [142]. A second mapping strategy for frequency coding is based on rate, that is, the frequency of the pulses delivered to the electrodes. Research has shown that the average maximum rate of pulses perceivable by CIs is approximately 300 Hz. Frequencies above about 300 Hz are represented by the site of stimulation (coding by place), whereas frequencies below about 300 Hz are represented by temporal variations in the modulation waveforms (rate) [142]. The amplitude information of the incoming signal is conveyed through pulse intensity. The dynamic range of CIs is typically around 60 dB, which is lower than that of HAs or the natural ear ($\approx$ 120 dB) [145]. This limitation affects the implant's ability to replicate the fine-grained auditory resolution and dynamic responsiveness of typical hearing (TH), underscoring the technological challenges in restoring complex auditory functions. Similarly to a vocoder [85], each electrode corresponds to a distinct channel that analyzes a specific frequency band of the incoming signal. Within each channel, the signal undergoes band-pass filtering, followed by envelope detection to extract its energy profile. This processed signal is then transformed through a nonlinear function and used to modulate a sequence of balanced biphasic pulses. This architecture ensures that stimulation parameters are tailored to the spectral characteristics of the input [142]. Although current CIs on the market provide up to 24 electrodes, existing strategies for electrode placement and signal processing do not appear to offer significant benefits beyond the use of more than eight channels [87, 142]. The electrical stimuli delivered to the cochlea activate relatively wide areas due to electrode placement, stimulating broader regions compared to the precise activation mechanism of IHCs in TH subjects [87].

In Denmark, eligibility for a CI requires a HL of at least 75 dB in the implanted ear, as determined by PTA thresholds at 1, 2, 3, and 4 kHz. Additionally, candidates must demonstrate limited benefit from HAs, defined as $\leq 50\%$ sentence recognition in the ear to be implanted. Despite these criteria, implantation decisions are made on a case-by-case basis by a multidisciplinary team, reflecting the complexity and individuality of HL.

2.1.4 Implications for the Children

In children, the implications of HL are comparable to those observed in adults; however, because childhood represents a critical period for the development of language, communication, and social skills, the impact of HL can be even more profound. If untreated, it is associated with adverse outcomes across multiple domains, including speech and language acquisition, educational achievement, social interaction, cognitive development, and overall QOL [80].

Timely interventions, can promote the achievement of developmental milestones in auditory perception that are comparable to those of their TH peers in children with severe-to-profound HL. Research indicates that early implantation supports the development of horizontal sound localization abilities at a similar age to TH children, suggesting the presence of intrinsic mechanisms guiding this aspect of auditory development [12]. Language development in children with CI shows promising progress within the first four years post-implantation, with performance increasingly resembling that of TH peers. However, between four and six years after implantation, challenges emerge in specific language domains, notably receptive vocabulary and expressive grammar. These findings underscore the necessity of sustained language intervention to support continued development and highlight the importance of comprehensive long-term monitoring [136]. The importance of tailored speech rehabilitation strategies–such as auditory-verbal therapy (AVT)–has been widely emphasized to enhance speech perception development in children with CIs, particularly during the early stages of auditory learning [131]. Notably, cochlear implantation alone does not guarantee age-appropriate speech and language outcomes, even when performed early and bilaterally [108].

In hearing disorders, compensatory neuroplasticity supports adaptation and enhances outcomes from interventions like CI [10]. Early childhood, particularly up to 3.5 years of age, represents a sensitive period of heightened plasticity [58, 126]. This phase is critical for auditory pathway development, which depends on early language exposure and social interaction [126]. Early cochlear implantation and auditory-verbal habilitation can promote age-appropriate language leveraging neuroplasticity [108]. Hearing deprivation leads to two distinct forms of cortical reorganization and it is a characteristic of HL [41]. Cross-modal plasticity occurs when intact sensory modalities, such as the visual or somatosensory systems, recruit auditory cortical areas to compensate for the missing input. While this reorganization may support visual speech processing, it can interfere with auditory outcomes in CI users [10, 58]. Intra-modal plasticity refers to reorganization within the auditory system itself. For example, in cases of unilateral HL, altered hemispheric activation patterns can emerge, leading to reduced interhemispheric asymmetry and impairments in spatial hearing [10].

Auditory deprivation and degraded CI input also affect neurocognitive functions. Children with CIs show persistent deficits in verbal working memory, which correlate with language outcomes [13]. The auditory neurocognitive model links reduced auditory experience to delays in executive functions and phonological processing [13]. Interventions leveraging neuroplasticity include auditory training (AT), music-based therapy, and bimodal stimulation [13]. Encouraging speech production and sensorimotor interaction further supports language development. Effective early intervention can reverse maladaptive reorganization and optimize outcomes [10, 58].

2.1.5 The Danish Way

When it comes to healthcare, it is essential to consider the cultural and societal context in which it is delivered, as well as the structure of the local healthcare system. This paragraph provides an overview of the Danish approach to addressing hearing challenges, which is particularly relevant to the studies presented in this thesis, as they were conducted within this context. Acknowledging the complexity of the topic–particularly in relation to sociocultural aspects of disability and community belonging–this section only offers a documented perspective on the circumstances under which this project was conceived and developed. The following paragraph is an overview of the landmarks in educational planning for deaf and hard of hearing (d/hoh) children in Denmark over the past few decades.

In 1993, the first children in Denmark underwent CI surgery for deafness [107], marking the beginning of a new era in hearing rehabilitation. By 2005, universal neonatal hearing screening was implemented, enabling the diagnosis of HL by three months of age and intervention by six months [76]. This milestone significantly advanced early identification and intervention efforts in Denmark. However, early implantation alone was not sufficient to ensure age-equivalent language performance, raising questions about the education system for children with CIs [106]. In 2012, the National Board of Health recommended AVT as the standard rehabilitation practice for children with HL [76]. Three years later, in 2015, the Danish National Board of Social Affairs extended this recommendation to include AV practices for the (re)habilitation of children with HL, regardless of the type of hearing technology used [66]. To minimize socio-economic influences, all medical, technological, and AVT interventions are provided at no cost to children with HL [76]. By 2020, most d/hoh children in Denmark were being educated in local community schools using AVT, with minimal or no use of sign language–marking a shift from the previously applied sign language-based model. This transition has sparked ongoing debate, particularly within Deaf communities, where cochlear implants are sometimes perceived as a threat to cultural identity. Nonetheless, the oralist model has become dominant in both policy and practice [35], supported by scientific evidence.

Research indicates that $90–95\%$ of deaf children have hearing parents [94], who are often motivated to enable communication with their children in their native language (i.e., mother tongue) [107]. This aspect is also reflected in studies where deaf children report difficulties communicating with their hearing parents [107]. The inability to share a fluent language with one's parents can hinder social-cognitive development [126]. These are among the driving reasons why AVT has been chosen by families with children with HL as a preferred intervention approach, eventually becoming the standard adopted by the Danish healthcare system.

This short overview serves as a memento for both myself and the reader that, when working with families, children, or people in general, we must remain mindful of the context in which we operate. This includes cultural, social, and individual factors that may shape their experiences, expectations, and needs.

2.1.6 Music and Hearing Challenges

Music encompasses a wide range of auditory experiences, from simple melodies played on a single instrument to complex compositions performed by a symphonic orchestra. It is characterized by elements such as pitch, rhythm, timbre, and dynamics. From a signal perspective, music exhibits an extremely rich

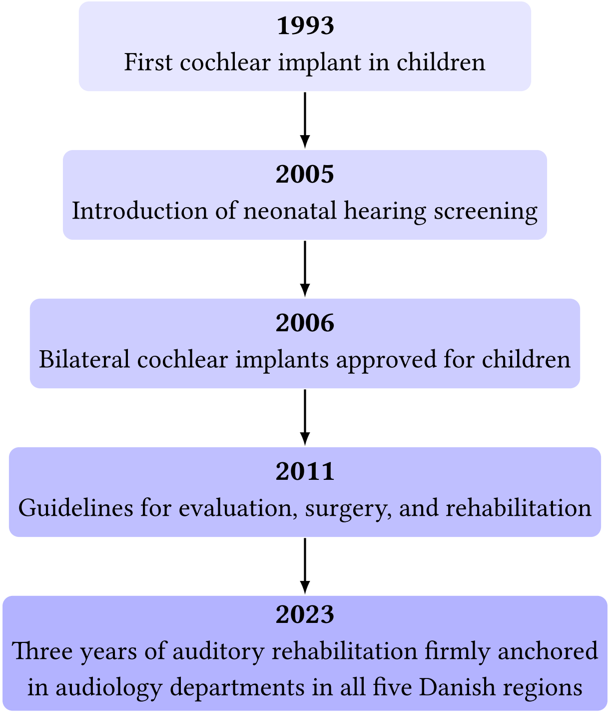


Fig. 2.3: Timeline of major milestones in pediatric cochlear implantation and auditory rehabilitation in Denmark. Adapted from [107].

spectral and temporal structure, often spanning the entire human audible frequency range (20 Hz to 20 kHz) and, at times, the full dynamic range (0 to 120 dB). The challenges associated with music perception in individuals with HL are deeply rooted in the anatomical and physiological structures of the auditory system. As outlined in Moore's review [96], dysfunction of the OHCs leads to an elevation of the absolute threshold, meaning that the quietest sound detectable by the listener becomes louder. This elevation not only affects basic audibility but also compromises the ability to resolve fine spectral details in complex acoustic signals. Specifically, individuals may struggle to perceive distinct frequency components and to identify the spectral shape of sounds, which are essential for recognizing musical timbre and harmony. Additionally, pitch clarity is often degraded, and the dynamic range–the span between the softest and loudest perceivable sounds–is significantly reduced. In contrast, impairments involving IHCs, synaptic transmission, and neural pathways primarily disrupt pitch perception and sound localization. These deficits are particularly detrimental to music listening, where accurate pitch discrimination and spatial awareness contribute to the richness of the experience. Hearing technologies attempt to address some of these limitations. Hearing aids, for instance, employ automatic gain control to compensate for elevated thresholds and use directional microphones to improve signal-to-noise ratio in complex listening environments. However, these strategies do not effectively restore the auditory abilities required for music appreciation. Similarly, CIs incorporate automatic gain control but are constrained by a reduced dynamic range intrinsic in the device's technology, which limits their ability to convey the subtle nuances of musical expression. Furthermore, current spread along the cochlea–where electrical stimulation activates broader regions than intended–compromises frequency selectivity. This phenomenon makes it difficult for users to distinguish closely spaced pitches, resulting in poor pitch perception. Reported consequences include difficulty identifying melodic contours, recognizing harmonic structures, and detecting consonance and dissonance. Despite these limitations, rhythm perception tends to remain relatively intact, suggesting that temporal processing is less affected by the anatomical constraints of HL and cochlear implantation.

Beyond anatomical limitations, individuals with HL encounter a range of perceptual and psychosocial challenges that influence their musical experiences. Even though the value of music remains undiminished, their ability to engage with it meaningfully is often compromised. Several factors influence music perception in this population, including the etiology of HL, the degree of residual hearing, the duration of auditory deprivation, and individual cognitive and physiological differences [4]. These variables interact to determine how music is perceived, interpreted, and enjoyed. Although HAs can enhance access to musical sounds, they do not fully restore the richness of the experience to the level enjoyed by TH individuals [8]. Moreover, the subjective music enjoyment tends to decline as HL becomes more severe [8]. For users of CIs, the limitations of the device can even result in music having a negative impact on overall QOL [4]. Interestingly, despite these perceptual challenges, many CI recipients continue to regard music as an important and meaningful aspect of life [4], underscoring the resilience of music's emotional and cultural significance.

In a UK-based study by Greasley et al. [62] involving both HA users and audiologists, 58 % of HA users reported never having discussed music listening with their audiologist. Additionally, half of the audiologists indicated that they initiated such conversations with only one in five patients or fewer. A larger proportion–72%–noted that fewer than one in five patients raised the topic themselves. These findings highlight the rarity of music-related dialogue in clinical settings and the inconsistent approaches to programming HA for music. The study also revealed that even minimal exposure to training–such as attending a single professional development session or conference–can positively influence clinical practice by increasing audiologists' confidence and the likelihood of tailoring HAs for music. Among the most frequently cited challenges were difficulties with pitch perception, restricted dynamic range, and reduced lyric clarity, making music listening problematic for approximately two-thirds of participants.

Adult CI users often report that their musical experiences have improved following implantation, particularly in terms of access and clarity. However, many still find music listening to be disappointing [40], suggesting that technological improvements alone are insufficient. This dissatisfaction highlights the need for personalized interventions and comprehensive clinical assessments that address both perceptual and emotional dimensions of music engagement [4]. Commonly cited challenges include the complexity and familiarity of musical material, with background music frequently described as a barrier to effective communication [56]. When listening demands exceed an individual's auditory capacity, such situations may become overwhelming, leading to withdrawal from social and musical activities. To counteract these effects, rehabilitation and counseling strategies should consider the listener's problem-solving orientation and motivational factors that promote active engagement. Social connection, cultural relevance, and personal preference in musical content can serve as powerful motivators [56]. Additionally, research has shown that musical training may enhance auditory processing under challenging conditions. Musicians, for example, outperform non-musicians in speech perception tasks involving two-talker maskers and demonstrate more efficient lexical decision-making in noisy environments [73], suggesting that musical expertise can confer cognitive advantages in complex auditory scenes.

A comprehensive systematic review by Bleckly et al. [16] further emphasizes the scope of these challenges. Analyzing 131 studies, the authors found that research on music and HL in adults predominantly focuses on perceptual accuracy, especially among CI users. Pitch, timbre, and melody recognition emerged as the most problematic domains, while rhythm perception remained relatively preserved. Despite this, overall musical fidelity was found to be poor. Crucially, the review identified a significant gap in the literature: few studies address music appreciation, emotional engagement, or active participation, even though music is widely recognized for its psychosocial and cultural importance. This narrow focus limits the development of holistic rehabilitation strategies. The authors conclude that while hearing devices are effective in restoring speech perception, they fall short in restoring music perception, which can negatively affect emotional well-being and social connectedness. Given that CIs and HAs are primarily optimized for speech, music perception remains significantly degraded. The review calls for future research to explore the broader psychosocial role of music, develop targeted rehabilitation programs, and improve device technology to enhance musical fidelity.

Improving the music experience for individuals with HL has long been a challenge, due to the complex interplay between anatomical factors, music perception, and technological limitations. Given that each individual is, by definition, unique also in their hearing profile, solutions that are effective for one person may not be suitable for another. Researchers and clinicians have identified training as a potential key to enhancing the music experience for people with HL. Section 2.3.2 provides a comprehensive overview of training programs and their efficacy.

2.1.7 Emotion Recognition Abilities

An important element of the social context is the capacity to perceive and interpret emotions conveyed through speech. This ability is crucial for successful communication and social engagement, as it allows individuals to respond appropriately to others' emotional states. In this PhD project, this topic is particularly interesting, as the vocal emotion recognition shares many acoustic features with music perception.

Vocal emotion communication is primarily conveyed through systematic variations in pitch, represented by fundamental frequency (F_0), as well as loudness, voice quality, and timing. High-arousal emotions such as anger, fear, and happiness typically exhibit elevated F_0 and intensity, whereas sadness is characterized by lower values [72]. Additional cues–such as high frequency energy, formant shifts, and speech rate–signal vocal tension or relaxation, further differentiating emotional states. These cues combine into emotion-specific acoustic patterns that allow listeners to infer both the type and intensity of emotions [72]. Without the use of HAs, individuals with HL face challenges in discerning these critical acoustic cues, leading to difficulties in accurately perceiving and interpreting vocal emotional content in speech and nonverbal vocalizations [42]. Even with assistive devices, the performance of individuals with HL in vocal emotion recognition tasks remains inferior to that of TH individuals [42]. Adults with bilateral HAs have been shown to exhibit lower identification accuracy and longer reaction times to vocal emotional stimuli compared to their TH peers [27]. Unsurprisingly, difficulties in recognizing vocal emotions are also experienced by CIs users. Vocal emotion recognition is impaired in both adults [134] and children with CIs [67]. These challenges extend to emotion recognition in music, where both CI and HA users rely more on tempo information than on F_0 or harmonic content to identify emotions [96]. Furthermore, congenital HL has been found to exert a more detrimental impact than acquired HL on emotion recognition in music [96]. Music exposure during early developmental stages with TH appears to facilitate the acquisition of emotion recognition skills in music, which persist even after the onset of HL [96].

Literature suggests that age is one of the strongest correlates of vocal emotion recognition ability in speech. In TH individuals, these abilities significantly improve with age, showing increased accuracy and more nuanced attribution of emotions in prosody from childhood to adolescence, even when the stimuli are linguistically meaningless [44]. A recent study by Rachman et al. [112] found that children with HL also exhibit developmental improvements, although these are less pronounced than in their TH peers. Notably, performance did not significantly differ between the two groups until the age of eight, with some individuals performing at age-typical levels and others falling below expectations–even at chance levels. Further complicating the picture, neither aided nor unaided hearing thresholds were found to predict vocal emotion recognition performance [112]. The authors hypothesize that additional factors–such as cognitive abilities, early auditory exposure, and language experience–may have had an effect on these outcomes. This highlights the high degree of variability among individuals and underscores the importance of considering individual differences when assessing vocal emotion recognition abilities in children with HC. In this complex context, the role of screening and targeted interventions using ad-hoc tools becomes crucial in supporting children with HL in developing their vocal emotion recognition skills.

2.2 Multisensory Integration

Everyone perceives the world not through a single sensory modality, but by integrating information from multiple senses. For example, when analyzing the characteristics of a fruit, we use sight to observe its color and shape, smell to detect its aroma, touch to assess its temperature and texture, and taste to experience its flavor. Unless we are particularly attentive, we typically integrate two or more senses to form a holistic experience without consciously realizing it. This process of combining information from different sensory modalities is known as multisensory integration, and it is fundamental for the brain to construct a coherent representation of the external world.

Multisensory integration is crucial for various cognitive functions, including perception, attention, learning, and memory. By integrating inputs from multiple senses, the brain enhances the accuracy and reliability of its interpretations, leading to improved decision-making and adaptive behavior [100]. It also helps resolve ambiguities that may arise from relying on a single sensory modality, thereby enriching the overall perceptual experience [100].

One of the pioneering studies in this field was conducted by Stein and Meredith [124]. They demonstrated that when inputs from different modalities are spatially and temporally aligned, the resulting neural response is often greater than the sum of its parts–a phenomenon known as superadditivity. This enhancement is not arbitrary; it follows specific principles, such as the spatial rule (stimuli must originate from the same location) and the temporal rule (stimuli must occur close together in time). These rules ensure that the brain prioritizes signals likely to originate from the same external event. The authors also showed that these integrative processes are reflected in behavior and that multisensory integration develops through experience. This behavioral relevance is further illustrated by findings that congruent environmental sounds can enhance memory for object locations–even when those sounds lack explicit spatial cues–highlighting the role of associative mechanisms in strengthening visual memory through multisensory input [88]. Early in life, the brain learns to associate sensory inputs based on their co-occurrence, gradually refining its integrative capabilities [124]. This developmental perspective provides insight into how multisensory processing can be shaped, disrupted, or enhanced across the lifespan.

Multisensory integration doesn't happen in just one part of the brain or follow a single pathway. A recent review by Gao et al. [54] combined data from 121 brain imaging studies and found that audiovisual integration uses a wide network of regions. These include early sensory areas, subcortical structures like the thalamus, and higher-level association areas. Among these, the superior temporal cortex stood out as a key hub, playing an important role in coordinating multisensory processing. The study also showed that integration changes depending on the situation–such as the type of analysis, the complexity of the stimulus, and where attention is focused. This supports a flexible model where integration can happen at different levels of the brain depending on the task and sensory input. This view fits with the idea that experience shapes multisensory integration and highlights the need to consider both bottom-up and top-down influences when studying how the brain combines information from different senses.

2.2.1 Touch and Sound

In this PhD project, the interaction between touch and sound is of particular interest due to its potential applications in auditory rehabilitation and assistance for individuals with HL. By leveraging the principles of multisensory integration just outlined, the auditory dimension may be enhanced through targeted tactile feedback. To better understand the potential of this specific sensory integration, it is essential to briefly examine the underlying mechanisms of the somatosensory system.

The Somatosensory System

The somatosensory system, which processes tactile information, is a complex network comprising various receptors in the skin, muscles, and joints, along with neural pathways that transmit this information to the brain. The largest organ of the somatosensory system–and of the entire body–is the skin [132]. The skin is divided into two main types: glabrous skin, which is hairless and found on the palms of the hands and soles of the feet, and hairy skin, which covers the remaining part of the body. Glabrous skin is densely populated with mechanoreceptors that are highly sensitive to tactile stimuli and commonly targeted by haptic feedback devices, such as those used on the fingertips. It is innervated by four primary types of mechanoreceptors: Merkel cells, Meissner's corpuscles, Ruffini endings, and Pacinian corpuscles [132], each specialized to detect distinct aspects of tactile stimuli. The sensations arising from the activation of these receptors contribute to our perception of touch and encompass multimodal experiences such as light and deep touch, pressure, vibration, itch, tickle, and other complex sensory blends [132]. Among the various types of sensory input that can be delivered through the somatosensory system, haptics represent the most general category. L. Jones defines haptics as "[...] 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" [15]. Within haptics, vibrations can be identified as a specific type of tactile stimulus, with vibrotactile referring to the sensation produced by applying vibrations to the surface of the skin [15].

Vibrotactile stimuli can convey various types of information, including frequency, amplitude, duration, and temporal patterns—properties typically used to describe sound. Depending on the frequency of stimulation, different sensations can be elicited. Low frequencies (up to 4–8 Hz) are typically perceived as slow up-and-down movements, mid frequencies (up to 40–60 Hz) produce fluttering or wobbling sensations, and higher frequencies are experienced as buzzing or diffuse vibrations [84]. Russo et al. [116] investigated whether humans can distinguish musical timbre using only vibrotactile input. Across five experiments, participants received complex vibrations through a chair equipped with voice coils and judged whether pairs of tones were the same or different. The tones included real instrument samples (cello, piano, trombone) and synthesized tones differing only in spectral centroid (brightness), with rigorous masking to eliminate auditory cues. Both hearing and deaf participants performed well above chance, even when fundamental frequency, duration, and amplitude envelope were controlled, indicating sensitivity to spectral content. These findings suggest that frequency-tuned skin receptors function similarly to auditory critical bands, filtering complex tones into components and enabling tactile perception of timbre. This work highlights the potential of vibrotactile feedback in assistive music technology, multimodal composition, and communication systems for individuals with HCs.

Interaction Between Touch and Sound

When comparing human tactile and auditory capabilities, several similarities and key differences emerge. Merchel et al. [92] conducted a psychophysical study comparing frequency discrimination in both modalities. They found that while tactile and auditory systems share frequency-dependent perception thresholds and similar temporal resolution, they differ markedly in frequency range, dynamic range, and sensitivity to aging. Vibrotactile perception is restricted to lower frequencies and exhibits a steeper growth in perceived intensity with level, whereas auditory perception spans a broader range and offers greater dynamic flexibility. The highest perceivable frequency is still debated, since previous studies report values up to 1 kHz [92], but most recent literature highlights a limit around 2 kHz [25]. Age-related decline affects both systems but is more pronounced in the tactile domain at higher frequencies, potentially reducing the effectiveness of vibrotactile feedback for older users. Additional differences include masking effects, frequency discrimination, and spatial localization—critical factors for designing effective auditory-tactile systems. For example, the auditory just noticeable difference in frequency (JNDF) for sounds longer than 200 ms is typically below 1 Hz at low frequencies and increases with frequency, whereas tactile JNDF ranges from 4 to 100 Hz across 20-250 Hz, with some variability detected across studies. The highest sensitivity for hearing occurs around 2–5 kHz, while tactile sensitivity varies depending on the body location [84, 92], with wrist peak sensitivity near 160 Hz [31], and hand (thenar eminence, i.e. under the thumb) around 250 Hz [133]. These findings underscore the need to account for modality-specific characteristics when developing multisensory systems integrating touch and sound.

Considering the applications of touch-sound integration, several studies have demonstrated the potential benefits of combining these modalities across various auditory tasks. One of the most extensively investigated areas is speech intelligibility, where tactile cues have been shown to enhance speech comprehension [29, 63, 117, 128]. Additionally, vibrotactile stimulation has been found to improve pitch perception [86], timbre recognition [116], and localization abilities [61]. Beyond perceptual enhancements, tactile feedback has also been shown to increase immersion in virtual environments and multimedia experiences [122]. These findings underscore the versatility and effectiveness of touch-sound integration in augmenting auditory perception and experience across a range of contexts.

Another body of literature has focused on the mechanisms underlying the interaction between sound and touch. Crommett et al. [34], for instance, investigated how auditory adaptation influences tactile frequency discrimination. They found that adapting to an auditory stimulus improved tactile frequency discrimination when the auditory and tactile frequencies overlapped. This crossmodal effect was feature-specific, as adaptation to a particular auditory frequency enhanced tactile discrimination only within that frequency range, suggesting shared neural mechanisms for processing frequency information across modalities. In another study, Crommett et al. [33] explored how auditory frequency sweeps affect tactile judgments of sweep direction. They demonstrated that auditory sweeps biased tactile perception in a manner dependent on the depth of the auditory sweep, with stronger biases observed for deeper sweeps. Notably, this effect was absent when the auditory and tactile stimuli occupied non-overlapping frequency ranges, and intensity sweeps did not produce similar effects, indicating that multisensory integration is governed by feature-specific interactions rather than general decisional biases. These studies highlight the intricate neural interplay between auditory and tactile systems, emphasizing their functional coupling in encoding temporal frequency information.

Building on this line of research, Darki et al. [36] examined how tactile stimulation influences auditory stream segregation, a fundamental process for parsing complex acoustic scenes. Through psychophysical experiments and computational modeling, they demonstrated that vibrotactile pulses synchronized with a single tone sequence (either high or low frequency tones) in an interleaved auditory pattern enhanced perceptual segregation, whereas pulses aligned with both tone sequences promoted integration. Their neuromechanistic model suggests that tactile input modulates tonotopic responses in the auditory cortex, sharpening contrast when paired with one tone and smoothing activity when paired with both. These findings reveal that temporal alignment of tactile cues can dynamically bias auditory organization, highlighting a spatiotemporal mechanism of multisensory integration with potential applications in assistive technologies for improving speech perception in noisy environments.

2.2.2 Vibrotactile Technology

To complete this brief overview of the multimodal interaction between sound and touch, it is essential to highlight some technical aspects and practical implications that support the theoretical and experimental findings discussed thus far. A moderate but growing body of research focuses on designing and developing prototypes that integrate touch and sound to enhance auditory experiences. These prototypes often employ vibrotactile actuators to deliver tactile feedback accompanied with auditory stimuli, aiming to improve or complement various aspects of sound perception.

Actuators' Overview

Kim et al. [77] describe the most common actuator technologies used in this context. Electromagnetic actuators are widely adopted due to their simplicity, versatility, and ease of integration. Two common types are eccentric rotating mass (ERM) and linear resonant actuator (LRA). An ERM consists of a DC motor with an off-center mass attached to its shaft; when activated, the rotation of this mass generates centrifugal forces that produce multidirectional vibrations on the skin. Although ERM actuators can deliver strong vibrations, their torque and speed are coupled, limiting independent control of amplitude and frequency and resulting in a narrow optimal frequency band. In contrast, LRA devices operate using voice coils and magnets suspended on springs, similar to loudspeakers. These actuators generate linear vibrations normal to the skin and allow more precise control over amplitude and frequency within a resonance bandwidth. Beyond electromagnetic designs, other technologies include piezoelectric actuators, which deform under applied voltage to produce high frequency vibrations; electrostatic actuators, suitable for thin and flexible interfaces; pneumatic systems, which use air pressure for tactile stimulation; and shape memory alloys, which exploit phase changes to generate motion. Each technology offers distinct mechanical characteristics and interaction dynamics with the skin, influencing haptic performance and perceptual outcomes. Paisa et al. [102] found that the two most commonly used actuators in auditory-tactile for applications targeting individuals with HL are ERM and LRA. From their review, we can also notice that the most targeted body locations for delivering vibrotactile feedback are the fingers, hands, and the wrist, likely due to their high tactile sensitivity and ease of access, and the preferred amount of actuators for these locations is low, spanning from one to six. These findings are also consistent with the overview provided by Flores Ramones et al. [48].

Mappings

A fundamental aspect of designing effective auditory-tactile systems is the selection of appropriate vibrotactile signals. In the context of multimodal integration between touch and sound, the term “mappings” is commonly used to describe the relationship between auditory features and tactile stimuli, or in other words how the auditory stimuli are translated to vibrations (and vice-versa). Effective mappings are essential to ensure that tactile feedback is intuitive and enhances the auditory experience.

One of the simplest approaches, particularly common in earlier work, is the so-called “pitch-to-position” mapping, which involves a spatial arrangement of multiple actuators to represent different pitch levels [102]. This method has been utilized in different studies, such as the one by Karam et al. [74], where they reproduced the tonotopic organization of the human cochlea (see Section 2.1.1) using an array of eight voice coil actuators embedded in a chair. They are arranged so that low frequencies are provided near the lower back and high frequencies near the upper back. Two mapping strategies are employed: one divides the audio into bands, such as octaves, and assigns each band to a specific vibrotactile channel; the other uses separate instrument tracks when available. This design preserves timing and spatial cues, reduces masking effects, and enhances tactile resolution, enabling users–especially those who are Deaf and Hard of Hearing (D/HOH)–to perceive emotional aspects of music through touch. On the other hand, more actuators increase hardware complexity and costs, so a balance must be struck between perceptual benefits and practical constraints.

A more articulated approached from a DSP point of view, is the one proposed by Fletcher et al. [45], where a they use a tactile vocoder that transforms speech signals into vibrotactile patterns optimized for the skin’s sensitivity range. The process begins by filtering the audio into multiple frequency bands using a filterbank. For each band, the amplitude envelope is extracted and optionally enhanced through multi-band expansion to emphasize salient speech cues such as formants and rapid onset changes. These envelopes then modulate the amplitude of a set of fixed-phase vibrotactile carriers, each tuned to distinct low frequency tones suitable for tactile perception. This mapping preserves temporal dynamics and distributes spectral information across several tactile channels, aiming to improve phoneme discrimination and robustness in noisy environments while remaining computationally efficient for real-time wearable devices [45].

Another mapping approach utilizes a double-sideband suppressed carrier (DBSC) modulation scheme proposed by Park and Choi [64]. In this method, the audio signal is used to modulate the amplitude a high frequency carrier signal, which is then delivered to a single vibrotactile actuator. This technique allows the tactile system to convey complex auditory information through variations in vibration intensity, potentially enhancing the perception of pitch.

Other complex mappings have been explored over the years however, in many cases, no signal processing is applied to the sound signal before it is transmitted to the actuators [102]. One notable example of sound signal to vibrotactile feedback is the study by Fontana et al. [49], where bass guitar notes were recorded and paired to form melodic intervals classified as unison, consonant, or dissonant. These audio signals were reproduced as vibrations using tactile transducers mounted on a chair and a floor platform. The original notes were converted into vibrotactile stimuli without additional processing, preserving their timing and pitch differences. Interestingly, the study proved that users could distinguish between consonant and dissonant intervals in a similar manner as happens with the auditory channel, with consonant intervals generally easier to identify than dissonant ones [49].

Among the various approaches to designing mappings, Aker et al. [6] sought to identify which audio-tactile congruences are most important for enhancing music listening experiences through vibrotactile stimulation. They found that, in individuals with TH, congruence in intensity and timing between auditory and tactile stimuli is crucial for enhancing music perception; when disrupted, it can significantly reduce the perceived quality of the auditory stimulus. In a more recent study, Aker et al. [5] demonstrated that while TH listeners consistently preferred music accompanied by haptic stimulation congruent in timing and intensity, only a subset of CI users shared this preference. The variability in responses suggests that factors such as tactile sensitivity, musical experience, and individual perceptual integration influence the degree of benefit. Notably, congruent timing and intensity were more impactful than frequency congruence, underscoring the need for personalized and adaptable haptic systems. These findings highlight the potential of multisensory approaches in auditory rehabilitation while also emphasizing the complexity of designing universally effective vibrotactile music devices.

2.2.3 Vibrotactile Feedback and Hearing Loss

After exploring the hearing system and the challenges posed by HL in Section 2.1, and reviewing the fundamentals of multisensory integration and the interaction between touch and sound in Section 2.2.1, we can now examine how vibrotactile feedback can assist individuals with HL.

Vibrotactile feedback has been investigated for decades as a means of supporting communication and perception in this population. One of the earliest notable examples is the Tickle Talker™, introduced in the 1980s [32]. This system consisted of a lavalier microphone, a handset, and a speech processor. The handset featured eight electrodes arranged in four rings, positioned to align with the digital nerve bundles along the sides of the fingers. Users received information through electrocutaneous stimulation of these nerve pathways, enabling the development of a tactile-only vocabulary through training [52]. Interestingly, the original Tickle Talker™ processor was derived from the first-generation Nucleus 22-channel cochlear implant processor [52, 53], highlighting technological parallels in the way sound has been mapped to tactile stimuli and the tonotopic approach used in CIs. Today, research has shifted toward vibrotactile motors rather than electrocutaneous stimulation, making vibrotactile feedback the primary method for conveying tactile information [48].

In general terms, vibrotactile feedback is being explored as a means to enhance aspects of auditory perception compromised by HL. As discussed in Section 2.2.2, the hands and fingers are the most common target locations for delivering vibrotactile stimulation, which explains why many early prototypes focused on these areas. For example, Shin et al. [118] introduced a vibrotactile glove that encodes musical tones into tactile patterns using nine motors spanning three octaves (C3-B5). Nine CIs and HAs users trained over ten sessions with programs for tactile tone learning, sing-a-tone, and melodic singing. Post-training results showed significant improvements in interval pitch discrimination, melodic perception, and singing accuracy, with average interval deviation reduced from 6.39 to 1.5 semitones. Participants reported increased confidence and enjoyment, though challenges remained in recognizing complex tactile tones and fast melodies. The study underscores the potential of multisensory integration for auditory rehabilitation and music accessibility, suggesting future improvements in tactile encoding and visual feedback design.

Other studies with comparable prototypes have pursued different goals, such as enabling deafblind individuals to comprehend speech in real time [57], exploring audio-tactile mappings [141], or activating emotional processes during audio-visual events for individuals with HL [9]. Fletcher and Zgheib [47] proposed a signal-processing strategy that converts audio signals into vibrotactile stimulation on the wrists to improve spatial hearing without bilateral implantation. Their study involved 32 participants with normal tactile perception showed that a trained group achieved a root-mean-square localization error of 22.6°, outperforming unilateral and bilateral CI users and matching bilateral hearing-aid performance. The approach relied on interaural level difference (ILD) cues delivered via multi-band dynamic-range compression and wrist sensitivity correction, with training benefits generalizing across stimuli and being most pronounced for central sound sources.

Alternative approaches to delivering vibrotactile feedback through wearable devices involve embedding actuators in furniture, enabling users to experience tactile stimulation through contact with surfaces such as chairs or sofas. Fletcher et al. [46] demonstrated that vibrotactile stimulation applied via chair armrests can enhance spatial perception, with participants integrating spatial information from auditory and tactile cues more effectively when both modalities were combined than when either was used alone. A notable example of furniture-based tactile stimulation is the Model Human Cochlea, already mentioned in Section A.10 [74]. This system supports a wide range of audio-tactile applications: for instance, the it has been employed to improve timbre perception in CI users [116] (see the end of Section 2.2.1 for overview of the study) and to deliver immersive musical experiences [75]. Another example of furniture-based vibrotactile technology is presented by Jack et al. [69], who designed and developed a suite of audio-tactile furniture for music listening targeted at Deaf users and designed in collaboration with a Deaf arts charity. Other examples include the kalimba-inspired wooden bed by Sierra et al. [121] for individuals with HL and the set of vibrotactile installations by Marozeau [89], which allow users to experience music through furniture. The museum installations featured full-body immersion platforms, chairs with frequency-specific resonance, and interactive structures such as a vibrating flower, tree, and spiral. The music was composed specifically for these installations, and audience feedback was collected during concerts at the Museum of Art and History in Geneva.

The use of vibrotactile feedback for individuals with HL does not solely imply positive outcomes and is not without challenges. Most studies conducted so far are exploratory in nature and have not been validated through clinical trials, large-scale studies, or in ecologically valid settings [102]. Consequently, they often provide interesting insights, but their results are based on a small number of subjects or on participants with TH [102]. In addition, most designs and solutions proposed in the literature, while innovative, often fall into the category of assistive devices that make users stand out and may therefore lead to social stigma [119]. The aesthetics and form factors have a significant impact on the acceptance and adoption of assistive technologies [119], yet they are frequently overlooked due to the early stage of development or a primary focus on technical performance.

In summary, vibrotactile feedback has demonstrated significant potential for enhancing auditory perception and experience in individuals with HL. A variety of prototypes, ranging from wearable devices to furniture-based systems, have been developed to provide vibrotactile stimulation that complements auditory input. These technologies exploit principles of multisensory integration and perception to improve speech intelligibility, pitch perception, spatial hearing, and overall music enjoyment. As research in this area continues to advance, further developments are expected to enhance both the effectiveness and accessibility of vibrotactile feedback for auditory rehabilitation.

2.3 Training

Since birth, humans continuously acquire knowledge, and one of the most common mechanisms for learning is training. According to the Oxford Dictionary of Education [135], training is defined as instrumental in preparing the learner for a specific task, job, or profession. It is often contrasted with education, which serves a broader developmental purpose. In the context of hearing, training has been investigated as a method to improve auditory skills and listening abilities in individuals with HC. In this section, we will explore the concept of training in the context of hearing, focusing on three main areas: auditory training 2.3.1, music training 2.3.2, and auditory-verbal therapy 2.3.3.

2.3.1 Auditory Training

> ”Auditory training can be defined as a purposeful and systematic presentation of sounds such that listeners are taught to make perceptual distinctions about those sounds” [101].

Auditory training is a structured approach to enhance auditory skills, particularly for individuals with HC with the goal to increase communication abilities. Literature shows that it can be beneficial for improving communication and cognitive skills when AT is offered along with HA fitting [125] or cochlear implantation [39].

Auditory training involves various techniques and strategies designed to improve auditory skills and listening abilities in individuals with HC. This can include activities such as sound discrimination, auditory memory exercises, and listening in noise. Recent studies on AT for individuals with HL have investigated a range of approaches and outcomes. Intensive auditory and cognitive training protocols—particularly when combined with sensory rehabilitation and HAs—can improve auditory communication skills, although long-term benefits remain inconsistent [125]. Evidence for AT in post-lingually deafened adults with CI is promising, especially for nonsense word training, which has shown improvements in sentence recognition in noise. However, further high-quality research is needed to identify the most effective training methods for specific auditory outcomes [21].

2.3.2 Music Training

Recent research highlights the effectiveness of music-based AT in enhancing auditory and cognitive functions in individuals with hearing impairments, particularly those with mild cognitive impairment [103]. Home-based musical rehabilitation for adolescents and adults has shown promising results in improving the QOL of CI users, supporting its potential integration into routine audiological care [50]. Music perception training programs have been found to enhance the recognition of tonal and emotional aspects of speech, indicating improved processing of spectrally rich auditory signals. Across various studies, musical interventions have consistently improved timbre discrimination, music appreciation, and emotional prosody perception, with socially interactive modalities yielding particularly strong cross-domain benefits [2]. Also, auditory music training has been shown to enhance musical perception, with effects influenced by factors such as age, training duration, and type of hearing device [120].

In pediatric populations, younger CI users tend to benefit more from AT than older children, likely due to greater neural plasticity [2]. Moreover, the impact of music training varies with age, intervention duration, and type of hearing device, with children exhibiting significantly greater improvements than adults, reinforcing the role of neuroplasticity in auditory rehabilitation [120]. In addition to traditional auditory rehabilitation, music training has been identified as a beneficial supplementary approach. Among different musical activities, singing has emerged as the most effective for improving rhythm, pitch, and harmony perception, followed by instrumental play and passive listening [1]. It enhances the perception of music and emotional speech prosody, contributing positively to auditory and linguistic outcomes following cochlear implantation [60]. Musical education is proposed to foster general auditory processes such as attention and perception, which in turn support broader cognitive and linguistic development [114]. Even modest engagement in music training has been shown to yield improvements in both music and speech perception, reinforcing its value as a complementary habilitation strategy for children with HL [82]. Beyond auditory and linguistic benefits, music training also appears to positively influence psychosocial well-being and QOL in children with HL [83]. Given their increased vulnerability to poorer psychosocial outcomes, participation in group-based musical activities offers a promising avenue for holistic support [83].

2.3.3 Auditory-Verbal Therapy

The formal definition of AVT describes it as a specialized educational approach aimed at supporting children with HCs in developing spoken language skills. It prioritizes auditory input and verbal communication over sign language to facilitate language acquisition, regardless of the degree of hearing impairment or the type of hearing device used [137]. Auditory-verbal therapy is recognized as a family-centered, evidence-based practice with clearly defined and measurable goals. The activities conducted during AVT sessions include parent-child interactions, auditory discrimination tasks, and language modeling techniques to enhance listening skills and spoken language development. These sessions are typically scheduled once a month and serve as training for parents (and children) to enable them to carry out the exercises at home. Intervention is delivered by speech and language pathologists who are either certified auditory-verbal practitioners or have completed the three-year training program offered by the AG Bell Academy for Listening and Spoken Language [137]. Thanks to this program, Danish children with hearing loss born after 2012 represent a population with optimal conditions for achieving high levels of functional hearing, enabling participation in the hearing world [76]. Recent studies have shown that children with HC undergoing AVT have the potential to develop foundational language skills within the typical range and exhibit strong social well-being [137]. Wischmann et al. [137] further emphasize that, with high-quality technical and educational support, these children can be fully integrated into their local hearing communities. The positive outcomes of such interventions suggest that children with HC not only improve their language abilities but also achieve levels of social well-being comparable to their peers with TH [65], displaying similar emotional and behavioral profiles as well as social strengths [138]. Nevertheless, some children may continue to face challenges related to specific aspects of their HL, particularly in tasks requiring complex auditory processing, such as sound localization or music perception. These difficulties are often observed by parents and therapists working closely with the children.

Accordingly to Estabrooks et al. [43], the principles of AVT emphasize early identification and intervention for children with hearing impairment, advocating prompt audiologic management and the use of advanced hearing technologies to maximize auditory input. Central to this approach is the active involvement of parents, who are guided to become primary facilitators of their child's listening and spoken language development through individualized therapy and consistent engagement. The therapy promotes hearing as the primary sensory modality for language acquisition, encouraging the integration of listening and spoken language into all aspects of daily life. Developmental patterns of audition, speech, language, cognition, and communication are leveraged to support natural learning, with ongoing diagnostic assessments ensuring tailored treatment plans and progress monitoring. Ultimately, the goal is to enable children to thrive in mainstream educational settings alongside their hearing peers, supported by appropriate services.

2.4 Gamification and Serious Games

Playing games is a natural activity for children and one of the most effective ways for them to learn and explore the surrounding world. It positively influences mental and social development, contributing to good psychosocial health [59] and facilitates skill acquisition [28]. For these and more reasons, the principles of gamification have been adopted in educational contexts to enhance the training experience [19, 115], promoting positive emotions and skill development, often associated with virtual or physical rewards while performing a given task [19]. The main characteristic of this approach is the incorporation of game elements to promote motivation, engagement, and stimulate individual behavior in environments, products, and services that do not primarily focus on games [19, 28].

A formal definition of gamification is provided by Deterding et al. [38]: "the use of game design elements in non-game contexts". They highlight three key aspects of this definition: the focus on design elements rather than full games, the application in non-game contexts, and the emphasis on gamefulness (structured, goal-oriented experiences) rather than playfulness (freeform, exploratory interaction). Gamification components can include interface patterns (e.g., badges, leaderboards), mechanics (e.g., time constraints, limited resources), design principles (e.g., clear goals, enduring play), models (e.g., GameFlow), and methods (e.g., playtesting, value-conscious design) [38]. The related concept of serious games is defined by Valenza et al. [130] as games designed for a primary purpose other than pure entertainment–a mental contest played with a computer according to specific rules, which uses entertainment to achieve other objectives. The distinction between gamification and serious games is therefore subtle, making it difficult to establish a clear boundary [38]. In this work, both gamification and serious games are very relevant to the context of AT for children with HL, as they can enhance engagement, motivation, and learning effectiveness.

2.4.1 Gamification and Auditory Training

In the context of this PhD project, gamification plays an important role in developing engaging and effective AT tools for children with HL. Gamification has been successfully applied in various fields, including education and healthcare [28, 38], demonstrating its potential to enhance learning outcomes and improve user engagement. A recent example is the study by Xiang et al. [140], which demonstrates the efficacy of gamified AVT for 31 preschool children with hearing impairments, particularly during intermediate and advanced stages of rehabilitation. The study introduces a custom-designed digital game that integrates multimodal interaction, real-time feedback, and narrative elements to enhance engagement and learning outcomes. Empirical results show that gamification significantly improves auditory memory, articulation, and expressive language skills, especially in children at the intermediate stage, while also providing a more enjoyable and motivating experience compared to traditional therapy. These findings underscore the potential of stage-specific gamified interventions to complement professional AVT and support personalized rehabilitation strategies.

Other examples of gamified AT applications for children with HL is the work from Cano et al. [22], where they introduce a serious game that integrates physical and digital interfaces to foster engagement and learning in children with HL. The game combines a physical board with a mobile application, utilizing QR codes and vibrotactile feedback to enhance interaction. The qualitative study involving seven children aged 7 to 11 from the Institute for Deaf and Blind Children in Colombia demonstrated increased motivation, collaboration, and gradual improvement in computational thinking skills such as problem-solving, reasoning, planning, and communication. The findings suggest that serious games, when designed with inclusive methodologies and multimodal interaction, can effectively support cognitive development in children with specific needs [22].

Mobile games have been also recently investigated, aiming for the integration of gamification into AT for CI users, aiming to enhance engagement and perceptual learning. The study by Zhou et al. [147] introduces MOGAT (Mobile Games with Auditory training), a mobile-based system designed to improve musical auditory habilitation for pre-lingually deafened children with cochlear implants. Unlike traditional AT focused on speech, MOGAT targets pitch perception and production, which are critical for music perception but challenging for CI users. The system comprises three interactive games: Higher Lower for pitch interval discrimination, Vocal Matcher for single-pitch production and breath control, and Ladder Singer, which integrates pitch, lyrics, and visual cues to guide melody singing. A cloud-based web service complements the games, enabling teachers to monitor progress, provide feedback, and personalize training remotely. Development was informed by a pilot study comparing CI and TH children, which revealed significant deficits in pitch perception and intonation among CI users. A two-week user study with 15 CI children demonstrated statistically significant improvements in pitch discrimination, pitch matching speed, and melody singing accuracy, alongside high ratings for enjoyment, motivation, and usability ($> 4.3/5$). Teachers also reported positive experiences with the web platform for managing habilitation. Overall, MOGAT represents the first integrated mobile solution for music-focused AT in CI children, offering an engaging, cost-effective, and scalable approach to complement conventional therapy.

Garadat [55] developed and beta-tested SoundScape1 and SoundScape2, two smartphone-based serious games designed for adult CI recipients. These games, based on the infinite runner genre, target early auditory skills—sound detection and discrimination—using structured, repetitive exercises. The development followed a participatory action research framework across three phases, involving iterative design improvements and user feedback. Final evaluations with 51 participants revealed high satisfaction ($> 90\%$) with features such as graphics, feedback, and training mode, indicating strong usability and motivational potential for self-directed rehabilitation. In contrast, Çetinkaya et al. [91] focused on preschool children (ages 3-5) with prelingual HL, creating World of Sounds, a mobile application comprising four games–Crucible Sound, Mole Hunting, Find the Sound, and Choo-Choo–targeting auditory awareness, discrimination, recognition, and comprehension. Usability testing with 40 children (20 CI users, 20 TH) and their parents demonstrated high ratings for ease of use, motivation, and design, while performance analysis showed CI users achieving success primarily at easier levels. Positive correlations between game scores, duration of CI use, and developmental measures underscore the app's developmental appropriateness. Collectively, these studies highlight gamified AT as a promising adjunct to traditional rehabilitation, offering accessible, engaging, and customizable tools to improve auditory skills in both adults and children with cochlear implants.

2.4.2 Games and Haptics

When playing on gaming consoles such as PlayStation or Xbox, it is hard not to notice that these devices provide vibrotactile feedback through their controllers, specifically via actuators that convey synchronized vibrations to the players' hands. This feature has been shown to effectively enhance immersion and engagement [51]. It represents one of the most evident examples of haptic feedback in gaming and has been widely adopted by console manufacturers over the decades. More advanced approaches have also been explored, such as systems with six degrees of freedom that incorporate force and torque to deliver more realistic haptic sensations to players [68]. A great body of research focuses on the integration of haptic feedback in virtual reality (VR) gaming to enhance user experience and immersion, since the VR environment primarily relies on visual and auditory stimuli [18].

Especially in the context of VR gaming, haptic feedback plays a crucial role in enhancing the sense of presence and immersion within virtual environments [70]. It has also demonstrated its effectiveness in delivering more congruent experiences through multisensory integration and supporting rehabilitation for children with special needs [18]. Various haptic devices have been developed for VR applications, including gloves, thimbles, belts, and vests that provide tactile sensations corresponding to virtual interactions [3, 18, 20, 70]. These devices can simulate a wide range of sensations, from simple vibrations to complex textures and forces, thereby enriching the gaming experience.

As introduced in Section 2.2.2, mappings between audio and haptic stimuli can be employed to convey auditory information through tactile sensations. This approach has been explored in gaming to enhance user experience and accessibility. A notable example of mappings applied to haptic feedback in games, is the work from Yun et al. [143], in which they present a real-time system for generating selective and multimodal haptic effects from gameplay sound to enhance user experience in video games. The proposed approach employs a machine learning classifier based to determine whether a sound event should trigger haptic feedback and to select the appropriate modality among vibration, impact, or a combination of both. A user study comparing the proposed method with state-of-the-art algorithms demonstrates significant improvements in adequacy, synchrony, and overall preference. Furthermore, the integration of multimodal feedback enriches immersion, although user preferences vary depending on context. This work establishes the first machine learning-based, explicitly selective, and real-time sound-to-haptic conversion framework, paving the way for more immersive and scalable haptic experiences in interactive media.

CHAPTER 3

Research Questions

3.1 Problem Area 1 – State-of-the-Art and Clinical Needs

In order to address the overarching aim of this PhD project, an investigation of the state of the art regarding the use of haptics for individuals with HC was conducted, particularly in training and gamified contexts. This investigation was guided by the following research questions:

Problem Area 1 - Research Questions

Q1_1 What is the current state of the art regarding the use of haptics for individuals with hearing challenges, particularly in training?

Q1_2 What are the needs and interests of the clinical professionals working with children with HL regarding the use of tools and prototypes for the auditory training?

Problem Area 1 lays the foundation for the following two Problem Areas, as they propose practical solutions based on the identified gaps and needs.

3.2 Problem Area 2 – Vibrotactile Applications

The second problem area focuses on the design and development of vibrotactile applications for children with HL, aiming to enhance their auditory experience of music. The research questions guiding this area are as follows:

Problem Area 2 - Research Questions

Q2_1 How can vibrotactile feedback be effectively designed and implemented to enhance the music experience for children with HL?

Q2_2 What are the perceived benefits and challenges of using vibrotactile applications among children with HL and their caregivers?

3.3 Problem Area 3 – Training Tools

The last problem area aims to create specialized training tools that can support the auditory development of children with HL. The research questions relevant for this area are the following:

Problem Area 3 - Research Questions

Q3_1 How can these training tools be tested and evaluated with children with HL in real-world settings?

Q3_2 What are the outcomes of using these training tools in terms of auditory skills, engagement, and overall user experience among children with HL?

CHAPTER 4

Summary of Papers

The work conducted during this PhD project has been documented in six research papers, each addressing different aspects of the research questions outlined in Chapter 3. This chapter provides an overview of each included paper, summarizing their motivation, methodology, key findings, conclusions, and a small reflection regarding the project. Furthermore, two other additional papers, which are not part of the PhD thesis, are included. These studies are related to the research topic and were conducted by the author in collaboration with other researchers during the same period. The full papers are presented in Part II of this dissertation, while this chapter serves as a concise guide to their content and contributions to the overall research objectives.

The order of the papers is not following a chronological sequence but is instead organized thematically based on the tools used and the specific research questions they address. The first paper is a systematic review that lays the foundation for the subsequent empirical studies. The following three papers focus on the design, development, and evaluation of vibrotactile applications and training tools for children with HL, exploring various aspects such as user experience, effectiveness, and practical implementation in real-world settings. The last two papers propose two novel training tools that aim to be used in AT, one addressing the training of localization abilities, and the second the training of vocal emotions recognition.

4.1 Paper A

The Role of Haptics in Training and Games for Hearing-Impaired Individuals: A Systematic Review

Motivation

This systematic review addresses the emerging field of haptic feedback for individuals with HC, particularly in training and gamified contexts. This specific topic has been chosen to understand the current state of the art, identify gaps in the literature, and inform the development of the different vibrotactile applications and training tools created during this PhD project.

Methodology

The review followed PRISMA 2020 guidelines. Searches were conducted in Scopus® and PubMed® using keywords grouped into three categories: tactile interaction (e.g., haptics, vibrotactile), hearing impairment (e.g., deaf, hearing-impaired), and training/gamification (e.g., game, education). Inclusion criteria required English-language, peer-reviewed primary research published within the last 25 years, addressing hearing-impaired populations. After screening 294 unique records, 35 manuscripts were included. Each study was categorized by research focus (design or effectiveness), study type (experimental, quasi-experimental, pre-post test, usability evaluation, mixed methods), haptic usage (sensory substitution or augmentation), body location, mapping strategies, vibrotactile processing, target impairment, training duration, and data collection methods.

Findings

The analysis revealed a growing interest in haptics for training in individuals with HL, particularly in the last five years. Most studies (77.14%) focused on effectiveness, commonly measuring task accuracy. Haptic feedback was used for sensory substitution in 54.29% of cases and for sensory augmentation in 45.71%. Hands and fingertips were the most frequent stimulation sites, and sound-to-vibrotactile mapping was employed in over 50% of studies. Vibrotactile processing often involved temporal envelope extraction or synthetic generation. Smartphones were the most common devices, typically using eccentric rotating mass actuators, while high-fidelity laboratory setups employed electrodynamic shakers. Participant recruitment was challenging, with an average of 8.31 hearing-impaired individuals per study. Outcomes were predominantly positive: 74.28% of studies reported beneficial effects, and 54.29% achieved statistical significance. All six studies with significant positive results used sound-to-vibrotactile mapping, confirming that perceptual aspects of sound can be transmitted through touch.

Conclusions

The review highlights the potential of haptic feedback as a viable solution for rehabilitation and training of individuals with HL, particularly when using sound-derived vibrotactile stimuli applied to sensitive body areas such as hands or wrists. However, the literature on gamification combined with haptics remains sparse, with only three studies addressing this intersection. Device diversity and lack of standardization in vibrotactile processing complicate generalization of findings. Recruitment challenges and small sample sizes limit statistical power, underscoring the need for collaborations with hospitals and care centers. Future research with haptic technologies should prioritize ecologically valid environments, and combine qualitative and quantitative assessments to better capture user experience and training effectiveness.

Reflection

This review highlighted the challenges of recruiting participants with HL, which I later encountered in the studies presented in the subsequent papers. I also learned that utilizing ecologically valid settings and involving clinicians in the design process are crucial considerations when developing vibrotactile training tools for this specific target group. Recognizing that few studies have explored the combination of gamification and haptics further motivated me to incorporate game elements into most of the prototypes I developed during this PhD and made me aware of the gap in the literature to which my research could contribute.

4.2 Paper B

Tickle Tuner - Haptic Smartphone Cover for Cochlear Implant Users' Musical Training

Motivation

This paper contains the partial outcome of the author's Master Thesis, and wants to address the challenge of designing a portable device that provides vibrotractile feedback for musical training of CI users. Specifically, this project has been driven by the interest in exploring how vibrotactile feedback can enhance pitch perception, which is often limited in CI users due to the device's technical constraints. During the early stages of this PhD project, a second design iteration brought a more refined version of the prototype, with an improved ergonomic design.

Methodology

The authors developed the *Tickle Tuner*, a haptic smartphone cover designed for musical training of CI users. The prototype consists of two high-quality haptic actuators (HapCoil-One) connected to a stereo class-D amplifier and powered via a modified digital-to-analog converter (DAC) through a USB-C plug. A 3D-printed ergonomic shell and handles designed for comfortable grip and optimal vibration transmission. The device's frequency response was measured using an accelerometer and MATLAB-generated sine sweeps in an anechoic room. Filters were designed to compensate for internal resonances and fingertip sensitivity. A test bench was implemented in Pure Data with a GUI via MobMuPlat to create and evaluate audio-haptic mappings. The study tested three mappings during a melodic contour identification task: amplitude modulation (AM), full audio through haptics, and no haptics. Fifteen normal-hearing participants using a CI simulation completed the task in an anechoic room.

Findings

Results showed that AM mapping significantly improved performance in melodic contour identification compared to other conditions. Participants achieved with the AM mapping the 65% of correct answers, with full audio

(CI simulation) the 45%, while without haptics only the 38%. Statistical analysis (ANOVA and t-tests) confirmed significant differences ($p\text{-value} < 0.05$) between all mappings. The improvement was particularly evident for complex timbres (e.g., viola), while sine waves were easier to identify across all mappings. The study found no correlation between musical experience and task performance under CI simulation.

Conclusions

The prototype demonstrated the potential of vibrotactile feedback, specifically with AM mapping, to convey pitch information and enhance musical training for CI users without prior training. However, limitations include the small sample size and reliance on hardware-based CI simulation, which cannot replicate individual variability among CI users. Future work will involve hardware and software improvements and further testing with actual CI users to validate effectiveness in real-world scenarios.

Reflection

This prototype served as an entry point for my journey into the field of vibrotactile feedback, and I learned a great deal about the design and technical challenges of creating portable haptic devices. The positive results obtained with AM mapping motivated me to further explore vibrotactile applications for children with HL, but this also meant abandoning the prototype in favor of more child-friendly designs. If time had allowed, I would have liked to further test this prototype with the actual target group during a second design iteration, as I received positive feedback from a few CI users during demo sessions at conferences and events.

4.3 Paper C

Vibrotactile Memory: A Case Study of Timbre Perception Training in Children with Cochlear Implants Using a Video Game

Motivation

CI have significantly improved speech perception in individuals with profound HL, but music appreciation remains challenging due to hardware and physiological limitations that affect timbre recognition, dynamic range, and frequency resolution. These limitations are particularly relevant for children with ANSD, who experience disrupted neural activity that impacts timing and low-frequency perception. Clinicians at CHBC have indicated that this specific target group and challenge could potentially benefit from vibrotactile feedback to enhance music perception. This study aims to explore the feasibility and effectiveness of a vibrotactile-enhanced video game for timbre perception training in a child with CI and AN.

Methodology

A case study was conducted with a 7-year-old child using a bimodal hearing configuration (CI in the left ear and HA in the right ear) diagnosed with ANSD. The intervention involved a custom-designed video game inspired by the memory card game, where players match musical instruments based on timbre. The game was developed in Unity and integrated vibrotactile feedback via the PlayStation 5 DualSense controller using FMOD for synchronized audio-haptic streams. The training lasted two weeks, with three home sessions per week supervised by parents. A timbre perception test was administered pre- and post-training. Data collected included clicks per card, completion time, and level progression. Interviews with the child and parents were conducted post-training to gather qualitative feedback.

Findings

Quantitative results showed no significant improvement in timbre perception after training (p-value = 0.2466). Interaction metrics indicated that the number of clicks per card averaged 2.42, approximately twice the optimal moves, and completion times increased with level difficulty. No clear effect of vibrotactile feedback was observed on performance metrics. Qualitative feedback revealed that the child experienced boredom and attempted to bypass the game mechanics, while parents noted some awareness of musical instruments but suggested improvements in engagement and reward systems.

Conclusions

The absence of measurable improvement suggests that the short training duration, repetitive gameplay, and limited gamification elements reduced engagement and learning efficacy. Vibrotactile feedback did not appear to influence outcomes, possibly due to cognitive load or insufficient understanding of its role. The study highlights the need for extended training periods and more engaging game design. Future work should include larger participant samples, control groups, and diversified mini-games to enhance motivation and ecological validity. Despite limitations, the case study provides valuable insights into designing AT interventions for children with CI and AN.

Reflection

This study introduced me to video game development for the first time, which was both exciting and challenging. I realized the importance of incorporating gamification elements to keep children engaged, particularly when addressing repetitive training tasks. Although the quantitative results were not as expected, the qualitative feedback provided valuable insights that informed the design of subsequent prototypes. This experience underscored the complexity of creating effective training tools for children with HL, given the fierce competition with mainstream games and their extremely high standards. In retrospect, I would have liked to involve an experienced game designer in the project to enhance the game's appeal and offer a more varied set of activities.

4.4 Paper D

Vibrotactile Teddy Bear: Enhancing Musical Experiences For Children With Cochlear Nerve Deficiency Through A Vibrotactile Soft Toy

Motivation

Children with hearing challenges, particularly those with ANSD and CND, face significant barriers to accessing and enjoying music, which impacts their overall development and QOL. While HAs and CIs improve speech comprehension, they do not provide the full spectrum of musical experiences due to limitations in frequency range, dynamic range, and timbre recognition. As in Paper C (4.3), the target group and specific need were highlighted by clinicians at CHBC. Integrating vibrotactile feedback into AVT sessions for very young children with ANSD could offer a more holistic and engaging experience without relying on visual cues, which tend to dominate in CI users.

Methodology

A soft toy prototype was developed to integrate sound and vibrotactile feedback synchronized to music. The toy, shaped like a stuffed animal, included a speaker and a vibrotactile motor controlled via Bluetooth. Two design iterations were implemented: the first used a stereo Bluetooth receiver and amplifier powered by AA batteries, while the second integrated an ESP-32 AudioKit with a rechargeable LiPo battery and programmable features. We studied two children with CND: Little Star (aged 2 years 10 months) and Bright Star (aged 4 years 5 months). Both were bilateral CI recipients enrolled in AVT. Sessions were conducted at CHBC during regular therapy, with the therapist guiding interactions. Activities included listening to music and playing a sound identification game. Observations were recorded, and parents provided feedback after extended home use of the toy.

Findings

Both children showed positive reactions to the prototype, but engagement varied by age and developmental stage. The younger child, Little Star, displayed initial curiosity and enjoyment of vibrations but required frequent redirection and lost interest over time. In contrast, the older child, Bright Star, remained focused, using the toy effectively during listening tasks and games. Parents reported that Little Star played extensively with the toy at home, though it was eventually damaged due to heavy use. Overall, vibrotactile feedback appeared to enhance engagement and auditory experience, particularly for older children who could understand the connection between sound and vibration.

Conclusions

The study highlights the potential of integrating vibrotactile feedback into AVT for children with severe hearing challenges, offering a multisensory approach to music perception. Age and developmental stage significantly influence engagement, suggesting the need for personalized and adaptable designs. The therapist's role is critical in guiding effective use and integrating the toy into therapy and home routines. Limitations include the small sample size, short study duration, and heterogeneity of participants. Future work should explore customization options, advanced vibrotactile mappings, and clinical integration, as well as conduct longitudinal studies with larger cohorts. While vibrotactile feedback may not be necessary for most d/hoh children, it could provide meaningful benefits for those with profound auditory deficits, such as CND.

Reflection

This project adopted a slightly different approach compared to the other studies, as the participants were very young children and the intervention aimed to expose them to music using the vibrotactile aid rather than training a specific skill. One of the most valuable experiences was observing how differently the two children responded to the same prototype, which reinforced the importance of considering age and developmental stage when designing such tools. The positive feedback from parents and therapists was encouraging and highlighted the potential for vibrotactile toys to become a staple in AVT for children with specific conditions incentive their sound exposure.

4.5 Paper E

SoundCubes - Preliminary Evaluation of a Spatial Hearing Training Tool and a Sound Localization Test for Children with Hearing Loss

Motivation

After being introduced by one of the audiologists at CHBC to the teddy bear used for training localization abilities, we decided to design a tool to improve this component of the AVT program. Children with HL often experience degraded spatial hearing, which impacts communication, safety, and overall QOL. Sound localization depends on cues such as ITD, ILD, and spectral filtering, which are challenging to replicate with hearing technologies like HA and CI. Training-based interventions leveraging neuroplasticity have shown promise in enhancing spatial hearing; however, there is a lack of child-specific, engaging tools for localization training. This study introduces SoundCubes, a tangible and portable system designed to support AT within frameworks such as AVT, with the goal of improving spatial hearing in children with HL.

Methodology

A mixed-methods design was employed. Quantitatively, a pre-post evaluation of sound localization abilities was conducted using a custom localization test in two configurations (120° span with 15° spacing; 360° span with 45° spacing). Six children with TH completed the test twice to assess reliability, while five children with HL underwent a two-week home-based SoundCubes training program. Qualitatively, usability and engagement were assessed through the System Usability Scale (SUS) questionnaire completed by parents and clinicians, and smiley-scale ratings from children. The SoundCubes system consists of up to nine wireless cubes controlled via an app, playing sound stimuli that children identify by picking up the active cube. Data logging, adaptive features, and feedback mechanisms were integrated. Statistical analyses included Pearson correlation, root mean square error (RMSE), and interclass correlation coefficient (ICC) for reliability, and paired tests for performance changes.

Findings

The localization test demonstrated high test-retest reliability in the TH group: for the 120° configuration, Pearson $r \approx 0.96$, RMSE $\approx 12°$, and ICC(A,1)=0.709; for the 360° configuration, $r \approx 0.91$-0.93, RMSE ≈ 42–47°, and ICC(A,1)=0.616. Training outcomes for children with HL were heterogeneous: two participants showed significant improvements in correlation for the 120° test, while one exhibited a significant deterioration in RMSE for the 360° test. Group-level changes were not statistically significant, likely due to small sample size and variable adherence. Usability ratings were high: parents (SUS $M = 86.25$), clinicians ($M = 75.63$), and children reported strong enjoyment (mean rating for “Did you like the SoundCubes?” = 4.75/5). Engagement was positive, though sustained interest was moderate.

Conclusions

The study confirms the reliability of the localization test and the feasibility of SoundCubes as an engaging, portable training tool for home and clinical use. Preliminary evidence suggests potential benefits for spatial hearing, but variability in outcomes highlights the need for larger samples, longer training durations, and standardized protocols. Challenges included limited recruitment and inconsistent adherence, underscoring the importance of structured support for families. Positive usability feedback from parents, clinicians, and children indicates strong potential for integration into AVT and other AT programs. Future work should focus on enhancing long-term engagement, improving ecological validity, and embedding SoundCubes into routine rehabilitation workflows to maximize impact on pediatric spatial hearing development.

Reflection

SoundCubes is undoubtedly the most comprehensive and extended project I worked on during this PhD. It required considerable effort in development, prototype construction, recruitment, and testing. As one of the final two studies, I had the chance to apply lessons learned from previous work, particularly regarding the importance of involving clinicians, parents, and children in the design process. The evaluation of the test has been also a valuable experience, and allowed me to assist to the tests, observing behaviors, difficulties, and strengths of the children while performing the task. The positive feedback received from all stakeholders was highly rewarding and reinforced my belief in the potential of training for children with HL. If I had more time and resources, I would have liked to conduct a longer-term study with a larger cohort to better assess the training effects and refine the system based on user feedback. I believe SoundCubes has strong potential for integration into clinical practice, and I hope future research will allow me to continue developing and evaluating it.

4.6 Paper F

EmotiCubes - A Training Tool for Vocal Emotion Recognition in Children with Hearing Loss

Motivation

This project originated from a collaboration with the db SPL group in Groningen, led by Dr. Laura Rachman and Professor Deniz Başkent. Upon introduction of the SoundCubes concept, the researchers identified the potential for a similar tangible approach to be applied to vocal emotion recognition training, an area of ongoing interest within their research group.

Children with HL often experience difficulties in recognizing emotions from speech prosody, which negatively impacts social communication and QOL. While evidence suggests that training can improve prosodic perception, few interventions specifically target vocal emotion recognition in this population. Existing research highlights variability in performance among children with HL, with some achieving age-typical scores and others lagging behind. The lack of dedicated training programs for vocal emotion recognition underscores the need for innovative solutions. Tangible user interface (TUI) have been identified as promising tools for engaging young learners, leveraging physical interaction to support auditory training. To address this gap, the EmotiCubes system was developed as a playful, home-based training tool aimed at improving vocal emotion recognition in children with HL.

Methodology

A pilot study was conducted using EmotiCubes, a TUI designed to facilitate vocal emotion recognition training through interactive auditory tasks. The prototype consists of a cube displaying three emotions—angry, happy, and sad—on its faces. Children listen to pseudospeech stimuli and rotate the cube to select the perceived emotion, receiving auditory feedback and rewards to enhance engagement. The hardware includes an ESP32 A1S AudioKit microcontroller, IMU MPU9250, loudspeaker, and microSD storage, while the software manages stimulus presentation and logs training data. Two Danish participants (ages 4-12) with bilateral CI and bimodal hearing completed a pre-training EmoHI test, engaged in three weeks of home-based EmotiCubes training (three sessions per week, 12 trials per session), and then completed a post-training EmoHI test. Usability was assessed via the SUS questionnaire completed by one participant's parents. Performance was measured using the sensitivity index (d'), calculated from hit and false alarm rates for each emotion.

Findings

Both participants demonstrated improvements in vocal emotion recognition after training, with average gains of $\Delta d' \approx 0.37$. Participant P1 improved from $d' = -0.2958$ to 0.0718, showing modest gains primarily for Happy ($\Delta = 1.2918$), while Angry and Sad showed smaller or negative changes.

Participant P2 improved from $d' = 0.8576$ to 1.2252, with notable gains for Angry ($\Delta = 0.4841$) and Happy ($\Delta = 0.6186$), while Sad remained unchanged. Accuracy increased from 25.00% to 36.11% for P1 and from 61.11% to 72.22% for P2. Usability feedback indicated good acceptability (SUS score: 72.5), with parents appreciating reward mechanisms and suggesting slower pacing between trials. Despite variability in training performance trends, both participants improved in the EmoHI test, suggesting benefits from EmotiCubes training.

Conclusions

This pilot study provides preliminary evidence that EmotiCubes can support vocal emotion recognition training in children with HL. Improvements were observed for both participants, with one achieving performance comparable to age-typical peers. These findings highlight the potential of gamified, tangible interfaces for auditory training and early intervention programs. However, discrepancies between training and test performance, small sample size, and lack of a control group limit generalizability. Future work should focus on larger-scale studies, adaptive difficulty, multimodal feedback, and remote monitoring to enhance usability and effectiveness. Overall, EmotiCubes represents a promising approach for improving socio-emotional development in children with hearing challenges.

Reflection

I learned a lot from this collaboration, especially about the challenges faced by children with HL in social communication. During this project I faced major difficulties in recruiting participants, which limited the validation of the tool. However, the positive results from this pilot study are encouraging, and I believe that EmotiCubes have great potential for further development and research. If given more time, I would have liked to conduct a larger-scale study with more participants and a control group to better assess the effectiveness of the training tool. I also know that the db SPL team is interested in expanding their project with more participants coming from different countries, and I would really like to understand if EmotiCubes can be effective across different languages and cultures.

4.7 Additional Paper G

Enhanced Neural Phase Locking Through Audio-Tactile Stimulation

Motivation

This study originated from a collaboration with a talented visiting master student at the ME Lab, whose background in engineering and neuroscience enabled the exploration of neurophysiological mechanisms underlying audio-tactile interactions. The project extended previous research by investigating neural correlates of audio-tactile integration, an area not previously addressed in the context of this thesis.

Numerous psychophysical and neurophysiological parallels exist between auditory and vibrotactile modalities, particularly in pitch perception. Both systems exhibit phase locking behavior in their peripheral structures, enabling temporal coding of stimulus frequency. Prior research has demonstrated perceptual audio-tactile interactions, such as reciprocal biasing of pitch perception and shared cortical activation patterns. However, the neural mechanisms underlying these effects remain unclear, especially at subcortical levels. Considering the efficacy of the frequency following response (FFR) in capturing subcortical phase locking activity, this study aims to investigate whether concurrent vibrotactile stimulation enhances auditory phase locking in humans. The hypothesis posits that vibrotactile stimuli complement auditory stimuli and improve phase locking acuity at the fundamental frequency (F_0), which is associated with pitch perception.

Methodology

FFRs were recorded from 22 healthy young adults (28 ± 6 years; 11 female) with no reported hearing loss. The stimulus consisted of a 40-ms /da/ speech syllable with dynamic formant transitions (F_0: 103-125 Hz; F1: 220-720 Hz; F2: 1,700-1,240 Hz; F3: 2,580-2,599 Hz). Auditory stimuli were presented monoaurally through insert earphones at 80 dB SPL, while vibrotactile stimuli were delivered bimanually using a DualSense controller calibrated to 0.85 m/s^2 peak-to-peak. Each session comprised 40 blocks of 100 trials at 10.9 Hz, with randomized inclusion of vibrotactile stimulation in 50% of blocks. FFRs were recorded using a vertical electrode montage (Cz-A2-Fpz) and processed with bandpass filtering (100-2,000 Hz), artifact rejection ($\pm 23.8 \mu V$), and polarity addition. Spectral encoding was analyzed via fast Fourier transform (FFT) for F_0 (75-175 Hz), lower harmonics (175-750 Hz), and higher harmonics (750-1,050 Hz). Statistical significance was assessed using paired permutation tests with Bonferroni correction.

Findings

Preliminary validation confirmed the suitability of the DualSense controller, showing consistent frequency response and sufficient return to baseline between trials (SNR: x = 15.49 dB, y = 14.74 dB, z = 13.39 dB; $p < 0.001$). Time-domain analysis revealed no significant differences in peak timing between audio and audio-tactile conditions, indicating that vibrotactile stimulation did not confound temporal alignment. Frequency-domain analysis demonstrated a significant increase in spectral amplitude at F_0 under audio-tactile stimulation (mean difference = 3.30 ± 8.41 µV; $p = 0.033$; Cohen's $d = 0.39$), while no differences were observed for harmonic frequencies. Control experiments with vibrotactile-only stimulation produced responses indistinguishable from baseline, suggesting that the observed effect reflects a super-additive interaction rather than linear summation.

Conclusions

The findings support the hypothesis that vibrotactile stimulation enhances auditory phase locking at subcortical levels, specifically at F_0. This effect may be mediated by multisensory integration within the inferior colliculus, consistent with its role as a hub for crossmodal processing. However, the absence of effects at harmonic frequencies could stem from biological constraints (limited vibrotactile bandwidth) and methodological factors (polarity addition favoring temporal envelope encoding). Future research should explore alternative polarity processing, weaker unisensory stimuli to test inverse effectiveness, and correlations with behavioral measures of pitch perception. Expanding participant demographics, including musicians and individuals with hearing impairments (e.g., CI users), may reveal differential benefits. Ultimately, these results suggest promising applications for audio-tactile training in improving pitch intelligibility, with potential implications for clinical and educational contexts.

Reflection

The concept of FFR seemed almost too good to be true, as it enables the assessment of subcortical activity through a non-invasive method, where the subject does not have to perform any specific active task. The resulting recordings can be correlated with the input stimulus to objectively evaluate how well the brain encodes specific features of sound. In this study, we were able to start exploring the neural mechanisms behind the audio-tactile interactions that I had been investigating from a more applied perspective in the previous studies. The results were very encouraging, and I believe that this line of research has a lot of potential for future studies, especially in clinical populations that could benefit from audio-tactile stimulation.

4.8 Additional Paper H

Multisensory Integration Design in Music for Cochlear Implant Users

Motivation

This project was part of Dr. Razvan Paisa's PhD project. Cochlear implant (CI) users experience significant challenges when listening to music, including difficulties in recognizing timbre and pitch, sound localization, and segregating individual instruments in multi-instrument mixes. While CI technology is advanced for speech perception, music appreciation remains underwhelming. The diversity among CI users—due to factors such as age, residual hearing, hearing aid use, and musical training—requires individualized approaches. This research aims to enrich CI users' musical experiences by integrating multisensory feedback to support the perception of specific musical features and elements.

Methodology

The study employed a participatory design approach through exploratory workshops with three CI users. Each workshop lasted 60-120 minutes and included semi-structured interviews, think-aloud protocols, and observations. Participants explored three installations:

- *Installation 1:* A multi-channel speaker setup enabling instrument segregation by spatializing individual instruments.
- *Installation 2:* Audio-tactile seating configurations using actuators on seats, footrests, and handheld devices to enhance low-frequency perception.
- *Installation 3:* Embodied interaction through movement-based performance excerpts and in-air haptics inspired by the Felt Sound project.

Feedback was collected via pre- and post-experiment interviews and continuous verbal commentary during sessions.

Findings

Participants demonstrated varied preferences and experiences:

- *Installation 1:* All participants improved instrument identification when allowed to move closer to individual speakers. Spatial separation facilitated auditory streaming.
- *Installation 2:* Tactile feedback was generally appreciated, especially through footrests and handheld devices. Overly strong vibrations or unsuitable music styles reduced enjoyment. Participants emphasized the importance of music preference in enhancing experience.
- *Installation 3:* Gestural and movement-based performance was perceived as supportive for understanding and enjoying music. Participants expressed interest in integrating gestures into their own musical practices.

Overall, tactile and spatial cues enriched music perception, but interpersonal differences in musical taste and hearing profiles strongly influenced outcomes.

Conclusions

The study highlights the need for process-oriented and personalized design for CI users. Flexibility and customization in multisensory devices are essential to accommodate diverse hearing abilities and musical preferences. Designers should ensure independent control of tactile and auditory stimuli and consider modular hardware configurations. Integrating visual feedback, such as gestures or movement-based performance, can further bridge perceptual gaps. Building communities and fostering collaborations between CI users, musicians, and researchers remain critical for advancing inclusive music technologies. Future work should focus on creating rehabilitation and practice frameworks that combine enjoyment with skill development.

Reflection

I took part in this project during the very beginning of my research journey, and it was a great opportunity to learn about participatory design methods and the importance of involving end-users in the design process. One of the things that might seem obvious but that I learned through this experience is how diverse CI users are, and how important it is to consider their individual preferences and needs when designing multisensory music experiences. The strong effect of music preference on the enjoyment of the installations was quite surprising, and it made me realize that even the best-designed technology might not be effective if it does not align with the user's tastes.

CHAPTER 5

Conclusion

The scope of this PhD project has been to explore and develop interactive music training tools for children with HC. Over more than three years of research and development, five different prototypes were created, each addressing specific challenges faced by individuals with HC in their auditory and musical experiences. Throughout this process, I provided answers and new questions related to the three problem areas outlined in Chapter 3. I have encountered several obstacles, ranging from technical issues to recruitment difficulties. These challenges yielded invaluable learning experiences, contributing to my growth as both a person and a researcher, while also shaping the direction of the whole project. Numerous rewarding moments made this journey worthwhile, such as witnessing children's enthusiasm while interacting with the prototypes and receiving positive feedback from clinicians regarding the potential impact of these tools. A single smile from a child successfully using one of the prototypes made all the effort meaningful.

Nevertheless, this project has only laid the groundwork for future research and development in this area. In the following sections, I will outline the problem areas addressed, propose future directions, and highlight emerging questions for further exploration.

5.1 Problem Area 1

Research Question Q1_1

The systematic review presented in Paper A laid the foundation for the entire PhD project by identifying the current state of the art in haptic and multimodal training and gaming for individuals with HL. The review highlighted the potential benefits of incorporating haptic feedback into training tools, as well as the existing gaps in research and development in this area. One of the key insights from the systematic review was the customary use of sound-to-haptic mappings without complex signal processing (e.g. audio stream separation, or F_0 extraction, see 2.2.2 for more), along with the focus on highly sensitive areas for vibrations, such as the hands. These findings were instrumental in designing the various prototypes presented in this PhD project and helped providing an answer to the research question Q1_1. Additionally, the review emphasized the lack of research on gamified training tools that provide haptic feedback for individuals with HL, which motivated the development of engaging and interactive prototypes capable of effectively leveraging vibrotactile feedback to enhance the training experience. A warning from the review concerned the challenge of recruiting participants with HL, as most of the included studies had a limited number of participants, making it difficult to draw definitive conclusions about the effectiveness of the proposed interventions. This challenge was encountered multiple times throughout the PhD project, underscoring the need for more extensive recruitment strategies and collaborations with clinical partners to ensure sufficient participant numbers in future studies. This aspect must also be put into perspective, as the studies in this PhD project present novel prototypes that are not yet suitable for large-scale clinical trials.

Research Question Q1_2

The research question Q1_2, in contrast, required a longer time to be addressed, as it involved extended interaction with clinical professionals and the children assisted at the CHBC. The time available for each patient during a clinical session is limited, requiring clinicians to carefully select the tools and interventions they employ. Consequently, any new training tool must be efficient, easy to set up, and require minimal calibration to be adopted in clinical practice. Clinicians expressed a preference for solutions that can be used with multiple children with small adaptation, emphasizing the need for scalability and robustness. Real-time feedback and quick responses were also identified as important features to maintain high engagement during training sessions. The AVT approach teaches and encourages parents to conduct most of the training at home; therefore, training tools should ideally support home-based use. This introduces additional requirements, such as user-friendliness and the ability to track progress remotely. The latter aspect is a great feature for clinicians to monitor training effectiveness and make necessary adjustments in the intervention. Logging and progress-tracking functionalities are thus essential components for any training tool intended for home use.

When collaborating with larger teams not part of the same institution, other considerations should be made. Involving clinicians from the beginning of the project requires time from their side, and resources from the project budget to allocate for their participation. This sometimes is overlooked, but it is crucial for the fair and successful involvement of clinical professionals in research projects. Also, the clinicians' perspective is different from that of developers, as their primary focus is on patient care. Therefore, it is essential to align the goals of the research project with the needs and priorities of the clinical practice to ensure successful collaboration and adoption of the developed tools. In other words, using the clinicians as mere testers of the prototypes is not sufficient; they need to be actively involved in the design and development process to ensure that the final products meet the practical requirements of clinical settings with the ultimate goal of benefitting the patients.

5.2 Problem Area 2

Research Question Q2_1

The insights gathered from interactions with the clinicians provided an understanding of the major advancements achieved in recent decades in terms of speech acquisition and language development for children with HL. Thanks to the AVT program, recent generations of children with CI have been able to develop spoken language skills comparable to those of their TH peers. Consequently, the need for multimodal training tools (e.g. vibrotactile feedback prototypes) to support speech acquisition is no longer as critical as it was in the past. One of the core principles of the AVT program is to prioritize the hearing sense, thus the “Sound first” rule. However, for specific conditions such as ANSD or CND, or for more complex auditory skills (e.g. sound localization), children with HL still face challenges. With the mentioned specific conditions, the use of vibrotactile feedback could be beneficial. In both Paper C and Paper D, I explored the potential of vibrotactile feedback to enhance timbre recognition and music perception respectively.

From a design perspective, my anecdotal experience of speaking with adults with HL provided valuable insights into the preferences and needs of potential users. Adults with HL expressed a strong desire for training tools that are non-intrusive and easy to use. Ideally, they would prefer not to wear or handle any additional devices. For some, having a hearing device is already one device too many, and this perspective can be extended to children as well. These considerations should be taken into account when designing vibrotactile training tools, as the goal is to create solutions that seamlessly integrate into users’ daily lives without adding extra burdens. Moreover, the idea of being dependent on a vibrotactile device for an extended period is not appealing. When talking to clinicians, we agreed that the ultimate goal of using vibrotactile feedback should be to facilitate learning and hearing skill acquisition, enabling users to eventually rely solely on their auditory capabilities. Therefore, training tools should be designed as temporary aids that support users in developing their auditory skills, rather than as permanent fixtures in their lives.

The creation of customized vibrotactile training tools necessitated close collaboration with clinicians to understand individual needs. Conversely, the generalizability of these prototypes was limited, as each child had unique requirements that needed to be addressed individually. The etiologies of HL are highly diverse, making the creation of one-size-fits-all solutions nearly impossible. Furthermore, differences in age, cognitive abilities, and personal preferences add complexity to the design of effective training tools for a generalized population. On a larger scale, these interventions would require significant resources to adapt and personalize the tools for each user, posing challenges in terms of budget and scalability.

Research Question Q2_2

The three vibrotactile prototypes developed during this PhD project were built to explore different aspects of music perception and auditory training. In Paper C, the focus was on enhancing timbre perception through vibrotactile feedback. The main issue encountered during the study was the low engagement of the child in the training sessions, which may have affected the overall effectiveness of the intervention. Upon returning from the training period, the child stated that they did not want to continue, as the video games used for entertainment were far more appealing and enjoyable than the training sessions. I learned the hard way that providing a fun and engaging experience is crucial, but creating one that meets children's expectations is not an easy task. In retrospect, a larger project with an appropriate budget should have been established to develop a more engaging training experience in collaboration with a team of game designers and developers. With a larger workforce, we might have been able to replicate some video game mechanics and aesthetics that are highly appreciated by children. A final note on the timbre training prototype concerns doubts about the proper use of the PlayStation 5 controller for vibrotactile feedback delivery. Despite providing instructions to the child and her parents on how to hold the controller, there is uncertainty regarding whether the child consistently held it correctly during the training sessions. This might have influenced the effectiveness of the vibrotactile feedback and, consequently, the training outcomes. However, this remains speculative, as we do not know exactly what happened during the at-home training sessions. To address this issue in future studies, it would be beneficial to modify the controller to include contact sensors that can detect whether it is being held properly.

Paper D had a different outcome, as the teddy bear was generally well received by the children. Its appealing appearance and the fact that it was a familiar interface made it easy to accept. The study had an exploratory nature, aiming to observe how children interact with vibrotactile feedback in a playful context rather than quantify improvement in a specific musical skill. This approach also made it difficult to draw definitive conclusions about the effectiveness of the intervention. Nevertheless, the qualitative feedback gathered from both children and clinicians provided valuable insights into the potential benefits and challenges of using vibrotactile feedback for music perception. One notable case involved a very young child who initially showed great enthusiasm for the teddy bear and its vibrotactile feedback. However, during subsequent sessions, the child appeared to lose interest and became less engaged with the prototype. This change in attitude could be attributed to various factors, such as the novelty wearing off or individual preferences. Keeping interest in a single toy over time is not a common trait in children, where repeated exposure, habituation, and novelty are decisive factors [129]. One strategy could be to alternate different vibrotactile toys to provide variety while maintaining consistent exposure to sound and vibrotactile feedback. The children involved in the study, in fact, had CND and were not able to benefit fully from auditory feedback alone, and clinicians suggested that sound and vibrotactile exposure would be beneficial for them.

The Tickle Tuner prototype presented in Paper B aimed to explore the potential of vibrotactile feedback for enhancing melodic contour identification. The prototype was generally well received; however, it was not tested with children with HL. Its design may be more suitable for teenagers, as it requires coupling with a smartphone. This prototype could be particularly useful for teenagers with a specific interest in music and severe conditions (e.g., CND) who wish to improve their music perception skills. Furthermore, the prototype could serve as a platform for composing music with the aid of vibrotactile feedback, as the smartphone interface enables a wide range of applications. The recruitment of teenagers with HL for testing the Tickle Tuner would require a different approach compared to recruiting younger children, as the CHBC primarily focuses on younger children that undergo the AVT. Collaborating with hearing associations and leveraging social media platforms could be effective strategies for reaching this demographic.

To sum up, the three vibrotactile prototypes developed during this PhD project provided valuable insights into the potential benefits and challenges of using vibrotactile feedback for music perception and auditory training. Each prototype addressed different aspects of music perception, and their development highlighted the importance of engagement, usability, and individualization in designing effective training tools for children with HC.

5.3 Problem Area 3

Research Question Q3_1

Research question Q3_1 introduces an important aspect of this thesis: the testing and evaluation of the developed prototypes. Recruitment challenges have already been mentioned, but I would like to elaborate further. Despite strong collaboration with the CHBC and support from clinical professionals, recruiting a sufficient number of participants for testing and evaluation proved to be a significant hurdle. Several factors contributed to this difficulty, including the limited pool of eligible participants and logistical challenges faced by families. This issue underscored the importance of establishing robust recruitment strategies and fostering strong relationships with clinical partners to facilitate participant engagement in future studies.

I attempted to overcome these problems through various approaches, such as distributing flyers via the CHBC network, reaching out to local hearing associations, and leveraging social media platforms to disseminate calls for participants. Despite these efforts, recruitment remained a persistent obstacle, highlighting the need for more effective strategies in the future. The recruitment process proved highly time-consuming, and if I could go back in time, I would initiate recruitment alongside prototype development and ideally allocate a budget for hiring a dedicated person for this task. Additionally, I would involve personnel from the clinic to assist with the process, as they have established relationships with potential participants and can facilitate communication and trust. To achieve this, I would ensure that the project budget includes funds specifically allocated for recruitment efforts, recognizing the critical role this plays in the success of research studies involving participants with HL.

Another side of the recruitment challenge relates to the limited time and availability of families. The frantic pace of daily life, combined with the specific needs of children with HL, made it difficult for families to commit to the multiple sessions required for training and evaluation. This limitation affected the depth of testing and refinement that could be conducted on the prototypes, ultimately impacting their effectiveness and usability. The iterative approach necessary for developing effective training also depended on children's motivation and willingness to engage with the prototypes. A single failed attempt could result in a loss of interest, drastically reducing the likelihood of subsequent iterations to improve the solution. For this reason, it is crucial to interview, observe, and involve participants in the early design stages to ensure that the prototypes aligned with their preferences and needs from the very first interaction.

In the context of evaluation, beyond obtaining more participants, I would also delegate the administration of the tests to clinical personnel. This approach would help mitigate potential biases that may arise when the researcher conducting the study is also responsible for administering the tests. My involvement in the process significantly simplified project administration and reduced costs, but most likely influenced participants' responses, either consciously or unconsciously, leading to skewed results. Furthermore, my dual role as both developer and evaluator of the prototypes may have introduced additional biases, as I had a vested interest in the success of the interventions. By having clinicians administer the tests, we can ensure a more objective and standardized evaluation process, ultimately enhancing the validity and reliability of the findings.

Research Question Q3_2

The development and evaluation of the SoundCubes prototype presented in Paper E provided insights into the design and implementation of gamified music training tools for children with HC. The SoundCubes prototype was designed to be engaging and interactive, incorporating game-like elements to motivate children to participate in music training. The evaluation of the prototype revealed several strengths, including its ability to capture children's attention and provide a fun learning experience. To some extent, the SoundCubes demonstrated that it is possible to create a gamified music training tool that is both engaging and effective with I'd say "one-size-fits-many". Of course it is far from being perfect, as the actual home usage was lower than the expected one, even though the qualitative feedback from all the involved parties was positive. In order to sustain long-term engagement, I imagine the SoundCubes would benefit from additional game elements, such as time challenges or competitive features, to keep children motivated over extended periods. This, of course, while keeping in mind the primary goal of spatial hearing training. The prototype provided a reward system in the form of collectible trophies, but further enhancements could be made to make the rewards more varied and thus motivating for children. From a training perspective, results were mixed, with three out of five participants showing improvement in sound localization skills in one condition. The small sample size makes it difficult to draw definitive conclusions about the effectiveness of SoundCubes for spatial hearing training, which underscores the importance of the considerations above for increasing the chances of success in future iterations.

Similar considerations apply to EmotiCubes (Paper F), as it is another gamified training tool developed almost concurrently with SoundCubes. The prototype aimed to enhance vocal emotional recognition through interactive gameplay. The evaluation of EmotiCubes highlighted its potential, with both participants with HL showing measurable improvements in the vocal emotion recognition abilities. However, similar to the SoundCubes, maintaining long-term engagement proved to be a challenge. Future iterations of the EmotiCubes could benefit from incorporating adaptive difficulty levels, allowing the game to adjust to the child's skill level and provide a more personalized experience.

5.4 Personal Reflections and Future directions

I believe this PhD project has only begun to explore the potential of interactive music training tools for children with HC. The prototypes developed throughout this work have provided valuable insights into their design and implementation; however, much remains to be done. Numerous ideas, projects, interactions, and collaborations have shaped this journey, and many more questions have emerged—questions that will likely require many more future PhD projects to address.

One recurring challenge has been participant recruitment for the larger studies necessary to validate the effectiveness of these prototypes. Researchers, myself included, should pay particular attention to this aspect in future projects, as it is critical for obtaining robust and generalizable results. A key step toward overcoming this obstacle will be close collaboration with clinical partners and the allocation of sufficient resources to support recruitment efforts.

Now that the research conducted during this PhD has been consolidated into a single document, it is clear that the prototypes developed are still in an early stage. Future work should focus on refining and improving these tools based on the feedback and insights gathered during the initial evaluation phase.

This may involve addressing technical issues, enhancing usability, sustaining engagement, and incorporating additional features to better meet the needs of children with HC. I would be particularly interested in seeing at least one of these prototypes progress to a clinical trial phase, enabling evaluation on a larger scale. Achieving this would likely require selecting a single prototype and investing significant time and resources into its further development and testing, shifting from the exploratory approach adopted in this PhD to a more confirmatory one.

From a long-term perspective, additional questions arise, such as how to maintain these prototypes in clinical settings, where intensive use over time may necessitate repairs and updates. These considerations are rarely addressed in research projects, yet they are essential for ensuring the success and sustainability of interventions in clinical practice.

I am deeply grateful for the opportunity to conduct this PhD project, which has allowed me to explore the intersection of music, technology, and HC. This journey has been both challenging and rewarding, and I have gained significant insight into the complexities of designing effective training tools for children with HC. The work in this area is far from complete, and I hope that the insights and prototypes developed here will serve as a foundation for future research and development. As WHO predicts [139], HL will become increasingly prevalent in the coming decades, primarily due to population aging and exposure to loud sounds. This trend underscores the importance of developing effective tools for auditory training and rehabilitation.

My hope is that the growing relevance of this topic will attract more researchers, funding, and industry engagement, ultimately leading to a broader range of solutions and improved outcomes for individuals with HC, enabling them to freely choose to engage with music in their lives.

References

[1] N. F. Ab Shukor, W. Han, J. Lee, and Y. J. Seo, “Crucial Music Components Needed for Speech Perception Enhancement of Pediatric Cochlear Implant Users: A Systematic Review and Meta-Analysis,” *Audiology and Neurotology*, vol. 26, no. 6, pp. 389–413, 2021.

[2] H. Abdulbaki, J. Mo, C. J. Limb, and N. T. Jiam, “The Impact of Musical Rehabilitation on Complex Sound Perception in Cochlear Implant Users: A Systematic Review,” *Otology & Neurotology*, Sep. 2023.

[3] A. Adilkhanov, M. Rubagotti, and Z. Kappassov, “Haptic Devices: Wearability-Based Taxonomy and Literature Review,” *IEEE Access*, vol. 10, pp. 91 923–91 947, 2022.

[4] A. A. Akbulut, A. Karaman Demirel, and A. Çiprut, “Music Perception and Music-Related Quality of Life in Adult Cochlear Implant Users: Exploring the Need for Music Rehabilitation,” *Ear & Hearing*, Sep. 2024.

[5] S. C. Aker, K. F. Faulkner, H. Innes-Brown, M. Vatti, and J. Marozeau, “Some, but not all, cochlear implant users prefer music stimuli with congruent haptic stimulation,” *The Journal of the Acoustical Society of America*, vol. 155, no. 5, pp. 3101–3117, May 2024.

[6] 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, Dec. 2022.

[7] A. Alahmadi, Y. Abdelsamad, N. A. AlAmari, M. Y. Alyousef, M. Al-Momani, F. N. Altamimi, S. F. Alhabib, and A. Hagr, “Hearing implants in pediatrics with cochlear nerve deficiency: An updated systematic review,” *European Archives of Oto-Rhino-Laryngology*, Dec. 2024.

[8] I. L. Alter, A. Chern, M. W. Denham, and A. K. Lalwani, “Improving Music Appreciation and Enjoyment in Hearing Aid Users,” *The Hearing Journal*, vol. 77, no. 1, pp. 12,13,15, Jan. 2024.

[9] G.-L. Alvaro, V. B. Ricardo, S. P. Jose Manuel, O. Tomás, and C. Víctor, “New haptic systems for elicit emotions in audio-visual events for hearing impaired people.” *Procedia Computer Science*, vol. 237, pp. 533–543, 2024.

[10] M. Alzaher, N. Vannson, O. Deguine, M. Marx, P. Barone, and K. Strelnikov, “Brain plasticity and hearing disorders,” *Revue Neurologique*, vol. 177, no. 9, pp. 1121–1132, Nov. 2021.

[11] S. R. Anderson, R. Jocewicz, A. Kan, J. Zhu, S. Tzeng, and R. Y. Litovsky, “Sound source localization patterns and bilateral cochlear implants: Age at onset of deafness effects,” *PLOS ONE*, vol. 17, no. 2, p. e0263516, Feb. 2022.

[12] F. Asp, E. Karltorp, and E. Berninger, “Development of Sound Localization in Infants and Young Children with Cochlear Implants,” *Journal of Clinical Medicine*, vol. 11, no. 22, p. 6758, Nov. 2022.

[13] A. M. AuBuchon, D. B. Pisoni, and W. G. Kronenberger, “Elementary Cognitive Processes Underlying Verbal Working Memory in Pre-lingually Deaf Children with Cochlear Implants,” in *Pediatric Cochlear Implantation: Learning and the Brain*, N. M. Young and K. Iler Kirk, Eds. New York, NY: Springer New York, 2016, pp. 257–273.

[14] E. E. Babajanian and R. K. Gurgel, “Cognitive and behavioral effects of hearing loss,” *Current Opinion in Otolaryngology & Head & Neck Surgery*, vol. 30, no. 5, pp. 339–343, Oct. 2022.

[15] M. D. Binder, N. Hirokawa, and U. Windhorst, Eds., *Encyclopedia of Neuroscience*, ser. SpringerLink Bücher. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[16] F. Bleckly, C. Y. Lo, F. Rapport, and R. Clay-Williams, “Music Perception, Appreciation, and Participation in Postlingually Deafened Adults and Cochlear Implant Users: A Systematic Literature Review,” *Trends in Hearing*, vol. 28, p. 23312165241287391, Jan. 2024.

[17] E. D. Borre, K. Kaalund, N. Frisco, G. Zhang, A. Ayer, M. Kelly-Hedrick, S. D. Reed, S. D. Emmett, H. Francis, D. L. Tucci, B. S. Wilson, A. S. Kosinski, O. Ogbuoji, and G. D. Sanders Schmidler, “The Impact of Hearing Loss and Its Treatment on Health-Related Quality of Life Utility: A Systematic Review with

Meta-analysis," *Journal of General Internal Medicine*, vol. 38, no. 2, pp. 456–479, Feb. 2023.

[18] I. Bortone, D. Leonardis, N. Mastronicola, A. Crecchi, L. Bonfiglio, C. Procopio, M. Solazzi, and A. Frisoli, "Wearable Haptics and Immersive Virtual Reality Rehabilitation Training in Children With Neuromotor Impairments," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 7, pp. 1469–1478, Jul. 2018.

[19] R. I. Busarello, V. R. Ulbricht, L. M. Fadel, and A. V. de Freitas e Lopes, "Gamification approaches to learning and knowledge development: A theorical review," in *New Advances in Information Systems and Technologies.* Springer, 2016, pp. 1107–1116.

[20] C. Camardella, D. Chiaradia, I. Bortone, A. Frisoli, and D. Leonardis, "Introducing wearable haptics for rendering velocity feedback in VR serious games for neuro-rehabilitation of children," *Frontiers in Virtual Reality*, vol. 3, p. 1019302, Jan. 2023.

[21] G. Cambridge, T. Taylor, W. Arnott, and W. J. Wilson, "Auditory training for adults with cochlear implants: A systematic review," *International Journal of Audiology*, vol. 61, no. 11, pp. 896–904, Nov. 2022.

[22] S. Cano, J. S. Naranjo, C. Henao, C. Rusu, and S. Albiol-Pérez, "Serious Game as Support for the Development of Computational Thinking for Children with Hearing Impairment," *Applied Sciences*, vol. 11, no. 1, p. 115, 2021.

[23] M. L. Cantuaria, E. R. Pedersen, F. B. Waldorff, L. Wermuth, K. M. Pedersen, A. H. Poulsen, O. Raaschou-Nielsen, M. Sørensen, and J. H. Schmidt, "Hearing Loss, Hearing Aid Use, and Risk of Dementia in Older Adults," *JAMA Otolaryngology–Head & Neck Surgery*, vol. 150, no. 2, p. 157, Feb. 2024.

[24] M. G. Carpenter and J. L. Campos, "The Effects of Hearing Loss on Balance: A Critical Review," *Ear & Hearing*, vol. 41, no. Supplement 1, pp. 107S–119S, Nov. 2020.

[25] Y.-T. Chen, D. de Thomas Wagner, A. J. Loutit, A. Nourizonoz, M. Ulanova, D. Graikou, M.-S. Chang, M.-C. Croisier-Coeytaux, S. Clerc-Rosset, J. Blanc, G. Knott, K.-S. Lee, and D. Huber, "Lamellar Schwann cells in the Pacinian corpuscle potentiate vibration perception," *Science Advances*, vol. 11, no. 24, p. eadt5110, 2025.

[26] S. Chennareddy, K. H. Liu, M. A. Mavrommatis, D. D. Kao, A. Govindan, Z. G. Schwam, and M. K. Cosetti, "Cochlear Implantation in Pediatric Cochlear Nerve Deficiency: A Systematic Review and Meta-Analysis," *The Laryngoscope*, vol. 135, no. 4, pp. 1259–1266, Apr. 2025.

[27] J. A. Christensen, J. Sis, A. M. Kulkarni, and M. Chatterjee, “Effects of Age and Hearing Loss on the Recognition of Emotions in Speech,” *Ear & Hearing*, vol. 40, no. 5, pp. 1069–1083, Sep. 2019.

[28] A. Christopoulos and S. Mystakidis, “Gamification in Education,” *Encyclopedia*, vol. 3, no. 4, pp. 1223–1243, Oct. 2023.

[29] K. Cieśla, T. Wolak, A. Lorens, M. Mentzel, H. Skarżyński, and A. Amedi, “Effects of training and using an audio-tactile sensory substitution device on speech-in-noise understanding,” *Scientific Reports*, vol. 12, no. 1, p. 3206, Dec. 2022.

[30] A. Ciorba, C. Bianchini, S. Pelucchi, and A. Pastore, “The impact of hearing loss on the quality of life of elderly adults,” *Clinical Interventions in Aging*, p. 159, Jun. 2012.

[31] Q. Consigny, N. Ouvrai, A. Paté, C. Fritz, and J.-L. Le Carrou, “Vibrotactile Thresholds on the Wrist for Vibrations Normal to the Skin,” *IEEE Transactions on Haptics*, vol. 16, no. 4, pp. 628–633, Oct. 2023.

[32] R. S. Cowan, J. Z. Sarant, K. L. Galvin, J. I. Alcantara, P. J. Blamey, and G. M. Clark, “The Tickle Talker: A speech perception aid for profoundly hearing impaired children,” *Scientific publications, vol. 5, 1989-1990, no. 300*, 1990.

[33] L. E. Crommett, D. Madala, and J. M. Yau, “Multisensory perceptual interactions between higher-order temporal frequency signals.” *Journal of Experimental Psychology: General*, vol. 148, no. 7, pp. 1124–1137, Jul. 2019.

[34] L. E. Crommett, A. Pérez-Bellido, and J. M. Yau, “Auditory adaptation improves tactile frequency perception,” *Journal of Neurophysiology*, vol. 117, no. 3, pp. 1352–1362, Mar. 2017.

[35] J. Dammeyer and S. E. Ohna, “Changes in Educational Planning for Deaf and Hard of Hearing Children in Scandinavia over the Last Three Decades,” *Scandinavian Journal of Disability Research*, vol. 23, no. 1, pp. 114–123, Apr. 2021.

[36] F. Darki, J. Rankin, and P. Słowiński, “Tactile stimulations reduce or promote the segregation of auditory streams: Psychophysics and modeling,” *PLOS Computational Biology*, vol. 21, no. 8, p. e1012701, Aug. 2025.

[37] R. D. De Siati, F. Rosenzweig, G. Gersdorff, A. Gregoire, P. Rombaux, and N. Deggouj, “Auditory Neuropathy Spectrum Disorders: From Diagnosis to Treatment: Literature Review and Case Reports,” *Journal of Clinical Medicine*, vol. 9, no. 4, p. 1074, Apr. 2020.

[38] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, “From game design elements to gamefulness: Defining ”gamification”,” in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 2011, pp. 9–15.

[39] J. R. Dornhoffer, C. M. Lohse, T. N. Tamati, A. C. Moberly, and M. L. Carlson, “Current practices and opinions on auditory training in adult cochlear implant recipients,” *American Journal of Otolaryngology*, vol. 45, no. 4, p. 104339, Jul. 2024.

[40] Ward. 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, no. 2, pp. 114–123, Feb. 2015.

[41] J. J. Eggermont, “Neuroplasticity of the Auditory System,” in *Textbook of Tinnitus*, W. Schlee, B. Langguth, D. De Ridder, S. Vanneste, T. Kleinjung, and A. R. Møller, Eds. Cham: Springer International Publishing, 2024, pp. 149–163.

[42] M. Ekberg, J. Andin, C. Signoret, S. Stenfelt, and Ö. Dahlström, “Effects of mild-to-moderate sensorineural hearing loss and signal amplification on vocal emotion recognition in middle-aged–older individuals,” *PLOS ONE*, vol. 20, no. 7, pp. 1–26, Jul. 2025.

[43] W. Estabrooks, H. M. Morrison, and K. MacIver-Lux, *Auditory-Verbal Therapy: Science, Research, and Practice.* San Diego: Plural Publishing, Incorporated, 2020.

[44] M. Filippa, D. Lima, A. Grandjean, C. Labbé, S. Y. Coll, E. Gentaz, and D. M. Grandjean, “Emotional prosody recognition enhances and progressively complexifies from childhood to adolescence,” *Scientific Reports*, vol. 12, no. 1, p. 17144, Oct. 2022.

[45] M. D. Fletcher, E. Akis, C. A. Verschuur, and S. W. Perry, “Improved tactile speech perception and noise robustness using audio-to-tactile sensory substitution with amplitude envelope expansion,” *Scientific Reports*, vol. 14, no. 1, p. 15029, Jul. 2024.

[46] M. D. Fletcher, R. O. Cunningham, and S. R. Mills, “Electro-Haptic Enhancement of Spatial Hearing in Cochlear Implant Users,” *Scientific Reports*, vol. 10, no. 1, p. 1621, Dec. 2020.

[47] M. D. Fletcher and J. Zgheib, “Haptic sound-localisation for use in cochlear implant and hearing-aid users.” *Scientific reports*, vol. 10, no. 1, p. 14171, Aug. 2020.

[48] A. Flores Ramones and M. S. del-Rio-Guerra, “Recent Developments in Haptic Devices Designed for Hearing-Impaired People: A Literature Review,” *Sensors*, vol. 23, no. 6, p. 2968, Mar. 2023.

[49] F. Fontana, I. Camponogara, P. Cesari, M. Vallicella, and M. Ruzzenente, “An exploration on whole-body and foot-based vibrotactile sensitivity to melodic consonance,” in *SMC 2016 - 13th Sound and Music Computing Conference, Proceedings*, ser. SMC 2016 - 13th Sound and Music Computing Conference, Proceedings, R. Grossmann and G. Hajdu, Eds. Zentrum fur Mikrotonale Musik und Multimediale Komposition (ZM4), Hochschule fur Musik und Theater, 2016, pp. 143–150.

[50] A. Frosolini, L. Franz, G. Badin, A. Mancuso, C. De Filippis, and G. Marioni, “Quality of life improvement in Cochlear implant outpatients: A nonrandomized clinical trial of an auditory music training program,” *International Journal of Audiology*, pp. 1–9, Nov. 2024.

[51] A. Gallace, C. Spence, and C. Spence, *In Touch with the Future: The Sense of Touch from Cognitive Neuroscience to Vitual Reality.* Oxford: Oxford University Press, 2014.

[52] K. L. Galvin, P. J. Blamey, R. S. C. Cowan, M. Oerlemans, and G. M. Clark, “Generalization of tactile perceptual skills to new context following tactile-alone word recognition training with the Tickle Talker™,” *The Journal of the Acoustical Society of America*, vol. 108, no. 6, pp. 2969–2979, Dec. 2000.

[53] K. L. Galvin, P. J. Blamey, M. Oerlemans, R. S. C. Cowan, and G. M. Clark, “Acquisition of a tactile-alone vocabulary by normally hearing users of the Tickle Talker™,” *The Journal of the Acoustical Society of America*, vol. 106, no. 2, pp. 1084–1089, Aug. 1999.

[54] C. Gao, J. J. Green, X. Yang, S. Oh, J. Kim, and S. V. Shinkareva, “Audiovisual integration in the human brain: A coordinate-based meta-analysis,” *Cerebral Cortex*, vol. 33, no. 9, pp. 5574–5584, Apr. 2023.

[55] S. N. Garadat, “Development and Beta Testing of Serious Game-Based Auditory Training Application to Enhance Perceptual Learning of Speech in Cochlear Implant Recipients,” *American Journal of Audiology*, vol. 32, no. 2, pp. 261–273, Jun. 2023.

[56] 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*, vol. 13, p. 1229, Nov. 2019.

[57] C. Giulia, D. V. Chiara, and H. Esmailbeigi, “GLOS: GLOve for Speech Recognition,” in *2019 41st Annual International Conference of the IEEE Engineering in*

Medicine and Biology Society (EMBC). Berlin, Germany: IEEE, Jul. 2019, pp. 3319–3322.

[58] H. Glick and A. Sharma, "Cross-modal plasticity in developmental and age-related hearing loss: Clinical implications," *Hearing Research*, vol. 343, pp. 191–201, Jan. 2017.

[59] N. R. R. Gomes, E. C. Maia, and I. V. D. Varga, "The benefits of playing for children's health: A systematic review," *Arquivos de Ciências da Saúde*, vol. 25, no. 2, p. 47, Jul. 2018.

[60] A. Good, K. A. Gordon, B. C. Papsin, G. Nespoli, T. Hopyan, I. Peretz, and F. A. Russo, "Benefits of Music Training for Perception of Emotional Speech Prosody in Deaf Children With Cochlear Implants," *Ear & Hearing*, vol. 38, no. 4, pp. 455–464, Jul. 2017.

[61] M. Gori, T. Vercillo, G. Sandini, and D. Burr, "Tactile feedback improves auditory spatial localization," *Frontiers in Psychology*, vol. 5, Oct. 2014.

[62] A. Greasley, H. Crook, and R. Fulford, "Music listening and hearing aids: Perspectives from audiologists and their patients," *International Journal of Audiology*, vol. 59, no. 9, pp. 694–706, Sep. 2020.

[63] P. Guilleminot and T. Reichenbach, "Enhancement of speech-in-noise comprehension through vibrotactile stimulation at the syllabic rate," *Proceedings of the National Academy of Sciences*, vol. 119, no. 13, p. e2117000119, Mar. 2022.

[64] Gunhyuk Park and Seungmoon Choi, "Perceptual space of amplitude-modulated vibrotactile stimuli," in *2011 IEEE World Haptics Conference.* Istanbul: IEEE, Jun. 2011, pp. 59–64.

[65] L. Hammer, N. R. Kamper, L. Jantzen, S. Serafin, and L. Percy-Smith, "Self-reported social well-being of children with hearing loss in 2023," *Cochlear Implants International*, pp. 1–7, May 2024.

[66] A. R. Hitchins and S. C. Hogan, "Outcomes of early intervention for deaf children with additional needs following an Auditory Verbal approach to communication," *International Journal of Pediatric Otorhinolaryngology*, vol. 115, pp. 125–132, Dec. 2018.

[67] T. M. Hopyan-Misakyan, K. A. Gordon, M. Dennis, and B. C. Papsin, "Recognition of Affective Speech Prosody and Facial Affect in Deaf Children with Unilateral Right Cochlear Implants," *Child Neuropsychology*, vol. 15, no. 2, pp. 136–146, Mar. 2009.

[68] X. Hou, O. Sourina, and S. Klimenko, "Haptic-Based Serious Games," in *2014 International Conference on Cyberworlds.* Santander, Cantabria, Spain: IEEE, Oct. 2014, pp. 39–46.

[69] R. Jack, D. A. Mcpherson, and D. T. Stockman, “Designing Tactile Musical Devices with and for Deaf Users: A Case Study,” p. 7, 2015.

[70] S. Jeong, H. H. Yun, Y. Lee, and Y. Han, “Glow the Buzz: A VR Puzzle Adventure Game Mainly Played Through Haptic Feedback,” in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. Hamburg Germany: ACM, Apr. 2023, pp. 1–6.

[71] L. Jürgensen, F. Denk, and H. Husstedt, “Visualization and analysis of multi-channel dynamic range compression in hearing aids,” *Acta Acustica*, vol. 8, p. 5, 2024.

[72] P. N. Juslin and P. Laukka, “Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion.” *Emotion*, vol. 1, no. 4, pp. 381–412, 2001.

[73] E. C. Kaplan, A. E. Wagner, P. Toffanin, and D. Başkent, “Do Musicians and Non-musicians Differ in Speech-on-Speech Processing?” *Frontiers in Psychology*, vol. 12, p. 623787, Feb. 2021.

[74] M. Karam, F. Russo, and D. Fels, “Designing the Model Human Cochlea: An Ambient Crossmodal Audio-Tactile Display,” *IEEE Transactions on Haptics*, vol. 2, no. 3, pp. 160–169, Jul. 2009.

[75] M. Karam, C. Branje, G. Nespoli, N. Thompson, F. A. Russo, and D. I. Fels, “The emoti-chair: An interactive tactile music exhibit,” in *CHI ’10 Extended Abstracts on Human Factors in Computing Systems*. Atlanta Georgia USA: ACM, Apr. 2010, pp. 3069–3074.

[76] N. E. Kepp, C. Schiøth, and L. Percy-Smith, “Timbre recognition in Danish children with hearing aids, cochlear implants or normal hearing,” *International Journal of Pediatric Otorhinolaryngology*, vol. 159, p. 111186, Aug. 2022.

[77] J.-T. Kim, H.-S. Shin, J.-Y. Yoo, R. Avila, Y. Huang, Y. H. Jung, J. E. Colgate, and J. A. Rogers, “Mechanics of vibrotactile sensors for applications in skin-interfaced haptic systems,” *Extreme Mechanics Letters*, vol. 58, p. 101940, Jan. 2023.

[78] B. J. Lawrence, D. M. P. Jayakody, R. J. Bennett, R. H. Eikelboom, N. Gasson, and P. L. Friedland, “Hearing Loss and Depression in Older Adults: A Systematic Review and Meta-analysis,” *The Gerontologist*, vol. 60, no. 3, pp. e137–e154, Apr. 2020.

[79] K. Liddle, E. J. Fitzgibbons, R. Beswick, and C. Driscoll, “Cochlear nerve deficiency is an important cause of auditory neuropathy spectrum disorder at a population level in children,” *International Journal of Pediatric Otorhinolaryngology*, vol. 158, p. 111171, Jul. 2022.

[80] J. E. C. Lieu, M. Kenna, S. Anne, and L. Davidson, "Hearing Loss in Children: A Review," *JAMA*, vol. 324, no. 21, p. 2195, Dec. 2020.

[81] F. R. Lin, "Age-Related Hearing Loss," *New England Journal of Medicine*, vol. 390, no. 16, pp. 1505–1512, Apr. 2024.

[82] C. Y. Lo, V. Looi, W. F. Thompson, and C. M. McMahon, "Music Training for Children With Sensorineural Hearing Loss Improves Speech-in-Noise Perception," *Journal of Speech, Language, and Hearing Research*, vol. 63, no. 6, pp. 1990–2015, Jun. 2020.

[83] ——, "Beyond Audition: Psychosocial Benefits of Music Training for Children With Hearing Loss," *Ear & Hearing*, vol. 43, no. 1, pp. 128–142, Jan. 2022.

[84] J. Lo¨fvenberg and R. Johansson, "Regional differences and interindividual variability in sensitivity to vibration in the glabrous skin of the human hand," *Brain Research*, vol. 301, no. 1, pp. 65–72, May 1984.

[85] P. C. Loizou, "Speech Processing in Vocoder-Centric Cochlear Implants," in *Advances in Oto-Rhino-Laryngology*, A. Møller, Ed. Basel: S. KARGER AG, 2006, vol. 64, pp. 109–143.

[86] 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, p. 1145, Oct. 2019.

[87] O. Macherey and R. P. Carlyon, "Cochlear implants," *Current Biology*, vol. 24, no. 18, pp. R878–R884, Sep. 2014.

[88] V. Marian, S. Hayakawa, and S. R. Schroeder, "Cross-Modal Interaction Between Auditory and Visual Input Impacts Memory Retrieval," *Frontiers in Neuroscience*, vol. 15, p. 661477, Jul. 2021.

[89] J. Marozeau, "An Artistic Audiotactile Installation Augmented Music," in *Proceedings of the 20th Sound and Music Computing Conference*, Stockholm, Sweden, 2023, pp. 419–425.

[90] H. McGurk and J. MacDonald, "Hearing lips and seeing voices," *Nature*, vol. 264, no. 5588, pp. 746–748, 1976.

[91] M. Meral Çetinkaya, Ö. Konukseven, and A. E. İralı, "World of sounds (Seslerin Dünyası): A mobile auditory training game for children with cochlear implants," *International Journal of Pediatric Otorhinolaryngology*, vol. 179, p. 111908, Apr. 2024.

[92] S. Merchel and M. E. Altinsoy, "Psychophysical comparison of the auditory and tactile perception: A survey," *Journal on Multimodal User Interfaces*, vol. 14, pp. 271–283, 2020.

[93] J. C. Middlebrooks, "Sound localization," in *Handbook of Clinical Neurology*. Elsevier, 2015, vol. 129, pp. 99–116.

[94] R. E. Mitchell and M. A. Karchmer, "Chasing the Mythical Ten Percent: Parental Hearing Status of Deaf and Hard of Hearing Students in the United States," *Sign Language Studies*, vol. 4, no. 2, pp. 138–163, Dec. 2004.

[95] A. R. Møller, "The Ear," in *Hearing: Anatomy, Physiology, and Disorders of the Auditory System*, 3rd ed. San Diego, CA: Plural Publishing, 2013, pp. 15–89.

[96] B. C. J. Moore, "The perception of emotion in music by people with hearing loss and people with cochlear implants," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 379, no. 1908, p. 20230258, Aug. 2024.

[97] F. E. Musiek and J. A. Baran, *The Auditory System : Anatomy, Physiology, and Clinical Correlates*, second edition. ed. San Diego, CA: Plural Publishing, Inc., 2020.

[98] N. Natarajan, S. Batts, and K. M. Stankovic, "Noise-Induced Hearing Loss," *Journal of Clinical Medicine*, vol. 12, no. 6, p. 2347, Mar. 2023.

[99] K. Niu, A. Brandström, S. Skenbäck, M. Duan, and I. Uhlén, "Risk factors and etiology of childhood hearing loss: A cohort review of 296 subjects," *Acta Oto-Laryngologica*, vol. 140, no. 8, pp. 660–666, Jul. 2020.

[100] U. Noppeney, "Perceptual Inference, Learning, and Attention in a Multisensory World," *Annual Review of Neuroscience*, vol. 44, no. 1, pp. 449–473, Jul. 2021.

[101] A. Olson, "Options for Auditory Training for Adults with Hearing Loss," *Seminars in Hearing*, vol. 36, no. 04, pp. 284–295, Oct. 2015.

[102] R. Paisa, N. C. Nilsson, and S. Serafin, "Tactile displays for auditory augmentation–A scoping review and reflections on music applications for hearing impaired users," *Frontiers in Computer Science*, vol. 5, p. 1085539, Apr. 2023.

[103] S. Park, K. H. Park, and W. Han, "The Effects of Music-Based Auditory Training on Hearing-Impaired Older Adults With Mild Cognitive Impairment," *Clinical and Experimental Otorhinolaryngology*, vol. 17, no. 1, pp. 26–36, Feb. 2024.

[104] A. N. Pearce, P. Eckard, A. R. Mangan, A. Baggett, J. Dornhoffer, and R. A. Saadi, "Cochlear implantation in cochlear nerve deficiency: A systematic review," *American Journal of Otolaryngology*, vol. 46, no. 5, p. 104694, Sep. 2025.

[105] J. E. Peelle and A. Wingfield, “The Neural Consequences of Age-Related Hearing Loss,” *Trends in Neurosciences*, vol. 39, no. 7, pp. 486–497, Jul. 2016.

[106] L. Percy-Smith, G. Busch, M. Sandahl, L. Nissen, J. L. Josvassen, T. Lange, E. Rusch, and P. Cayé-Thomasen, “Language understanding and vocabulary of early cochlear implanted children,” *International Journal of Pediatric Otorhinolaryngology*, vol. 77, no. 2, pp. 184–188, Feb. 2013.

[107] L. Percy-Smith, L. Jantzen, and P. Cayé-Thomasen, “Cochleaimplantation Til Børn,” https://ugeskriftet.dk/videnskab/cochleaimplantation-til-born, Dec. 2023.

[108] L. Percy-Smith, T. L. Tønning, J. L. Josvassen, J. H. Mikkelsen, L. Nissen, E. Dieleman, M. Hallstrøm, and P. Cayé-Thomasen, “Auditory verbal habilitation is associated with improved outcome for children with cochlear implant,” *Cochlear Implants International*, vol. 19, no. 1, pp. 38–45, Jan. 2018.

[109] N. K. Petersen, A. W. Jørgensen, and T. Ovesen, “Prevalence of various etiologies of hearing loss among cochlear implant recipients: Systematic review and meta-analysis,” *International Journal of Audiology*, vol. 54, no. 12, pp. 924–932, Dec. 2015.

[110] C. J. Plack, *The Sense of Hearing*, third edition ed. Abingdon, Oxon New York, NY: Routledge, 2018.

[111] G. R. Popelka, B. C. J. Moore, R. R. Fay, and A. N. Popper, Eds., *Hearing Aids*, ser. Springer Handbook of Auditory Research. Cham: Springer International Publishing, 2016, vol. 56.

[112] L. Rachman, G. Babaoğlu, B. Özkişi Yazgan, P. Ertürk, E. Gaudrain, L. Nagels, S. Launer, P. Derleth, G. Singh, F. Uhlemayr, M. Chatterjee, E. Yücel, G. Sennaroğlu, and D. Başkent, “Vocal Emotion Recognition in School-Age Children With Hearing Aids,” *Ear & Hearing*, Mar. 2025.

[113] G. Rance, B. Cone-Wesson, J. Wunderlich, and R. Dowell, “Speech Perception and Cortical Event Related Potentials in Children with Auditory Neuropathy,” *Ear and Hearing*, vol. 23, no. 3, pp. 239–253, Jun. 2002.

[114] F. Rochette, A. Moussard, and E. Bigand, “Music Lessons Improve Auditory Perceptual and Cognitive Performance in Deaf Children,” *Frontiers in Human Neuroscience*, vol. 8, Jul. 2014.

[115] J. M. Rodríguez-Ferrer, A. Manzano-León, J. M. Aguilar-Parra, and A. Cangas, “Effectiveness of gamification and game-based learning in Spanish adolescents with dyslexia: A longitudinal quasi-experimental research,” *Research in Developmental Disabilities*, vol. 141, p. 104603, Oct. 2023.

[116] F. A. Russo, P. Ammirante, and D. I. Fels, “Vibrotactile discrimination of musical timbre.” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 38, no. 4, pp. 822–826, 2012.

[117] A. Schulte, J. Marozeau, A. Ruhe, A. Büchner, A. Kral, and H. Innes-Brown, “Improved speech intelligibility in the presence of congruent vibrotactile speech input,” *Scientific Reports*, vol. 13, no. 1, p. 22657, Dec. 2023.

[118] S. Shin, S. Baek, S. Park, H. Yi, H. Shin, and W. Lee, “Enhancing Pitch Perception and Melodic Reproduction of Hard-of-Hearing Individuals Using a Vibrotactile Glove,” *International Journal of Human–Computer Interaction*, vol. 41, no. 19, pp. 12 032–12 049, Oct. 2025.

[119] K. Shinohara and J. O. Wobbrock, “In the shadow of misperception: Assistive technology use and social interactions,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* Vancouver BC Canada: ACM, May 2011, pp. 705–714.

[120] 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, no. 1, pp. 15–28, Feb. 2021.

[121] M. C. Sierra, J. Brunskog, and J. Marozeau, “An Audio-Tactile Artinstallation Hearing Impaired People,” in *2nd Nordic Sound and Music Conference.* Copenhagen: Zenodo, 2021.

[122] T. Singhal and O. Schneider, “Juicy haptic design: Vibrotactile embellishments can improve player experience in games,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–11.

[123] K. Slade, C. J. Plack, and H. E. Nuttall, “The Effects of Age-Related Hearing Loss on the Brain and Cognitive Function,” *Trends in Neurosciences*, vol. 43, no. 10, pp. 810–821, Oct. 2020.

[124] 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, no. 1, pp. 51–70, Dec. 1990.

[125] M. Stropahl, J. Besser, and S. Launer, “Auditory Training Supports Auditory Rehabilitation: A State-of-the-Art Review,” *Ear & Hearing*, vol. 41, no. 4, pp. 697–704, Oct. 2019.

[126] A. Sundqvist, B. Lyxell, R. Jönsson, and M. Heimann, “Understanding minds: Early cochlear implantation and the development of theory of mind in children with profound hearing impairment,” *International Journal of Pediatric Otorhinolaryngology*, vol. 78, no. 3, pp. 538–544, Mar. 2014.

[127] 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, Jan. 2024.

[128] H. Z. Tan, C. M. Reed, Y. Jiao, Z. D. Perez, E. C. Wilson, J. Jung, J. S. Martinez, and F. M. Severgnini, “Acquisition of 500 English Words through a TActile Phonemic Sleeve (TAPS).” *IEEE transactions on haptics*, vol. 13, no. 4, pp. 745–760, Dec. 2020.

[129] J. Trawick-Smith, J. Wolff, M. Koschel, and J. Vallarelli, “Effects of Toys on the Play Quality of Preschool Children: Influence of Gender, Ethnicity, and Socioeconomic Status,” *Early Childhood Education Journal*, vol. 43, no. 4, pp. 249–256, Jul. 2015.

[130] M. V. Valenza and I. Gasparini, “Serious Game Design for Children,” *Journal of Educational Technology & Society*, vol. 22, no. 3, p. 14, 2022.

[131] L. Van Bogaert, L. Machart, S. Gerber, H. Lœvenbruck, A. Vilain, and Consortium EULALIES, “Speech rehabilitation in children with cochlear implants using a multisensory (French Cued Speech) or a hearing-focused (Auditory Verbal Therapy) approach,” *Frontiers in Human Neuroscience*, vol. 17, p. 1152516, May 2023.

[132] R. T. Verrillo and S. J. Bolanowski, “Tactile Responses to Vibration,” in *Handbook of Signal Processing in Acoustics*, D. Havelock, S. Kuwano, and M. Vorländer, Eds. New York, NY: Springer New York, 2008, pp. 1185–1213.

[133] R. T. Verrillo, A. J. Fraioli, and R. L. Smith, “Sensation magnitude of vibrotactile stimuli,” *Perception & Psychophysics*, vol. 6, no. 6, pp. 366–372, Nov. 1969.

[134] C. I. Von Eiff, V. G. Skuk, R. Zäske, C. Nussbaum, S. Frühholz, U. Feuer, O. Guntinas-Lichius, and S. R. Schweinberger, “Parameter-Specific Morphing Reveals Contributions of Timbre to the Perception of Vocal Emotions in Cochlear Implant Users,” *Ear & Hearing*, vol. 43, no. 4, pp. 1178–1188, Jul. 2022.

[135] S. Wallace, *A Dictionary of Education*, 2nd ed. Oxford University Press, 2015.

[136] O. B. Wie, J. V. K. Torkildsen, S. Schauber, T. Busch, and R. Litovsky, “Long-Term Language Development in Children With Early Simultaneous Bilateral Cochlear Implants,” *Ear & Hearing*, vol. 41, no. 5, pp. 1294–1305, Sep. 2020.

[137] S. Wischmann, J. L. Josvassen, C. Schiøth, and L. Percy-Smith, “History rewritten for children with hearing impairment,” *International Journal of Pediatric Otorhinolaryngology*, vol. 152, p. 110991, Jan. 2022.

[138] S. Wischmann, C. F. Samar, M. K. Hestbæk, S. Serafin, P. Cayé-Thomasen, and L. Percy-Smith, “Quality Assurance of a Hospital-Based Auditory Verbal

Intervention for Children with Hearing Loss," *Journal of Clinical Medicine*, vol. 14, no. 4, p. 1098, Feb. 2025.

[139] World Health Organization, "WHO - Deafness and Hearing Loss," https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss, Feb. 2025.

[140] Y. Xiang, Z. Zhang, D. Chang, and L. Tu, "The Impact of Gamified Auditory–Verbal Training for Hearing-Challenged Children at Intermediate and Advanced Rehabilitation Stages," *Games for Health Journal*, vol. 13, no. 5, pp. 365–378, Oct. 2024.

[141] G. Young, D. Murphy, and J. Weeter, "Audio-tactile glove," in *Proceedings of the International Conference on Digital Audio Effects (Dafx), Maynooth, Ireland*, 2013.

[142] N. M. Young and K. Iler Kirk, Eds., *Pediatric Cochlear Implantation: Learning and the Brain.* New York, NY: Springer New York, 2016.

[143] G. Yun, M. Mun, J. Lee, D.-G. Kim, H. Z. Tan, and S. Choi, "Generating Real-Time, Selective, and Multimodal Haptic Effects from Sound for Gaming Experience Enhancement," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems.* Hamburg Germany: ACM, Apr. 2023, pp. 1–17.

[144] F.-G. Zeng, "Celebrating the one millionth cochlear implant," *JASA Express Letters*, vol. 2, no. 7, p. 077201, Jul. 2022.

[145] 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, no. 1, pp. 377–386, Jan. 2002.

[146] F.-G. Zeng, Y.-Y. Kong, H. J. Michalewski, and A. Starr, "Perceptual Consequences of Disrupted Auditory Nerve Activity," *Journal of Neurophysiology*, vol. 93, no. 6, pp. 3050–3063, Jun. 2005.

[147] Y. Zhou, K. C. Sim, P. Tan, and Y. Wang, "MOGAT: Mobile games with auditory training for children with cochlear implants," in *Proceedings of the 20th ACM International Conference on Multimedia.* Nara Japan: ACM, Oct. 2012, pp. 429–438.

Part II

Included Papers

Paper A

The Role of Haptics in Training and Games for Hearing-Impaired Individuals: A Systematic Review

F. Ganis, A. Gulli, F. Fontana, and S. Serafin

The article has been submitted to
Multimodal Technology Interaction, 2024.

© 2024, The Authors, under license to MDPI, Basel, Switzerland - CC 4.0

The layout has been revised.

Original Publication

F. Ganis, A. Gulli, F. Fontana, and S. Serafin, “The Role of Haptics in Training and Games for Hearing-Impaired Individuals: A Systematic Review,” *Multimodal Technologies and Interaction*, vol. 8, no. 1, p. 1, Dec. 2023, doi: 10.3390/mti8010001.

Abstract

Sensory substitution and augmentation are pivotal concepts in multi-modal perception, particularly when confronting the challenges associated with impaired or missing sense rehabilitation. The present systematic review investigates the role of haptics for the hearing impaired in training or gamified activities. We applied a set of keywords to the Scopus® and PubMed® databases, obtaining a collection of 35 manuscripts spanning 23 years. Each article has been categorized following a documented procedure and thoroughly analyzed. Our findings reveal a rising number of studies in this field in the last five years, mostly testing the effectiveness of the developed rehabilitative method (77.14%). Despite a wide variety in almost every category we analyzed, such as haptic devices, body location, and data collection, we report a constant difficulty in recruitment, reflected in the low number of hearing-impaired participants (mean of 8.31). This review found that in all six papers reporting statistically significant positive results, the vibrotactile device in use generated vibrations starting from a sound, suggesting that some perceptual aspects connected to sound are transmittable through touch. This fact provides evidence that haptics and vibrotactile devices could be viable solutions for hearing-impaired rehabilitation and training.

A.1 Introduction

Human perception is by nature multisensory, and an extensive amount of research has been dedicated to understanding how we perceive the world through the interaction of our senses, starting with the pioneering work presented in [68]. In Biocca et al. [10], the authors outline different ways in which the senses interact: for instance, cross-modal enhancement refers to the fact that stimuli from one sensory channel enhance or alter the perceptual interpretation of stimulation from another sensory channel. In certain situations individuals can lack or have reduced sensory modalities. This is the case of individuals with hearing impairment, visual impairment or tactile impairment. In these situations, technologies could help augment or substitute the missing modality.

In this paper we focus on individuals with hearing impairment and investigate how devices based on the sense of touch can help them to make better sense of different stimuli and in recent years, several devices have been proposed for this purpose. This approach is supported by the fact that hearing and touch present a higher temporal resolution if compared to vision, which is especially true when the sense of touch is experienced by the hands, which have a greater resolution than other body parts [43].

Since the 1920s, researchers conducted experiments with the goal of investigating the perception of vibrating objects through tactile sensitivity, and comparing their characteristics with hearing perception. These efforts also led to new research paths such as inquiring how deaf people experience sound through the sense of touch [42].

Recent research by Cieśla et al. [16] has shown that a speech-to-touch sensory substitution device significantly improves speech recognition in both cochlear implant (CI) users and individuals with normal hearing (NH). This finding aligns with the longstanding idea that the sense of touch can be effectively employed to substitute or enhance auditory experiences in such devices. One of the first experiments in this direction was the "hearing glove", a speech technology modeled on the cochlea but constrained by the limited sensitivity of human skin, presumably invented by Norbert Wiener in the 1940s [47]. Other prominent experiments were performed by Clark and colleagues who proposed the Tickle Talker [19], an eight channel electro-tactile speech processor. The Tickle Talker was used to reinforce residual hearing or to supplement lip reading; the device showed potential in rehabilitation of severe hearing-impaired children and adults. Since then, tactile feedback has been used for several applications aimed at aiding hearing-impaired individuals, such as music listening [52], and even tap dancing [63]. In the works considered by this review, vibrotactile devices have been used to enhance various dimensions of hearing, such as sound source localization [27], pitch discrimination [35, 64], and speech comprehension [25, 69]. Experiments have been proposed to improve non-auditory perceptual abilities, including environmental perception [55], voice tone control [59], as well as cognitive ones like braille perception [56], lip reading [9], or web browsing [39].

Most of the time, developers of games and video games most of the times do not take the needs of individuals with disabilities into account while creating their products [1]. Thus, accessible games have an important role to include a population that otherwise would be excluded [54]. In the last twenty years [32] a strong focus has been placed on creating accessible games for populations with different abilities. For individuals with severe hearing loss (HL) or those who may not benefit from traditional speech training, augmentative and alternative communication methods can support effective rehabilitation [22]. Among them, gamification principles have been demonstrated to be effective in strengthening children's learning performance and improving their training experience [12, 58]; this strategy has been widely applied in children's education and training products, bringing principles and mechanics from the gaming world to increase engagement and motivation of the user [15]. Therefore, a gamification approach to auditory-verbal training is also a promising direction for hearing rehabilitation [75], merging the fields of gamification and training to benefit individuals with hearing impairments.

We have briefly discussed the development of devices that enhance or replace acoustic signals, as well as the extensive use of tactile and vibrotactile feedback in games and video games over the past several decades [7]. These applications have roots dating back to the early days of gaming [74]. As a result of these developments, the intersection of rehabilitation and training techniques for the deaf, haptic and vibrotactile stimulation, and game dynamics emerges as a promising area of research that deserves further investigation.

A relatively recent review concludes that there is a lack of research in auditory or cognitive impairments compared with visual and motor disabilities, suggesting this as a topic for further research [1]. While devices that augment or substitute hearing using touch have been continuously developed, it is less known how training using such devices can help improve hearing skills. Systematic reviews have raised questions about the effectiveness of musical training [45] and investigated and individualized computer-based auditory training [34]. Some have underscored the influence of variables such as participants' age, training duration, and the type of hearing device used [65], while some enquired the use of tactile displays for music applications design for hearing impaired individuals [52]. Additionally, studies have explored the impact of gamification on the learning process [12], while others have focused on deaf students without incorporating the haptic aspect into the assessment [17]. Therefore, the evaluation of the impact of vibrotactile technology is a crucial consideration for providing assistance in training activities to individuals with hearing impairments.

In this paper, we present a systematic review of the literature regarding training and gamified experiences that use haptic feedback to help individuals with hearing impairments. Section A.2 introduces the (often ambiguous) terminology, Section A.3 addresses the research questions, Section A.4 the methodology, Section A.5 the results, Section A.6 the discussion, Section A.7 the limitations of this study and Section A.8 the conclusions.

A.2 Definitions

To establish a foundational understanding of this review and facilitate comprehension of the central concepts addressed in it, we will commence by providing relevant definitions. These definitions will serve as a framework for the subsequent analysis and discussion throughout the document.

Haptic In the Dictionary of Psychology, James M. Baldwin defined haptics as "[...] the concomitant sensations and perceptions [...] cover[ing] the whole range of function of skin, muscle, tendon, and even of the static sense—thus including the senses of temperature and pain, and the perceptions of position, movement, etc." [5].

Sensory augmentation Involves extending the individual perception of a sense by utilizing another sense or the same sense, and can involve various sensory systems [44].

Sensory substitution Is the replacement of a missing sensory perception by conveying the information typically acquired through one sense to another [4].

Tactile Is an umbrella term for the perception of vibrations, static pressure, skin stretch, or friction [14].

Vibrotactile Is a subcategory of tactile perception, where the tactile sensation is caused by an oscillating object [14].

A.3 Problem Statement and Research Questions

The problem tackled by this systematic review is to explore the methods we found that integrate haptic and vibrotactile stimulation into rehabilitation and aid in general for people with hearing problems. This research is motivated by a recognized gap in the existing literature, as discussed in the introduction. We want to update the corpus of existing devices and techniques (e.g., training) in this area and report their impact and effectiveness on the disabled population under study. Our systematic review focuses on the following research questions:

RQ1 What are the main methodological characteristics of the reviewed articles?

RQ2 What are the most common strategies for designing haptic-enhanced games or training programs to facilitate skill development, communication, or accessibility for individuals with varying degrees of hearing impairment?

RQ3 Are the studies successful in reproducing positive effects when haptic feedback is applied?

A.4 Methodology

In this section, we describe the methodology we employed to select relevant literature for the systematic review. The whole study has been conducted following the PRISMA 2020 guidelines, and the related checklist [51].

A.4.1 Keywords

We curated a set of keywords concatenated with the logical operator "AND", organized into three essential categories: tactile, hearing impairment, and games for rehabilitation and training. To cover various aspects of the core topics, we used the logical operator "OR" to connect alternative keywords. The keyword combination with the operators used for the database search is presented below (Listing A.1):

Listing A.1. Keywords combination used for the database search.

```
    haptic OR vibrotactile
    OR tactile OR touch
AND
    hearing-impaired OR deaf
    OR (hearing AND impaired)
AND
    game OR training OR education
    OR videogame OR gamification
```

The first group of keywords comprises terms that cover aspects of tactile interaction, such as *haptics* and *vibrotactile*. The term 'haptics' refers to the broad sense of touch, while 'vibrotactile' specifically entails the presence of a vibrating object that stimulates the tactile sensation. For a more in-depth explanation of these terms, we invite the reader to consult Section A.2. The second group of keywords focuses on the target group who are hearing impaired or deaf individuals. Lastly, the third group includes keywords related to both training and gamification, which are at the core of this research.

Thanks to this set of keywords we intended to cover the majority of terms that are commonly used in the research fields of tactile perception, hearing impairment, and game-based interventions.

A.4.2 Inclusion/Exclusion Criteria

Together with the keywords, we established specific inclusion criteria. The manuscripts accepted in our review had to be written in English, undergo peer review, present primary research, being published in the last 25 years, and be designed to address the specific needs of the hearing-impaired population.

During the analysis, we adopted different exclusion criteria codes to better track the process. Concerning the first iteration where we took into account only abstract, title, and keywords, we applied the codes that are reported in Table A.1 with the exception of the last two (missing validation or intervention), that have been used for the in-depth analysis.

Table A.1: Exclusion codes and data.

Description	No.	%	Note
Not Available	5	4.27%	Cannot find the manuscript
Not English	1	0.85%	
Not Game/Training/Edu	17	14.53%	
Not Hearing-Impaired	28	23.93%	
Not Last Publication	3	2.56%	Newer publications, same project
Not Primary Research	5	4.27%	e.g., review
Not Vibrotactile	26	22.22%	
Off Topic	32	27.35%	Multiple reasons (e.g., NV + NGTE + NHI)
No Intervention	1	0.85%	
No Validation	6	5.13%	
Total	117	100%	

A.4.3 Database Selection

To identify relevant literature, we conducted searches in two prominent electronic databases: Scopus® and PubMed®. We chose these databases due to their extensive and pertinent literature in the technical and medical domains. Since Scopus® includes more than 90 million records (Scopus blog, `https://blog.scopus.com/posts/scopus-now-includes-90-million-content-records`, accessed on 5 December 2023) and PubMed® more than 36 million (Pubmed about page, `https://pubmed.ncbi.nlm.nih.gov/about/`, accessed on 5 December 2023), we deemed incorporating additional data sources into this review unnecessary.

A.4.4 Data Collection

Using the aforementioned keywords and criteria, we retrieved a total of 187 entries from Scopus® and 180 from PubMed® databases (as of 26 September 2023). In the former database, one paper has been automatically removed by the Scopus® search engine due to lack of a peer review. The research results were stored in the references manager Zotero (`https://www.zotero.org/`, accessed on 5 December 2023); here, we merged the two collections and removed the duplicates, obtaining 294 unique records. Subsequently, we exported the results in a spreadsheet that allowed us to better organize the references and keep the relevant information only. We performed a second filtering operation by choosing only the manuscripts published after 1998 (i.e., within the last 25 years). For each of the 159 records obtained we analysed the title, the abstract, and the keywords, finally selecting only 42 relevant records. As a last step, we performed a comprehensive review of the full papers by narrowing down the eligible records for this review to 35 manuscripts. In Figure A.1 we report the diagram of the whole process for the data collection of this systematic review.

A.4.5 Coding and Analysis

To answer the research questions, we categorized the selected entries with the methodology illustrated in Figure A.1, analyzing the content of each manuscript. We reported the most relevant aspects of each research in a spreadsheet that contains, among other things, the following elements: study

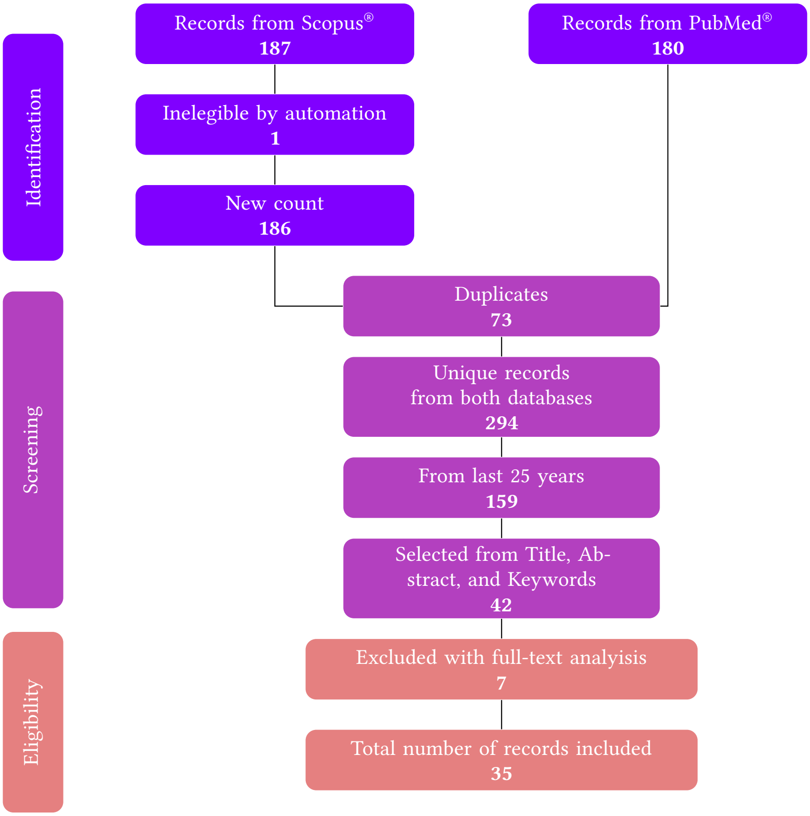


Fig. A.1: Block diagram of data inclusion method.

type, haptic body location, haptic usage, vibrotactile technology, mappings, vibrotactile processing, target impairment, training, and data collection method. This choice has been made to create an overview of the different applied methodologies and technologies, with the goal of answering the research questions.

A.4.6 Categorization

Here we introduce and explain some of the categories we applied. We divided the entries based on their research focus. The ones that are primarily aiming to demonstrate or validate the efficacy and outcomes of a particular approach or intervention are categorized as *effectiveness*. Others that focus on exploring and refining design methodologies and proving their validity fell into the *design* category. A second relevant category is the type of study. We differentiated the *experimental* from the *quasi-experimental* designs when we found that sample randomization, i.e., the random selection and assignment to a group of participants, was missing [41]. We indicated with *pre-post tests* the studies that analyzed the effect of an intervention measuring the subjects' performances before and after it. The *development and usability evaluation study* category has been created for less structured studies that mainly focused on the design and the functionality aspects rather than the effects of the treatment. Finally, *mixed methods* was the category chosen for the experiments where the design of the study featured different aspects of other studies.

Considerable attention has been paid to categorizing and describing the approaches for the haptic usage, the choices around mappings, and the processing techniques for the generation of vibrotactile stimuli (where present). In the literature, the use of haptic feedback mainly addresses the *substitution* or the *augmentation* of one or more senses (e.g., hearing).

Diverse mapping strategies include employing *full sound* for actuator feedback, generating vibrotactile stimuli through *text* or *gestures*, and utilizing *synthesis techniques* that diverge from traditional sound-based or input-related methods. Endless combinations can be chosen when referring to vibrotactile processing. The categories we used try to simplify the plethora of techniques, pointing at two main features that can be identified in most of them: *fundamental frequency (F0) extraction* and modulation of a carrier with a *temporal envelope*. For more complex choices, we invite the reader to refer directly to the related manuscripts, since it would have been impractical to put this information inside a table.

The manuscripts reviewed in this study employed various mappings in their research projects, which we categorized into three groups: *input-vibrotactile*, *input-location vibrotactile*, and *body location-output*. In the first category, an input source is recognized and mapped to a specific vibration output. For example, *sound-vibrotactile* mapping involves generating vibrations from a manipulated sound sample to achieve specific perceptual effects. Researchers also utilized other inputs such as visual input (e.g., associating a specific image with a vibration), gestural input (e.g., associating a specific movement with a vibration), and textual input (e.g., associating a specific word or group of words with a vibration). The second category involves the use of sound-vibration maps on a specific body area, where the information includes both the spatial position and the vibration itself. Lastly, the third category incorporates the use of body location as an input source (e.g., touch of a part of the hand) mapped to a text output (e.g., letter, phoneme).

A.5 Results

By applying the filters presented in Section A.4, we selected 35 out of 159 articles that were obtained with the screening process (22.01% rate of inclusion). Of the excluded items, 27.35% were marked as *out of topic* due to the lack of multiple key aspects for this research (i.e., more than one exclusion code applied). An additional 23.94% were not focusing on the hearing-impaired population, and 22.22% were missing haptic feedback. Other reasons for exclusion and the associated rates can be found in Table A.1. As a result, we present in Tables A.2–A.9 the papers that constitute this systematic review, highlighting key characteristics of each publication.

Table A.2: Summary of the selected articles (Part one, A).

Article	Year	Description	Study Type	Haptic Usage	Body Location	Mappings	Vibrotactile Processing	Target Group	Participants	Training
[35]	2023	Pitch discrimination study with training on amateur and professional musicians with normal or severely impaired hearing.	Pre-post test	Sensory substitution	Fingertip, forefoot	Sound—vibrotactile	Synthetic generation	Hearing impaired	19 participants, 15 NH, four hearing impaired	≤2 months
[20]	2023	Multisensory phonological and syntactic training	Pre-post test	Sensory augmentation	Wrist	Sound—vibrotactile	Not specified	Deaf	40 deaf and 28 hearing children	>2 months
[29]	2022	Design of a vibrotactile feedback device and test with melodic contour identification	Pre-post test	Sensory augmentation	Hand, fingertip	Sound—vibrotactile	Temporal envelope, full sound	Hearing impaired	15 NH participants	Pre-test
[37]	2022	Design of a vibrotactile feedback device for delivering customizable spatiotemporal tactile patterns	Pre-post test	Sensory substitution	Lower back	Text—vibrotactile	Synthetic generation	Sensory impairment	10 healthy participants	Pre-test
[76]	2022	Investigate algebra learning experience of university students with hypoacusis using tangible systems	Mixed methods	Sensory augmentation	Hand	None	None	Hearing impaired	One cochlear implanted, one NH	Pre-test

Table A.3: Summary of the selected articles (Part one, B).

Article	Year	Description	Study Type	Haptic Usage	Body Location	Mappings	Vibrotactile Processing	Target Group	Participants	Training
[21]	2021	Investigate whether temporal abilities can be enhanced using a novel Android app	Pre-post test	Sensory substitution	Hand	None	Synthetic generation	Sensory impairment	12 participants (no impairment specified)	$\leq$1 week
[70]	2021	Determine limits of underwater vibrotactile stimuli perception and measure training	Mixed methods	Sensory substitution	Full body	Sound—vibrotactile	Synthetic generation, full sound	Hearing impaired	five hearing impaired, 30 children, 15 with severe HL, 15 NH	None
[13]	2021	Design of a serious game for children with hearing impairment with physical and digital interfaces	Development and usability eval.	Sensory augmentation	Hand	Visual—vibrotactile	Synthetic generation	Hearing impaired	Seven children hearing impaired	Pre-test
[36]	2021	Design of a musical game to let the hearing impaired enjoy music playing	Development and usability eval.	Sensory substitution	Hand	Gesture—vibrotactile	Synthetic generation	Deaf, hearing impaired	Six deaf and hard of hearing	Pre-test

Table A.4: Summary of the selected articles (Part two, A).

Article	Year	Description	Study Type	Haptic Usage	Body Location	Mappings	Vibrotactile Processing	Target Group	Participants	Training
[69]	2020	Test a tactile phonemic sleeve for word recognition	Quasi-experimental	Sensory substitution	Forearm	Phoneme—vibrotactile	Complex	Hearing impaired	51 NH	≤1 month
[24]	2020	Assessing if electro-haptic stimulation substantially improves speech recognition in multi-talker noise when the speech and noise come from different locations	Experimental	Sensory augmentation	Wrist	Sound—vibrotactile	Temporal envelope	CI	Nine CI users, each of whom was implanted in only one ear	≤1 h
[64]	2020	Tactile glove that helps recognize pitch for hearing impaired individuals	Pre-post test	Sensory augmentation	Hand	Sound—location vibrotactile	Synthetic generation	Hearing impaired	Two CI users	≤1 month
[27]	2020	Improve haptic sound-localization accuracy using a varied stimulus set and assess whether accuracy improved with prolonged training	Experimental	Sensory augmentation	Wrist	Sound—location vibrotactile	Temporal envelope	Hearing impaired	32 adults with normal touch perception (16 experimental group, 16 control group)	≤1 month

Table A.5: Summary of the selected articles (Part two, B).

Article	Year	Description	Study Type	Haptic Usage	Body Location	Mappings	Vibrotactile Processing	Target Group	Participants	Training
[30]	2019	Tactile glove for speech-to-vibrotactile feedback	Development & usability eval.	Sensory substitution	Hand	Sound—location vibrotactile	Synthetic generation	Deaf-blind	Three NH	≤1 month
[16]	2019	Assessing if multisensory stimulation, pairing audition and a minimal-size touch device, improves intelligibility of speech in noise	Development & usability eval.	Sensory substitution	Fingertip	Sound—vibrotactile	Temporal envelope	Deaf, Hearing impaired	12 NH	≤1 h
[25]	2019	Vibrotactile feedback algorithm to improve speech-in-noise perception	Pre-post test	Sensory augmentation	Wrist	Sound—vibrotactile	Temporal envelope	Hearing impaired	10 CI users	≤2 weeks
[26]	2018	Tactile presentation of low-frequency sound information to improve speech-in-noise performance for CI users	Quasi-experimental	Sensory augmentation	Fingertip	Sound—vibrotactile	Temporal envelope	CI	Eight normal-hearing participants listened to CI simulated speech-in-noise	≤1 week
[31]	2017	EEG study on vibrotactile language discrimination in deaf and hearing individuals	Quasi-experimental	Sensory substitution	Fingertip	Sound—vibrotactile	Not specified	Deaf	14 deaf, 14 NH	≤1 month

Table A.6: Summary of the selected articles (Part three, A).

Article	Year	Description	Study Type	Haptic Usage	Body Location	Mappings	Vibrotactile Processing	Target Group	Participants	Training
[60]	2016	Design of an app for training of the Lorm-alphabet for facilitating communication between deaf-blind and sensory-abled individuals	Development and usability eval	Sensory augmentation	Fingertip	Location vibrotactile—text	Synthetic generation	Deaf-blind	Three NH	$\leq$1 h
[50]	2015	Design of a Morse code modulated haptics prototype for deaf-blind individuals to navigate web pages	Pilot study	Sensory substitution	Hand	Text—vibrotactile	Not specified	Deaf-blind	Four NH	$\leq$1 h
[53]	2014	Assessment of four signal processing methods in an app for environmental perception of sounds in deaf-blind people	Quasi-experimental	Sensory substitution	Ankle, Palm	Sound—vibrotactile	Temporal envelope	Deaf-blind	13 deaf, 5 deaf-blind	Pre-test
[55]	2013	Improve the ability of people with severe hearing impairment or deafblindness to detect, identify, and recognize the direction of sound-producing events	Field trial	Sensory substitution	Forearm, palm	Sound—vibrotactile	Temporal envelope	Hearing impaired, deaf-blind	Four with Usher syndrome I (deaf-blind)	Individual

Table A.7: Summary of the selected articles (Part three, B).

Article	Year	Description	Study Type	Haptic Usage	Body Location	Mappings	Vibrotactile Processing	Target Group	Participants	Training
[67]	2013	Intervention to teach three conceptually referenced tactile symbols for a child with multiple disabilities	Quasi-experimental	Sensory substitution	Hand	Shape/texture—word	None	Deaf-blind, intellectual disability	One deaf-blind	>1 month
[49]	2012	Vibrotactile chair to perform speech production training in deaf children	Experimental	Sensory augmentation	Full body	Sound—vibrotactile	Full sound	Deaf	Six deaf children; 20 deaf children	>2 months
[59]	2012	Investigate the effect of voice pitch training using a tactile feedback system	Quasi-experimental	Sensory substitution	Fingertip	Sound—vibrotactile	F0 extraction	Deaf, hearing impaired	Eight normal-hearing	None
[72]	2010	Vibrotactile feedback to improve speech production of Mandarin words	Experimental	Sensory augmentation	Fingertip	Sound—vibrotactile	F0 extraction	CI	12 cochlear implanted children	None
[38]	2010	Vibrotactile feedback to improve braille perception	Development and usability eval.	Sensory augmentation	Fingertip	Text—vibrotactile	Not specified	Deaf-blind, blind	Six deaf-blind, Three blind	Pre-test

Table A.8: Summary of the selected articles (Part four, A).

Article	Year	Description	Study Type	Haptic Usage	Body Location	Mappings	Vibrotactile Processing	Target Group	Participants	Training
[6]	2009	Real-time vibrotactile and visual feedback to train hearing impaired individuals	Experimental	Sensory augmentation	Fingertip	Sound—vibrotactile	F0 extraction	Deaf	53 hearing impaired	Pre-test
[40]	2006	Comparison of one-, tow- and seven- channel tactile aids for speech recognition in severely hearing impaired individuals	Quasi-experimental	Sensory substitution	Fingertip, wrist, neck, chest, abdominal skin	Sound—vibrotactile	Not specified	Hearing impaired	23 hearing impaired	Pre-test
[77]	2005	Design and evaluation of tactual display to reinforce lipreading	Experimental	Sensory augmentation	Fingertip	Sound—vibrotactile	Complex	Deaf	Four NH	Pre-test
[23]	2004	Design of a tactile pen and evaluation of tactons generation	Development and usability eval.	Sensory substitution	Hand	None	Synthetic generation	Deaf, blind	26 NH	Pre-test

Table A.9: Summary of the selected articles (Part four, B).

Article	Year	Description	Study Type	Haptic Usage	Body Location	Mappings	Vibrotactile Processing	Target Group	Participants	Training
[3]	2002	Verify that the deaf-blind people's tactile memory is better than that of sighted-hearing people through recognition and recall memory tasks and a matching pairs game	Quasi-experimental	Sensory substitution	Hand	None	None	Deaf-blind	10 deaf-blind and 10 sighted-hearing	Pre-test
[2]	2001	Investigate effects of tactile aids on visual lipreading task	Experimental	Sensory augmentation	Hand	Sound—vibrotactile	Temporal envelope	Hearing impaired	14 hearing impaired	Pre-test
[8]	2001	Investigate how speechreading is affected by hearing impairment and vibrotactile training	Experimental	Sensory substitution	Forearm	Visual—vibrotactile	Temporal envelope	Hearing impaired, NH	Eight NH; 8 hearing impaired	$\leq$2 months
[28]	2000	Investigate the potential value of tactile-alone training for hearing impaired	Experimental	Sensory substitution	Hand	Sound—electrotactile	Complex	Hearing impaired	Six NH	$\leq$1 week

In Figure A.2, we can observe the distribution of manuscripts over the years. An increase in publications concerning this review's topic is evident over the last four years, starting from 2019.

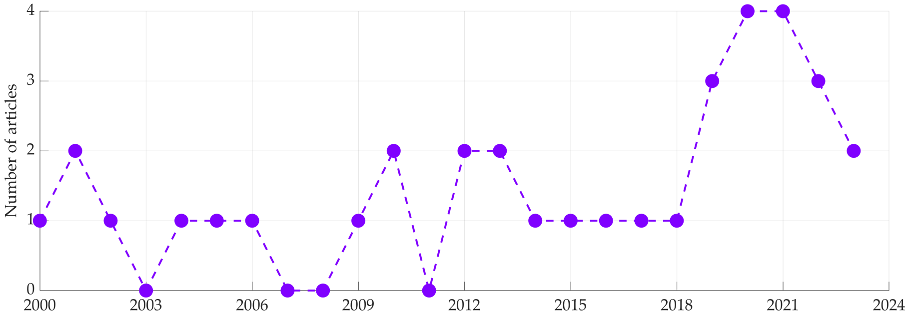


Fig. A.2: Amount of articles per year of publication.

In the following sections, we will present the data retrieved from the manuscripts and organized into charts and tables that categorize the main themes: Section A.5.1—metrics, Section A.5.2—methodologies, Section A.5.3—haptics, Section A.5.4—vibrotactile technologies, Section A.5.5—subjects, and Section A.5.6—outcomes. The consequent plots have been generated with MATLAB (version: 23.2.0 (R2023b), `https://www.mathworks.com/products/matlab.html`, accessed on 5 December 2023) using a combination of plot, scatter and bar functions.

A.5.1 Metrics

Here we display the metrics in terms of type of publication and amount of citations per article.

Publication Types

In Figure A.3 we report the type of publication of the included articles: the vast majority (68.57%) of them are journal articles, while only one is a book chapter. The remaining papers are conference proceedings.

Citations

Here, we present the citation count from Google Scholar along with the citations per year. The latter are calculated by dividing the total number of

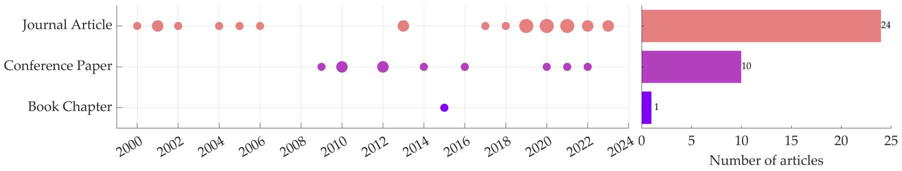


Fig. A.3: Publication type. The size of each data point represents the amount of articles per year.

citations by the number of years between the publication date and the current year. In Figure A.4, we also show the means for both categories: 18.43 for total citations and 3.01 for citations per year.

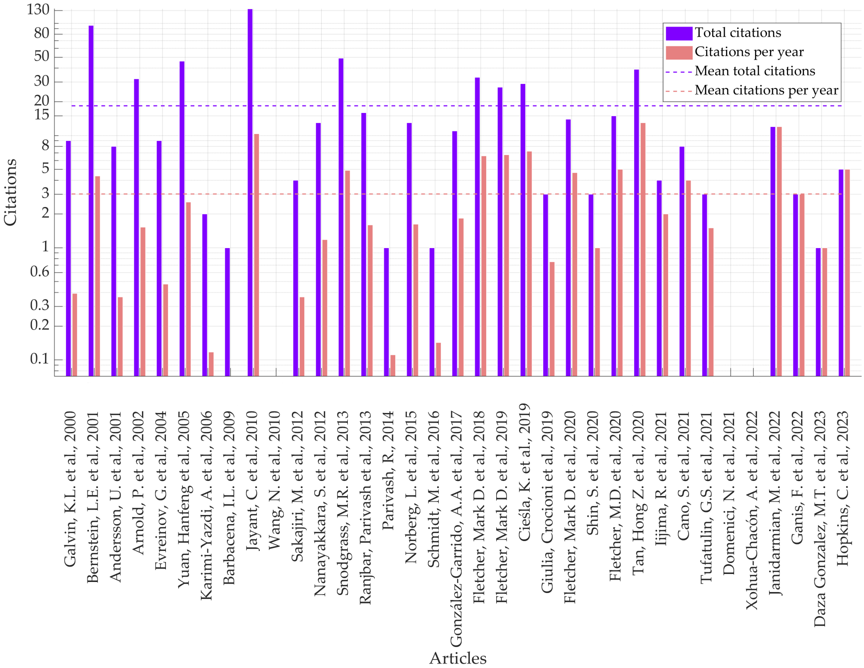


Fig. A.4: Citations per article [2, 3, 6, 8, 13, 16, 20, 21, 23–31, 35–38, 40, 49, 50, 53, 55, 59, 60, 64, 67, 69, 70, 72, 76, 77].

A.5.2 Methodologies

The first step of the analysis process included an investigation of the practices involved in the study. We considered the type of study design, the aim of the research, and the data collection procedure.

Study Type

We classified each article based on the study typology, as shown in Figure A.5. The majority of articles fall into two main categories: experimental (9, 25.71%) and quasi-experimental design (8, 22.86%). We employed this distinction to clearly identify studies randomizing the participants' groups (experimental) [61].

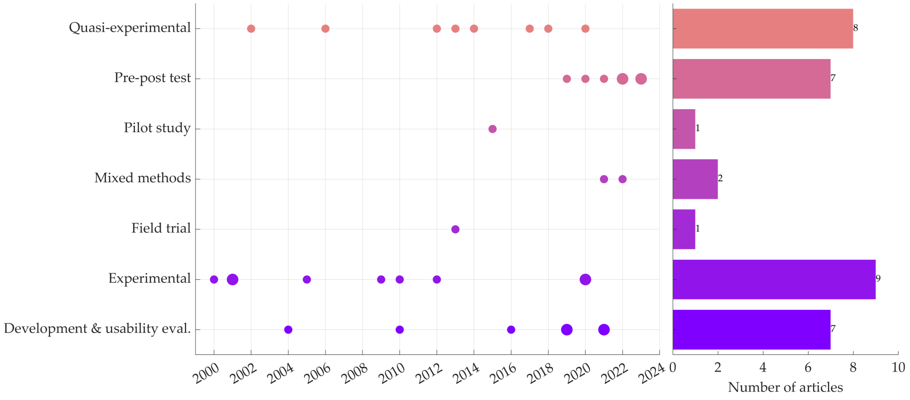


Fig. A.5: Study types. The size of each data point represents the amount of articles per year.

Two other frequently occurring study designs include the *pre-post test* and the *development and usability evaluation study*. The former examines the impact of a treatment by assessing performance before and after treatment administration [57]; based on our research criteria, we observe that this design has only been adopted during the last four years. The latter, as implied by its own name, is attributed to manuscripts whose aim is to design a process or device, and a test on a small group of participants is conducted.

We came across only one field trial, in which researchers aimed to enhance the ability of individuals with severe hearing impairment or deaf-blindness to detect, identify, and recognize the direction of sound-producing events [55]. Lastly, we encountered two mixed methods studies where both qualitative and quantitative evaluations have been made [70, 76].

Research Focus

In this section, we report the focus of each research included in this review. Despite the high variability in study design approaches and topics, we tried to summarize the principal goals in only two categories: *design* and *effectiveness*. The difference between the two groups is the main focus: in the first group, specific attention is payed to developing a solution, leaving the evaluation as a secondary aspect; in the second group, the core of the research is the assessment of a specific method/device in terms of its performances. We can see in Figure A.6 that the vast majority of the articles can be grouped in the *effectiveness* category (27 articles, 77.14%), and they can be found along the whole period of time that we took into account.

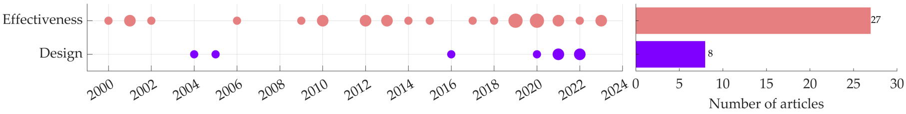


Fig. A.6: Research focus. The size of each data point represents the amount of articles per year.

Data Collection

Figure A.7 showcases a diverse range of data collection methods. Notably, *task accuracy* stands out as the most commonly employed method. In this category, we included all the manuscripts that contain accuracy measurements to evaluate user performance in specific tasks that are crucial for assessing the effectiveness of a treatment or a particular design. This prevalence of task accuracy as a method is not surprising, especially when compared to the findings in Figure A.6, which indicate that the majority of papers are exploring the effectiveness of novel solutions.

A.5.3 Haptics

The objective of this systematic review is to explore the utilization of haptic feedback in studies involving the hearing-impaired population and their training. To achieve this goal, it is essential to delve into various aspects of haptic feedback. In this section, we will examine the diverse roles of haptic feedback, investigate the specific body parts involved in this process, and provide an overview of the devices commonly used for this purpose

Usage

The breakdown in Figure A.8 reveals distinct patterns: 54.29% of the manuscripts have designed their studies to convey specific information through touch, completely bypassing other senses (*sensory substitution*). Conversely, approximately 45.71% use haptic feedback to enhance one or more senses falling in the category of *sensory augmentation*, as explained in Section A.2.

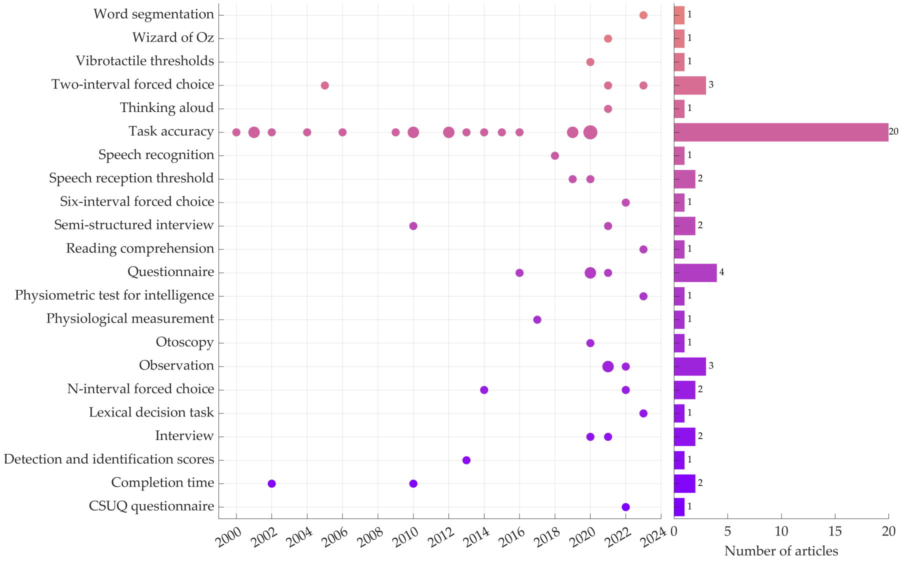


Fig. A.7: Type of data collection method. The size of each data point represents the amount of articles per year.

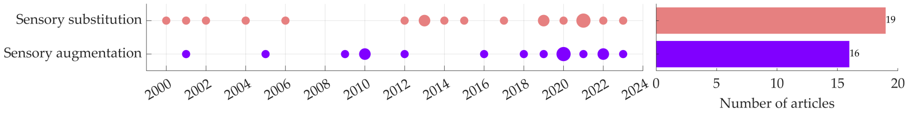


Fig. A.8: Use of haptics in the articles. The size of each data point represents the amount of articles per year.

Haptic Body Location

The human body presents different sensitivity to haptic stimuli depending on the body location involved. Therefore, we investigated the distribution on the body of stimulus application and presented the results in Figure A.9. A significant portion of the studies focused on stimulating either hands (13 studies) or fingertips (12 studies).

Mappings

A final aspect that has a great importance in the design of the experience with haptic feedback is the mapping, that is the way we connect a source stimulus with the haptic feedback. It is important to notice that haptic feedback can also play the role as an input, as in [67] where the shape/texture of a symbol was matched with a word. In Figure A.10, we can observe that 18 articles (51.42%) use the *sound-vibrotactile* mapping. The group *none* encompasses all the manuscripts that do not present a specific connection between a sensorial input and the vibrotactile feedback generated, but instead investigate a perceptual aspect related to haptic feedback.

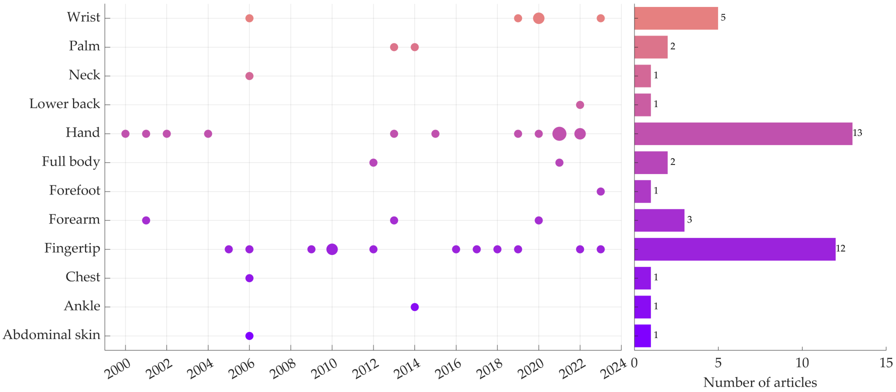


Fig. A.9: Body locations where haptic feedback has been applied. The size of each data point represents the amount of articles per year.

A.5.4 Vibrotactile Technology

The majority of the publications involved the use of some vibrotactile feedback technology. This can be provided by either a prototype conveying vibrations

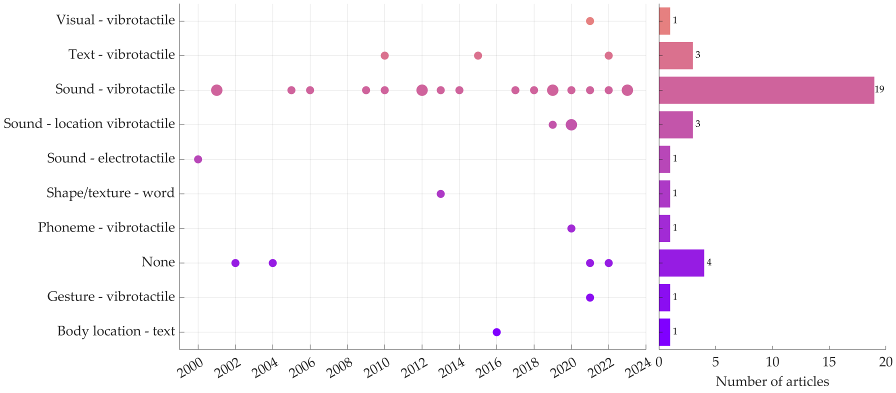


Fig. A.10: Haptic feedback mappings. The size of each data point represents the amount of articles per year.

or a commercially available device. In the following sections we are going to investigate which kind of solutions have been used.

Device

Figure A.11 displays the devices utilized in various articles. Smartphones are the most frequently used, with six publications employing their features to provide vibrotactile feedback. We can also observe that a great variety of solutions have been investigated, from measuring devices [24–27] to industrial products [35], and specifically designed devices for conveying vibrotactile feedback [2, 20, 29, 40, 72].

Actuators

Another important aspect of vibrotactile feedback is the actuator's technology. In Figure A.12, it is evident that the *eccentric rotating mass (ERM)* is the most commonly employed type of actuator. This aligns with our earlier discussion in Section A.5.4, where we discussed about using smartphones as tactile devices. Notably, ERMs are the most prevalent actuators found in smartphones due to their low cost and small dimensions. The studies that opted for some of the Tactaid devices have been tagged with *not specified*, since to the best of our knowledge it is not clear which technology operates behind these patented devices. The second most common type of actuator is the *electrodynamic*

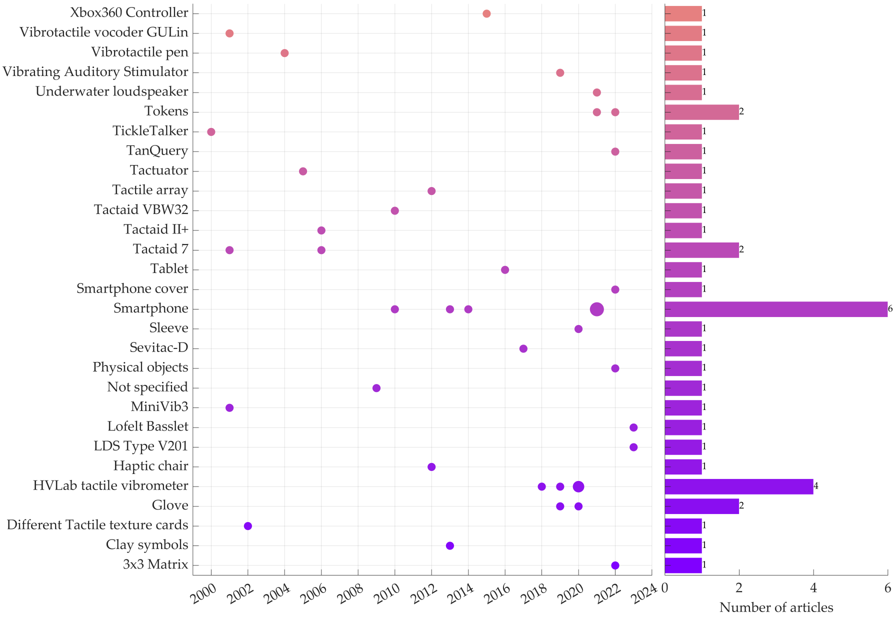


Fig. A.11: Vibrotactile devices. The size of each data point represents the amount of articles per year.

shaker, that is a high precision device for laboratory experiments and presents higher fidelity for greater cost and size compared to ERMs.

Vibrotactile Processing

The choice of a specific processing technique for generating vibrotactile stimuli is as crucial as selecting the target body part and the device. In Figure A.13, we can observe that 10 of the studies employed a temporal envelope to modulate a carrier signal, while an additional 10 generated bespoke signals without starting from a pre-existing sound or source; both such techniques have seen increased usage in the last decade. It is worth noting that five studies did not specify how vibrotactile stimuli were created. Three among those falling under the category *none* are primarily focused on haptic interactions [67, 76] or they measure the perception of a vibrotactile stimulus that is not associated with other sources [3]. Furthermore, three studies employed a unique and convoluted approach to derive vibrations from sound signals that did not fit any of the categories part of the figure. As a result, we categorized them as

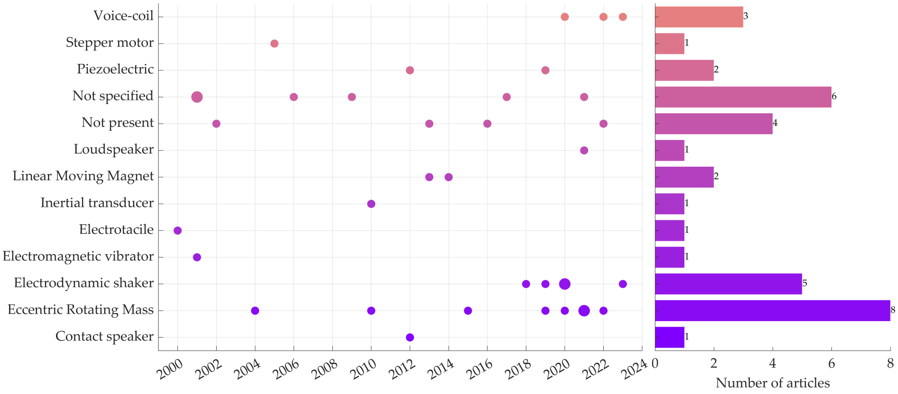


Fig. A.12: Actuator technology. The size of each data point represents the amount of articles per year.

complex [28, 69, 77].

A.5.5 Subjects

Upon examining the various participant groups in each study, we found that the mean number of total participants was 16.46 (STD = 15.67). By contrast, for studies including only sensory impaired participants, the mean value was 8.31 (STD = 12.13). Figure A.14 indicates the number of total and sensory-impaired participants in each article. It is evident that the sensory-impaired group shows less consistency compared to the non-impaired group across different experiments. Fourteen studies (40.00%) from this review are actually missing an impaired testing pool. This inconsistency can be attributed to the challenge of recruiting individuals with specific sensory impairments who are willing to participate in the tests. As a result, it is more common to simulate sensory impairments by depriving non-impaired individuals of a sense (e.g., using earplugs).

Target Impairment

The distribution of target groups is quite homogeneous, with the majority of the articles dedicating to the hearing impaired (17 studies) followed by the deaf (Nine studies) and the deaf-blind (Eight studies). Other categories included in Figure A.15 are associated with at least one of the above-mentioned groups.

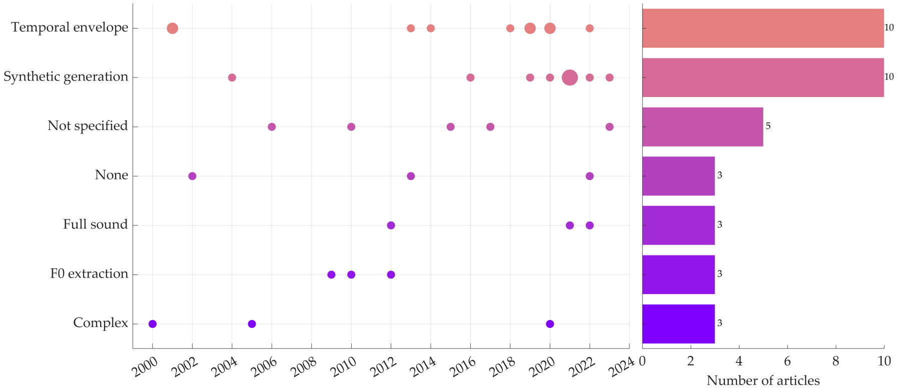


Fig. A.13: Vibrotactile processing generation techniques. The size of each data point represents the amount of articles per year.

Training

A final key point of our systematic review is the training aspect. The *pre-test* label reported in Figure A.16 indicates a short training experience conducted right before the test, with variable time, and often not specified. This condition has been reported by 13 articles (37.14%) while in four manuscripts we found an extended training experience that took at least one month [8, 20, 49, 67].

A.5.6 Outcome

In this section we examine the overall outcomes of the included articles and their statistical significance.

Positive/Negative

Figure A.17 depicts the results obtained in each study. None of the articles reported only negative effects of their treatments. On the contrary, it is quite surprising to see that 26 articles out of 35 (74.28%) obtained a positive result from their tests, and almost all of them have been published in the last 12 years. This fact might recall the effect of positive findings on the submission rate [18]. The category *complex* represents all the studies where more than one outcome has been found and not all of them were positive.

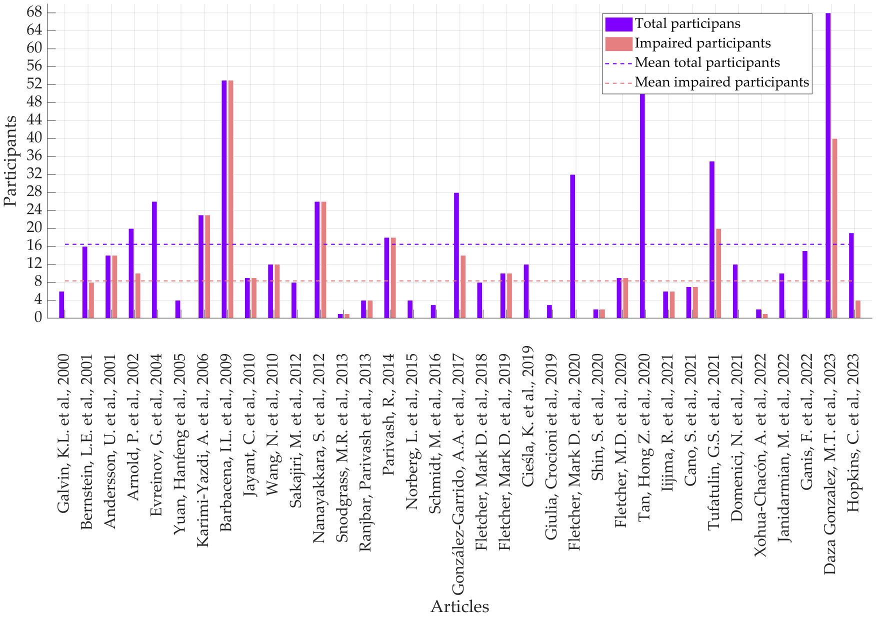


Fig. A.14: Sample size for each study. The two lines indicate the mean number of participants for each group [2, 3, 6, 8, 13, 16, 20, 21, 23–31, 35–38, 40, 49, 50, 53, 55, 59, 60, 64, 67, 69, 70, 72, 76, 77].

Statistical Significance

The outcomes obtained from each study could be statistically significant or not, and can be related to both a qualitative and quantitative measurement. In Figure A.18 we can see that in 19 (54.29%) articles there is an outcome that is statistically significant.

A.6 Discussion

The primary objective of this systematic review is to gather and analyze articles that propose haptic treatments or design solutions for the hearing-impaired population. In our methodology we detail the sampling approach, and in the subsequent chapter we evaluate 35 identified papers published between 2000 and 2023.

Two central themes underpin our exploration: training and gamification. To include all relevant literature where haptic technology intersects with gam-

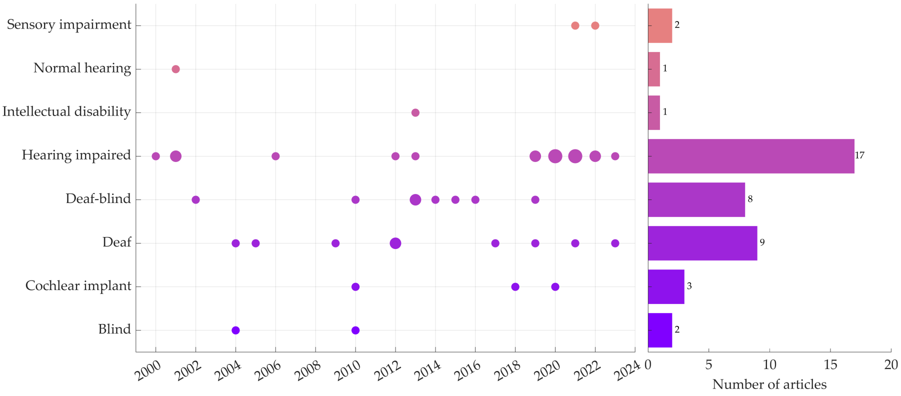


Fig. A.15: Target impairment. The size of each data point represents the amount of articles per year.

ification for hearing-impaired individuals, specific keywords such as game and gamification were introduced. The research in this particular domain yielded a limited number of papers. Despite the interest from industry and academia in video games equipped with vibrotactile feedback [66], the literature reporting their application to enhance the experience for the hearing-impaired population appears to be sparse. In our research, only three articles directly addressed gamification aspects in their design processes [13, 23, 36]. Cano et al. [13] focused on a table game for children aged 7 to 11 with hearing impairment. The game board and cards are the principal means of engagement. Additionally, a smartphone provides visual and vibrotactile feedback by reading QR codes on the physical interface. However, the latter is somewhat limited, offering a buzz-like sensation only when a child's answer is incorrect, given its secondary role since the smartphone screen simultaneously displays a corresponding sad face emoticon. The second paper [36] introduces a mobile app that offers vibrotactile feedback in response to a detected drumming gesture by the smartphone. The interaction and feedback are described clearly, but the study lacks emphasis on the gaming aspect, even if the keyword game is included. The final paper addressing gamification is authored by Evreinov et al. [23]. In this work, the authors showcase a pen which is able to provide vibrotactile feedback when connected to a pocket PC. The primary objective of the vibrotactile feedback in this context is to convey tactile icons (i.e., tactons [11]) to deaf or blind users during their interaction with two video games

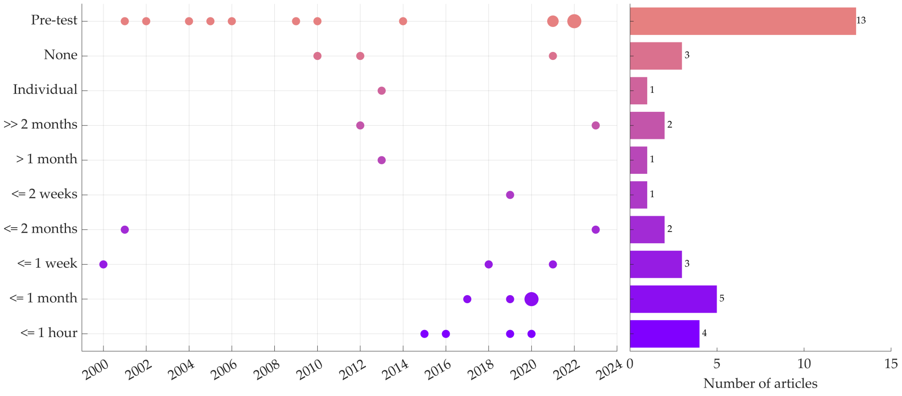


Fig. A.16: Users' training. The size of each data point represents the amount of articles per year.

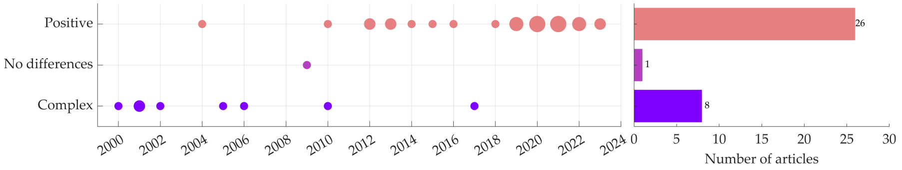


Fig. A.17: Articles outcomes. The size of each data point represents the amount of articles per year.

designed for this specific purpose. Bringing together these considerations, we noticed a gap in the literature regarding games and haptics for the hearing impaired, making this a valuable path to investigate in the future.

Figure A.2 reveals a growing interest in haptics applied to training for the hearing impaired. This can be paired with the increase in the number of new systems for music applications for the same target population [52]. The majority of papers focus on the effectiveness of the developed rehabilitative method, as stated in Section A.5.2. Seventeen out of 27 studies (62.96%) measure the user's accuracy on a specific task which is the most recurrent measurement, as reported in Figure A.7. The remaining ones rely on psychophysiological measurements (such as two-interval forced choice scores), speech comprehension evaluations, or qualitative observations. Conversely, all the studies that collected data regarding task accuracy have effectiveness as a research focus, except for three [23, 60, 64]. This finding can be read as a shared methodology construction; the experimental design of a rehabilitative

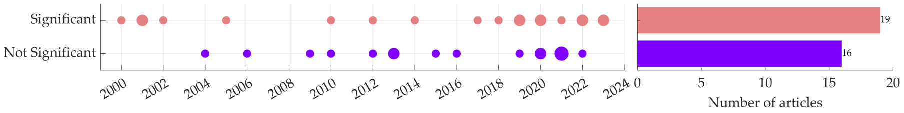


Fig. A.18: Statistical significance. The size of each data point represents the amount of articles per year.

or training method includes the definition of a task whose outcome serves as a measurable quantity that can be used as a metric for training effectiveness.

We relate the almost equal partition in Figure A.8 to some of the observed themes' main patterns. For the target population, we note that there are seven studies involving blind participants in which vibrotactile technology was used for sensory substitution; only two used it for sensory augmentation. It is reasonable to think that absence of sight drives this design choice. Conversely, all three studies involving users with CIs use haptics for sensory augmentation. CI users receive a new electrical stimulation to their auditory nerve that gives them a mode of perception; making it multisensory could be a way to acquaint them with hearing. We note that sensory substitution and augmentation have a symmetric distribution concerning the main trends in vibrotactile processing in Figure A.13. Among the studies using the augmentation approach, four used temporal envelope and seven synthetic generation; in the other group, six used temporal envelope, and only three generated vibrotactile stimulation synthetically. Even if distributed almost uniformly, sensory substitution leans toward creating the vibrotactile stimulus from scratch; sensory augmentation tends to use a temporal envelope perceivable by the other senses.

Determining the placement of vibrotactile stimuli is crucial for achieving the intended outcomes. The sensitivity distribution of our body to haptic stimuli is quite diverse, and considering these concepts is pivotal for good design. From Figure A.9 it can be seen that most devices deliver haptic feedback to hands, palms, and fingertips. This is because these areas are rich in mechanoreceptors such as Pacinian receptors and Meissner corpuscles, which are crucial for perceiving vibrotactile stimulation [71]. Notably, even in the early stages of human life, during infants' exploration, it has been demonstrated that we commonly rely on our hands and fingers to give sense to our surroundings. This tactile exploration allows humans to discern the objects' shapes, textures, and temperatures, even before having the ability to inquire them visually [73]. Another area of the body used in the selected manuscripts are the wrists. This is a more convenient area for conducting other activities while receiving haptic feedback, since our hands can be left free to perform other tasks. The drawbacks are the presence of body hair that affects sensitivity, and clothes that might interfere with the experience. It is worth mentioning that in the article by Tufatulin et al. [70], the researchers used a loudspeaker to convey a full-body haptic feedback experience through water, using it as a medium to provide a multisensory experience, combining sound and vibrations to improve children's hearing activation after hearing aid (HA) or cochlear implantation.

Since humans present limited tactile capabilities if compared to hearing ones (e.g., reduced frequency spectrum and resolution [46]), often digital signal processing (DSP) techniques are often applied to the input sound stimuli to extract specific features such as the fundamental frequency (F0), harmonics, and temporal envelope. This way, the vibrotactile stimulation can emphasize certain aspects of the sound input while omitting secondary ones, aligning with the research objectives and tactile capabilities. In Figure A.13, we can observe that one of the most common approaches involves extracting the temporal envelope from sound signals and applying it to the vibrotactile signal (that could, for instance, be generated using a synthesis method). If we focus on the articles that employed the *sound-vibrotactile* mapping, a remarkable pattern emerges: all of them administered vibrotactile stimuli to either the fingertip, the palm, or the whole hand, capitalizing on the high sensitivity of these body parts to vibrations [71]. Furthermore, stimulating the hand or fingertip requires minimal preparation from the participants, often eliminating the need for additional garments or wearable equipment that might increase the task duration and discomfort. Out of the 19 studies applying the *sound-vibrotactile* mapping, nine present positive statistically significant results; and additional six show more complex results with negative and positive outcomes [2, 8, 31, 40, 72, 77]. These outcomes are tightly linked with both the design choices and the characteristics of the participants. Upon examining individual experiments, a common trend emerges in eight of the 14 studies that reported a positive or statistically significant result: the temporal envelope processing technique. This technique involves extracting the amplitude of sound stimuli over time and applying it to the vibrotactile signal, aiming at a clear amplitude correlation between the two. Summarizing these findings, one could argue that employing *sound-vibrotactile* mapping with temporal envelope processing techniques and delivering this stimulus to the hand (or fingertip or palm) may result in positive and statistically significant outcomes.

Moving to the device choice, we can observe that almost every article adopts a unique approach. The most common device is the smartphone [21, 36, 38, 51, 53], given its near-ubiquity; with most people owning one or at least being familiar with it, smartphones serve as convenient and portable tools for training and enhancing experiences. However, the compact size of smartphones comes with some drawbacks, particularly concerning vibrotactile performance. Due to their small form factor, the actuators in these devices must also be small, resulting in reduced frequency performance. Additionally, the design focus for the vibrotactile experience on smartphones has consistently prioritized conveying simple messages or notifications rather than complex sounds. To reduce costs and keep them as compact as possible, the majority of smartphones are equipped with ERM actuators that usually operate on one single frequency (resonant frequency) [33]. Furthermore, using such devices in this field introduces significant challenges in controlling potentially confounding variables that are typically less pronounced in controlled laboratory settings and equipment, hence complicating and reducing the reliability of experiments and evaluations. A contrasting approach is evident in the studies by Fletcher et al. [24–27], where electrodynamic shakers are employed to convey vibrations through a complex and high-fidelity piece of equipment. Specifically, electrodynamic shakers are closely linked to voice-coils and find extensive use in industrial applications. Using this method, the HVLab device reproduces the input signal with good quality, covering a frequency range of 16 to 500 Hz with a low tolerance for frequency deviation (¡0.1%). Given the variety of tools and devices available, researchers should exercise caution when choosing an actuator technology, bearing in mind that each has its pros and cons. Broadly, two major categories can be distinguished: piezoelectric, ERM, and linear moving magnets favor small size and low cost, whereas electromagnetic vibrators, voice-coils, loudspeakers, and inertial transducers emphasize high-quality performance.

Our research has unveiled haptic solutions that have evolved over the years, often utilizing unique vibrotactile processing techniques tailored to specific devices, as shown in Sections A.5.4 and A.5.4. The lack of documentation on both hardware and software for the patented solutions generated issues concerning transparency and replicability. The lack of standardization of processing (Section A.5.4) and device technology raises concerns about the generalizability of the findings to broader user populations. This issue becomes even more evident when considering the target population (Section A.5.5).

The retrieved data reveals a significant disparity in the participants involved in these experiments: most studies either include a limited number of individuals with target impairment, or simulate impairments by depriving people of one or more senses. Thirteen publications present more than eight participants with impairments (above the mean of the whole study group; see Figure A.14), and ten of these studies declared an affiliation with a hospital or collaboration with a school, health institution, or association for impaired individuals [2, 3, 6, 8, 20, 24, 25, 31, 40, 49, 53, 70]. While recognizing the substantial challenges in the recruitment process, particularly within minority groups, we recommend that researchers establish close collaborations with hospitals, schools, and care centers to access a more diverse and representative population. Working closely within a clinical environment can also shed light on challenges that might not be apparent to academics alone. This collaborative approach can foster a better understanding of the real-world needs and experiences of the hearing-impaired population, ultimately leading to more effective haptic solutions.

A consistent pattern emerges when filtering the included articles to focus on those with positive statistically significant outcomes involving impaired individuals. All six studies meeting these criteria have been published within the last eleven years and employed sound-to-vibrotactile feedback mapping. If we dig into the details, four of the six articles applied haptic technology to enhance another sensory modality by applying vibrotactile stimulation on the wrist [20, 24, 25] or full body, as observed by Nanayakkara et al. [49]. The remaining two studies used vibrations in other body parts for sensory substitution [35, 53]. In three of them, the vibrotactile processing techniques utilized temporal envelope-based methods [24, 25, 53], while the other three applied full sound [49] generated synthetic stimuli [35]. Since Daza Gonzalez et al. [20] utilized the Lofelt bracelet, the specifics of the DSP method for the vibrotactile generation were not disclosed.

Considering the studies with either no training or only a brief training experience before exposing the participants to the experiment (pre-test), we observed no relevant pattern relating the training length and the statistical significance of the results. Nine studies reported no significant results, whereas seven studies did.

In conclusion, the evidence that all the significant positive outcomes involved a sound-to-vibrotactile feedback mapping confirms the long-standing idea that the multiple perceptual aspects connected to sound are transmittable through touch. Thus, a sensory substitution of this type is a viable solution for hearing-impaired rehabilitation and training.

A.7 Limitations

For this systematic review, we exclusively used two databases: Scopus® and PubMed®. We did not employ alternative methods for the literature search, such as secondary references or websites, as we believed that these two databases comprehensively covered the available literature. However, it's worth noting that we may have missed some *grey* literature.

Given the wide range of topics covered in the selected manuscripts, we acknowledge that justifying the inclusion of some articles, even if they met the selection criteria, presented challenges. For example, we are aware that some literature primarily aims to measure perception thresholds rather than to assess the effectiveness of haptic treatments or designs. Another hurdle was comparing studies involving haptics with those focusing on vibrotactile feedback. Some employed categories may not perfectly align with studies that do not exclusively involve vibrotactile feedback. Moreover, we reviewed studies where haptics was used to train individuals with sensory abilities to communicate with those who have impairments, presenting a different perspective from the majority of the included studies. Despite this difference, we chose not to exclude these articles because they offered insights into valuable aspects relevant to all the manuscripts.

A.8 Conclusions and Future Research

This systematic review compiles a set of papers exploring the integration of haptic feedback in training and gamification protocols to enhance the auditory experience for individuals with hearing impairments. We initially identified 294 articles from two prominent databases using relevant keywords. After careful screening and eligibility checks, we included 35 manuscripts in our analysis. Our examination primarily centers on study design, hardware and software solutions, training protocols, and the resulting test outcomes. Finally, we derive insights from the findings to provide recommendations for future researchers and designers.

Within the literature review, we observed a notable scarcity of studies addressing games and haptics for hearing-impaired individuals, underlining the urgency for further exploration in this critical area. Furthermore, those that delved into the topic often had a limited focus, either on vibrotactile or gamification aspects, leaving the combination relatively unexplored.

A noteworthy discovery is a consensus on targeting hands and wrists with haptic feedback alongside temporal envelope-processed sound, yielding positive and statistically significant results. This presents a promising avenue for future research. On the contrary, the diverse array of devices conveying vibrotactile feedback adds complexity, making it challenging to establish clear correlations between treatment administration and observed outcomes.

We emphasize the importance of conducting research in real-world, ecologically valid environments, collaborating closely with end-users, rather than confining studies solely to controlled laboratory settings. While acknowledging the challenges of field research, we contend that testing in real-world scenarios offers a more accurate understanding of the practical challenges and benefits experienced by hearing-impaired individuals with haptic solutions. Inspired by the diversity of design choices for training programs (see Section A.5.5), we believe that combining qualitative assessments with quantitative data can provide a more comprehensive understanding of this multifaceted sensory domain and richer interpretation of the results.

Referring to the results in Section A.5.6, it is crucial to note that several papers were excluded from our analysis due to a lack of statistically significant quantitative findings, attributed to a low participant count. This challenge can be addressed by designing studies involving organizations and hospitals, thereby ensuring a more extensive population to collaborate with and emphasizing the importance of qualitative results alongside quantitative ones.

As a final remark for future research, we recommend exploring more engaging technologies tailored to the younger population. While researchers and industries have developed immersive technologies over the past decade, it is noteworthy that previous studies emphasize the importance of clinical environments [62]. However, there is a limited inclusion of haptic feedback in immersive technology specifically designed for hearing-impaired individuals, with only a few examples found in the literature [48]. Therefore, we propose further investigation into the potential benefits of immersive experiences coupled with haptic feedback for this demographic.

References

[1] J. Aguado-Delgado, J.-M. Gutiérrez-Martínez, J. R. Hilera, L. de-Marcos, and S. Otón, “Accessibility in video games: A systematic review,” *Universal Access in the Information Society*, vol. 19, no. 1, pp. 169–193, Mar. 2020.

[2] U. Andersson, B. Lyxell, J. Rönnberg, and K. E. Spens, “Effects of tactile training on visual speechreading: Performance changes related to individual differences in cognitive skills.” *Journal of deaf studies and deaf education*, vol. 6, no. 2, pp. 116–129, 2001.

[3] P. Arnold and K. Heiron, “Tactile memory of deaf-blind adults on four tasks,” *Scandinavian Journal of Psychology*, vol. 43, no. 1, pp. 73–79, 2002.

[4] P. Bach-y-Rita and S. W. Kercel, “Sensory substitution and the human–machine interface,” *Trends in cognitive sciences*, vol. 7, no. 12, pp. 541–546, 2003.

[5] J. M. Baldwin, “Dictionary of philosophy and psychology: Vol. 1 (1901), vol. 2 (1902), vol. 3 (1905),” 1901.

[6] I. Barbacena, R. Freire, A. Barros, B. Neto, E. Carvalho, and E. De Macedo, “Voice codification evaluation based on a real-time training system with tactile feedback applied to deaf people,” 2009, pp. 697–700.

[7] A. Bayousuf, H. S. Al-Khalifa, and A. Al-Salman, “Haptics-based systems characteristics, classification, and applications,” in *Advanced Methodologies and Technologies in Artificial Intelligence, Computer Simulation, and Human-Computer Interaction.* IGI Global, 2019, pp. 778–794.

[8] L. Bernstein, E. Auer Jr., and P. Tucker, “Enhanced Speechreading in Deaf Adults: Can Short-Term Training/Practice Close the Gap for Hearing Adults?” *Journal of Speech, Language, and Hearing Research*, vol. 44, no. 1, pp. 5–18, 2001.

[9] L. E. Bernstein, M. E. Demorest, D. C. Coulter, and M. P. O’Connell, “Lipreading sentences with vibrotactile vocoders: Performance of normal-

hearing and hearing-impaired subjects," *The Journal of the Acoustical Society of America*, vol. 90, no. 6, pp. 2971–2984, 1991.

[10] F. Biocca, J. Kim, and Y. Choi, "Visual touch in virtual environments: An exploratory study of presence, multimodal interfaces, and cross-modal sensory illusions," *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 3, pp. 247–265, 2001.

[11] S. A. Brewster and L. M. Brown, "Tactons: Structured tactile messages for non-visual information display," 2004.

[12] R. I. Busarello, V. R. Ulbricht, L. M. Fadel, and A. V. de Freitas e Lopes, "Gamification approaches to learning and knowledge development: A theorical review," in *New Advances in Information Systems and Technologies.* Springer, 2016, pp. 1107–1116.

[13] S. Cano, J. S. Naranjo, C. Henao, C. Rusu, and S. Albiol-Pérez, "Serious Game as Support for the Development of Computational Thinking for Children with Hearing Impairment," *Applied Sciences*, vol. 11, no. 1, p. 115, 2021.

[14] S. Choi and K. J. Kuchenbecker, "Vibrotactile display: Perception, technology, and applications," *Proceedings of the IEEE*, vol. 101, no. 9, pp. 2093–2104, 2012.

[15] A. Christopoulos and S. Mystakidis, "Gamification in Education," *Encyclopedia*, vol. 3, no. 4, pp. 1223–1243, Oct. 2023.

[16] 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, no. 2, pp. 155–166, 2019.

[17] C. Costa, L. Marcelino, J. Neves, and C. Sousa, "Games for education of deaf students: A systematic literature," in *ECGBL 2019 13th European Conference on Game-Based Learning.* Academic Conferences and publishing limited, 2019, p. 170.

[18] A. Coursol and E. E. Wagner, "Effect of positive findings on submission and acceptance rates: A note on meta-analysis bias." 1986.

[19] RS. Cowan, KL. Galvin, JZ. Sarant, R. Millard, PJ. Blamey, and GM. Clark, "Improved electrotactile speech processor: Tickle Talker." *The Annals of Otology, Rhinology & laryngology. Supplement*, vol. 166, pp. 454–456, 1995.

[20] M. Daza Gonzalez, J. Phillips-Silver, N. Maurno, L. García, and P. Ruiz-Castañeda, "Improving phonological skills and reading comprehension

in deaf children: A new multisensory approach," *Scientific Studies of Reading*, vol. 27, no. 2, pp. 119–135, 2023.

[21] N. Domenici, A. Inuggi, A. Tonelli, and M. Gori, "A novel Android app to evaluate and enhance auditory and tactile temporal thresholds()." *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2021, pp. 5885–5888, Nov. 2021.

[22] Y. Elsahar, S. Hu, K. Bouazza-Marouf, D. Kerr, and A. Mansor, "Augmentative and alternative communication (AAC) advances: A review of configurations for individuals with a speech disability," *Sensors*, vol. 19, no. 8, p. 1911, 2019.

[23] G. Evreinov, T. Evreinova, and R. Raisamo, "Mobile games for training tactile perception," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3166, pp. 468–475, 2004.

[24] M. D. Fletcher, R. O. Cunningham, and S. R. Mills, "Electro-Haptic Enhancement of Spatial Hearing in Cochlear Implant Users," *Scientific Reports*, vol. 10, no. 1, p. 1621, Dec. 2020.

[25] M. D. Fletcher, A. Hadeedi, T. Goehring, and S. R. Mills, "Electro-haptic enhancement of speech-in-noise performance in cochlear implant users." *Scientific reports*, vol. 9, no. 1, p. 11428, Aug. 2019.

[26] M. D. Fletcher, S. R. Mills, and T. Goehring, "Vibro-Tactile Enhancement of Speech Intelligibility in Multi-talker Noise for Simulated Cochlear Implant Listening." *Trends in hearing*, vol. 22, p. 2331216518797838, Dec. 2018.

[27] M. D. Fletcher and J. Zgheib, "Haptic sound-localisation for use in cochlear implant and hearing-aid users." *Scientific reports*, vol. 10, no. 1, p. 14171, Aug. 2020.

[28] K. L. Galvin, P. J. Blamey, R. S. C. Cowan, M. Oerlemans, and G. M. Clark, "Generalization of tactile perceptual skills to new context following tactile-alone word recognition training with the Tickle Talker™," *The Journal of the Acoustical Society of America*, vol. 108, no. 6, pp. 2969–2979, Dec. 2000.

[29] F. Ganis, M. Vatti, and S. Serafin, "Tickle Tuner - Haptic Smartphone Cover for Cochlear Implant Users' Musical Training," in *Haptic and Audio Interaction Design*, C. Saitis, I. Farkhatdinov, and S. Papetti, Eds. Cham:

Springer International Publishing, 2022, vol. 13417, pp. 14–24.

[30] C. Giulia, D. V. Chiara, and H. Esmailbeigi, "GLOS: GLOve for Speech Recognition," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).* Berlin, Germany: IEEE, Jul. 2019, pp. 3319–3322.

[31] A. González-Garrido, V. Ruiz-Stovel, F. Gómez-Velázquez, H. Vélez-Pérez, R. Romo-Vázquez, R. Salido-Ruiz, A. Espinoza-Valdez, and L. Campos, "Vibrotactile discrimination training affects brain connectivity in profoundly deaf individuals," *Frontiers in Human Neuroscience*, vol. 11, 2017.

[32] D. Grammenos, A. Savidis, and C. Stephanidis, "Designing universally accessible games," *Computers in Entertainment (CIE)*, vol. 7, no. 1, pp. 1–29, 2009.

[33] E. Granado, F. Quizhpi, J. Zambrano, and W. Colmenares, "Remote experimentation using a smartphone application with haptic feedback," in *2016 IEEE Global Engineering Education Conference (EDUCON).* IEEE, 2016, pp. 240–247.

[34] H. Henshaw and M. A. Ferguson, "Efficacy of individual computer-based auditory training for people with hearing loss: A systematic review of the evidence," *PloS one*, vol. 8, no. 5, p. e62836, 2013.

[35] C. Hopkins, S. Maté-Cid, R. Fulford, G. Seiffert, and J. Ginsborg, "Perception and learning of relative pitch by musicians using the vibrotactile mode," *Musicae Scientiae*, vol. 27, no. 1, pp. 3–26, 2023.

[36] R. Iijima, A. Shitara, S. Sarcar, and Y. Ochiai, "Smartphone Drum: Gesture-based Digital Musical Instruments Application for Deaf and Hard of Hearing People," in *Proceedings of the 2021 ACM Symposium on Spatial User Interaction*, 2021.

[37] M. Janidarmian, A. Roshan Fekr, K. Radecka, and Z. Zilic, "Wearable Vibrotactile System as an Assistive Technology Solution," *Mobile Networks and Applications*, vol. 27, no. 2, pp. 709–717, 2022.

[38] C. Jayant, C. Acuario, W. Johnson, J. Hollier, and R. Ladner, "V-Braille: Haptic Braille perception using a touch-screen and vibration on mobile phones," 2010, pp. 295–296.

[39] N. Kaklanis, D. Tzovaras, and K. Moustakas, "Haptic navigation in the world wide web," in *Universal Access in Human-Computer Interaction. Applications and Services: 5th International Conference, UAHCI 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009.*

Proceedings, Part III 5. Springer, 2009, pp. 707–715.

[40] A. Karimi-Yazdi, A.-A. Sazgar, A. Nadimi-Tehran, A. Faramarzi, F. Nassaj, and S. Yahyavi, "Application and usage of tactile aid in Iran," *Archives of Iranian Medicine*, vol. 9, no. 4, pp. 344–347, 2006.

[41] R. E. Kirk, "Experimental design," *Sage handbook of quantitative methods in psychology*, pp. 23–45, 2009.

[42] V. O. Knudsen, ""Hearing" with the sense of touch," *The Journal of General Psychology*, vol. 1, no. 2, pp. 320–352, 1928.

[43] S. J. Lederman and R. L. Klatzky, "Haptic perception: A tutorial," *Attention, Perception, & Psychophysics*, vol. 71, no. 7, pp. 1439–1459, 2009.

[44] F. Macpherson, "Sensory substitution and augmentation: An introduction," in *Sensory Substitution and Augmentation*, F. Macpherson, Ed., 2018.

[45] C. M. McKay, "No evidence that music training benefits speech perception in hearing-impaired listeners: A systematic review," *Trends in hearing*, vol. 25, p. 2331216520985678, 2021.

[46] S. Merchel and M. E. Altinsoy, "Psychophysical comparison of the auditory and tactile perception: A survey," *Journal on Multimodal User Interfaces*, vol. 14, pp. 271–283, 2020.

[47] M. Mills, "On disability and cybernetics: Helen Keller, Norbert Wiener, and the hearing glove," *Differences-a Journal of Feminist Cultural Studies*, vol. 22, no. 2-3, pp. 74–111, 2011.

[48] M. Mirzaei, P. Kan, 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, no. 5, pp. 2084–2093, May 2020.

[49] S. Nanayakkara, L. Wyse, and E. A. Taylor, "The haptic chair as a speech training aid for the deaf," in *Proceedings of the 24th Australian Computer-Human Interaction Conference*, ser. OzCHI '12. New York, NY, USA: Association for Computing Machinery, 2012, pp. 405–410.

[50] L. Norberg, T. Westin, P. Mozelius, and M. Wiklund, "Web accessibility by morse code modulated haptics for deaf-blind," in *Technology, Rehabilitation and Empowerment of People with Special Needs*, 2015, pp. 123–134.

[51] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan *et al.*, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *International journal of surgery*, vol. 88, p. 105906, 2021.

[52] R. Paisa, N. C. Nilsson, and S. Serafin, "Tactile displays for auditory augmentation–A scoping review and reflections on music applications for hearing impaired users," *Frontiers in Computer Science*, vol. 5, p. 1085539, Apr. 2023.

[53] R. Parivash, "Signal processing methods for improvement of environmental perception of persons with deafblindness," vol. 902, 2014, pp. 398–404.

[54] J. R. Porter and J. A. Kientz, "An empirical study of issues and barriers to mainstream video game accessibility," in *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, 2013, pp. 1–8.

[55] P. Ranjbar and I. Stenström, "Monitor, a vibrotactile aid for environmental perception: A field evaluation by four people with severe hearing and vision impairment." *TheScientificWorldJournal*, vol. 2013, p. 206734, 2013.

[56] J. Rantala, R. Raisamo, J. Lylykangas, V. Surakka, J. Raisamo, K. Salminen, T. Pakkanen, and A. Hippula, "Methods for presenting braille characters on a mobile device with a touchscreen and tactile feedback," *IEEE Transactions on Haptics*, vol. 2, no. 1, pp. 28–39, 2009.

[57] E. A. Robinson and H. J. Doueck, "Implications of the pre/post/then design for evaluating social group work," *Research on social work practice*, vol. 4, no. 2, pp. 224–239, 1994.

[58] J. M. Rodríguez-Ferrer, A. Manzano-León, J. M. Aguilar-Parra, and A. Cangas, "Effectiveness of gamification and game-based learning in Spanish adolescents with dyslexia: A longitudinal quasi-experimental research," *Research in Developmental Disabilities*, vol. 141, p. 104603, Oct. 2023.

[59] M. Sakajiri, K. Nakamura, S. Fukushima, S. Miyoshi, and T. Ifukube, "Effect of voice pitch control training using a two-dimensional tactile feedback display system," 2012, pp. 2943–2947.

[60] M. Schmidt, C. Bank, and G. Weber, "Zoning-based gesture recognition to enable a mobile lorm trainer," vol. 9759, 2016, pp. 479–486.

[61] H. J. Seltman, *Experimental Design and Analysis.* Carnegie Mellon University Pittsburgh, 2012.

[62] S. Serafin, A. Adjorlu, and L. M. Percy-Smith, "A review of virtual reality for individuals with hearing impairments," *Multimodal Technologies and Interaction*, vol. 7, no. 4, p. 36, 2023.

[63] M. Shibasaki, Y. Kamiyama, and K. Minamizawa, "Designing a haptic

feedback system for hearing-impaired to experience tap dance," in *Adjunct Proceedings of the 29th Annual ACM Symposium on User Interface Software and Technology*, 2016, pp. 97–99.

[64] S. Shin, C. Oh, and H. Shin, "Tactile Tone System: A Wearable Device to Assist Accuracy of Vocal Pitch in Cochlear Implant Users," 2020.

[65] 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, no. 1, pp. 15–28, Feb. 2021.

[66] T. Singhal and O. Schneider, "Juicy haptic design: Vibrotactile embellishments can improve player experience in games," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–11.

[67] M. Snodgrass, J. Stoner, and M. Angell, "Teaching conceptually referenced core vocabulary for initial augmentative and alternative communication," *AAC: Augmentative and Alternative Communication*, vol. 29, no. 4, pp. 322–333, 2013.

[68] B. E. Stein and M. A. Meredith, *The Merging of the Senses.* MIT press, 1993.

[69] H. Z. Tan, C. M. Reed, Y. Jiao, Z. D. Perez, E. C. Wilson, J. Jung, J. S. Martinez, and F. M. Severgnini, "Acquisition of 500 English Words through a TActile Phonemic Sleeve (TAPS)." *IEEE transactions on haptics*, vol. 13, no. 4, pp. 745–760, Dec. 2020.

[70] G. Tufatulin, I. Koroleva, S. Artyushkin, and Y. Yanov, "The benefits of underwater vibrostimulation in the rehabilitation of children with impaired hearing," *International Journal of Pediatric Otorhinolaryngology*, vol. 149, 2021.

[71] R. T. Verrillo and S. J. Bolanowski, "Tactile Responses to Vibration," in *Handbook of Signal Processing in Acoustics*, D. Havelock, S. Kuwano, and M. Vorländer, Eds. New York, NY: Springer New York, 2008, pp. 1185–1213.

[72] N. Wang and J. Huang, "The modulation of tactile stimulation on Mandarin tone production in children with cochlear implant," vol. 4, 2010, pp. 3199–3202.

[73] T. Wilcox, R. Woods, C. Chapa, and S. McCurry, "Multisensory exploration and object individuation in infancy." *Developmental psychology*,

vol. 43, no. 2, p. 479, 2007.

[74] M. J. Wolf, *The Video Game Explosion: A History from PONG to Playstation and Beyond.* Bloomsbury Publishing USA, 2007.

[75] Y. Xiang, Z. Zhang, D. Chang, and L. Tu, "The impact of gamified auditory-verbal training for hearing-challenged children at intermediate and advanced rehabilitation stages," *arXiv preprint arXiv:2310.11047*, 2023.

[76] A. Xohua-Chacón, E. Benítez-Guerrero, J. Muñoz-Arteaga, and C. Mezura-Godoy, "Using a tangible system to promote inclusive, collaborative activities to learn relational algebra for students with hearing impairment," *Universal Access in the Information Society*, 2022.

[77] H. Yuan, C. M. Reed, and N. I. Durlach, "Tactual display of consonant voicing as a supplement to lipreading." *The Journal of the Acoustical Society of America*, vol. 118, no. 2, pp. 1003–1015, Aug. 2005.

Paper B

Tickle Tuner - Haptic Smartphone Cover for Cochlear Implant Users' Musical Training

F. Ganis, M. Vatti, and S. Serafin

The article has been submitted to
Haptic and Audio Interaction Design London, United Kingdom, 2022.

© 2022, The Authors, Reproduced with permission from Springer Nature
The layout has been revised.

Original Publication

F. Ganis, M. Vatti, and S. Serafin, ”Tickle Tuner - Haptic Smartphone Cover for Cochlear Implant Users’ Musical Training,” in *Haptic and Audio Interaction Design, vol. 13417, C. Saitis, I. Farkhatdinov, and S. Papetti, Eds., in Lecture Notes in Computer Science*, vol. 13417. , Cham: Springer International Publishing, 2022, pp. 14–24. doi: 10.1007/978-3-031-15019-7_2.

Abstract

Cochlear implant (CI) allow hearing impaired individuals to understand speech with remarkable efficiency. On the other hand, they poorly perform in music perception. It may be possible to improve the music experience with the use of other senses such as touch. We present Tickle Tuner, a haptic feedback device suitable for musical training of CI users. The prototype is composed of two high-quality haptic actuators and an external digital-to-analog converter (DAC) hosted in a 3D printed enclosure coupled with a smartphone. We describe the design and implementation of the prototype, the analysis of its characteristics and we introduce a test bench for the design of different mappings between sound and vibrations.

B.1 Introduction

Cochlear implant (CI) are neuroprosthesis that allow people with severe or profound hearing loss to restore their sound perception. They are especially successful in reestablishing speech comprehension [6]. Music is a highly complex phenomenon and during the electrical stimulation of the auditory pathway, there are many factors that can influence CI users' access and enjoyment [5, 18, 19]. These limitations are translated into reduced pitch range and lack of low frequencies, incapability of distinguish consonance and dissonance [15]. However, there is emerging evidence that music training may improve music perception skills and enjoyment [15]. One of the novel techniques to improve musical listening performance is the use of haptic feedback [8]. While vibrotactile feedback for speech understanding has a long history, starting with the pioneering work of Cowan et al. [3], the same cannot be said for musical appreciation connected to hearing impairment. Studies have shown that haptics improve melody recognition [20], instrument identification [23] and emotional communication [17]. Additionally, haptics have been shown to improve speech-in-noise performance and sound localization [9, 10]. One of the most crucial parts of this project is how the haptics are coupled with sound. The vibrotactile sensitivity is affected by several parameters such as location of the stimuli, dimension of the contact area and type of signal [16, 26]. In addition, other perceptive and physiological characteristics such as temporal masking and energy integration make even more difficult to disentangle the correlation between haptics and sound [21]. Thus, we hope that different mappings and user tests will help to clarify the situation, finding correspondence between audio and haptic cues. Haptic feedback is widely used in video game consoles and smartphones to enhance the visual content and Singhal and Schneider [24] propose a vibrotactile embellishment to improve the overall player's experience. In Weber and Saitis' work [27] a conceptual framework and different methods are proposed to translate audio to vibrotactile feedback. Other researchers found inspiration directly from the natural haptic feedback conveyed through musical instruments [2] or propose to use digital signal processing (DSP) techniques to extrapolate audio features and re-map them to haptic features [1]. Finally, Dementyev et al. [4] propose a vibrotactile haptic platform with some practical applications.

In this paper we propose the Tickle Tuner, a vibrotactile feedback device that provides haptic information during musical training that can be performed using *ad hoc* mobile games. These applications are developed for improving CIs skills to better recognize musical features and thus increase music appreciation. The name of our prototype derives from the Tickle Talker, the first device that uses haptic feedback to aid speech understanding [3]. The prototype presented was developed after considering the limitations of different projects such as wristbands, vests [13, 28] or haptic actuators embedded in furniture [25] (chairs and sofas). The wearable solutions (e.g. wristbands and vests), require the user to wear an additional device and can elicit potentially uncomfortable experiences whereas haptic chairs and sofas bound the user to a specific location. Hence we focused on designing a prototype for mobile phones since they are part of most of our daily activities, from work to entertainment. We aimed for an external device, as the actuators available in smartphones are limited in bandwidth and linearity, and the overall performance varies a lot among models [27].

B.2 Prototype design and implementation

The smartphone cover is mainly composed by two parts: shell and handles. The shell features an adjustable rail system and a frame that holds in position the smartphone, connects the two handles and hosts part of the cables. The handles are the main contact area with the user's hands and should be shaped in order to have a stable and comfortable grip. We choose the 3D printing technology to obtain a durable and comfortable object since it is possible to use plastic materials and a good reproduction quality can be easily achieved. For our prototype we used polylactic acid (PLA), a common and affordable thermoplastic polyester used for 3D prints. The whole process involved four different software and an iterative approach between a print and the next design. The Tickle Tuner is tailored-made for the Android phone Poco X3 NFC but it can be easily adapted to other smartphones or even tablets.

The skeleton has a fairly simple form with only some cavities to accommodate cables and rubber bands. Its shape can be easily represented by sum and subtraction of primitives. For this reason, we choose OpenSCAD[1] since it allows the user to create 3D models with this intuitive technique. In future developments and adaptations to different devices (e.g. different smartphones), we will be able to rapidly change dimensions and proportions of our prototype between each iteration only using the user defined variables in the software. The skeleton is divided in two parts, each of them corresponding to one handle. To connect both sides, we used a single rail system where two cuboids slide inside each other pulled by the force exerted by a rubber band. Two T shaped hooks are placed on the smallest rail and a track is engraved on the biggest one to hold in place the rails and preventing them from moving on y and z axes. On the right side, we equipped the skeleton with a lid and some space underneath to place an external digital-to-analog converters (DACs). The lid can be easily secured or removed with two screws.

The handles have great importance in the design process since we aimed to have a steady and comfortable grip as possible. We used clay to mold a prototype of a handle, focusing on its ergonomic shape and on the fingers' grooves. Once we obtained a satisfying result, we used photogrammetry to reconstruct a digital version of the model. Finally, we loaded the pictures in Meshroom[2] to retrieve the digital model. The reconstruction was ported into a Blender[3] to remove some irregularities and artifacts created during the photogrammetry process. Since the handles are in direct contact with the user's skin, we inserted a small box that perfectly contains the Haptuators. We filled the gaps with PLA connecting three sides of each haptic actuator box to the handle in order to conduct the highest amount of vibrations as possible.

---

[1] `https://openscad.org`, last access December 17, 2025

[2] `https://alicevision.org/#meshroom`, last access December 17, 2025

[3] `https://www.blender.org/`, last access December 17, 2025

(a) *Front view* **(b)** *Back view*

Fig. B.1: The Tickle Tuner prototype.

Once we obtained the digital version of the skeleton and the handles, we ported them into Blender project to produce a model suitable for the 3D print. Using a Ultimaker 3 printer[4] with the fastest settings as possible, we were able to print the whole prototype in approximately 12.30 hours using 100 grams of PLA.

B.2.1 Circuitry

The Tickle Tuner features two HapCoil-One (Haptuator Mark II-D) actuators produced by Actronika[5] that can reproduce frequencies from 10 to 10000 Hz. They are both connected to a 3W stereo class-D amplifier PAM8403 chip on a DFR0119 board. This amplifier perfectly fits our needs since features low THD+N, low power consumption and two output channels. The smartphone feeds the Haptuators through a digital to analog converter chip (DACs) that receives the audio stream through a USB-C plug. We slightly modified the DACs soldering two cables from the pin-out of the USB-C connector to retrieve DC (+5V) and ground (GND) for powering the amplifier. In this way, with a single plug the Tickle Tuner is able to retrieve the audio signal and the power recalling the *plug and play* concept.

---

[4] `https://ultimaker.com/3d-printers/ultimaker-3`, last access December 17, 2025

[5] `https://www.actronika.com/`, last access December 17, 2025

B.2.2 Analysis

In order to understand the acoustical characteristics of the Tickle Tuner, we measured its frequency response. Following the guidelines of Farina at al. [7], we wrote a MATLAB script that generates a exponentially-swept sine signal following the equation:

$$x(t) = \sin\left(\frac{\omega \cdot T}{\ln\left(\frac{\omega_2}{\omega_1}\right)} \cdot \left(e^{\frac{t}{T}\cdot\ln\left(\frac{\omega_2}{\omega_1}\right)-1}\right)\right) \tag{B.1}$$

The sine sweeps are repeated in order to perform an average of all the sweeps reducing the noise effect of the recordings and thus improving the Signal-To-Noise ration (S/N) [7]. We performed the measurements in an anechoic room located at Aalborg University Copenhagen to prevent any noise from altering the results. We placed an analog accelerometer (Sparkfun ADXL335) on seven different areas of the device and recorded the vibrations from one single output axe of the accelerometer. The areas taken into account are: top, bottom and side of each handle and the center of the smartphone screen. The signal was captured using a Stainberg UR44C audio interface at 44.1 kHz sampling rate. The impulse responses retrieved from the different locations are fairly similar and present some common characteristics such as a peak in the low range around 70 Hz and lower energy in the highest section of the spectra. The average of all the repetitions in all the seven locations is reported in figure B.2 and from now on we will refer to it as the frequency response of the Tickle Tuner.

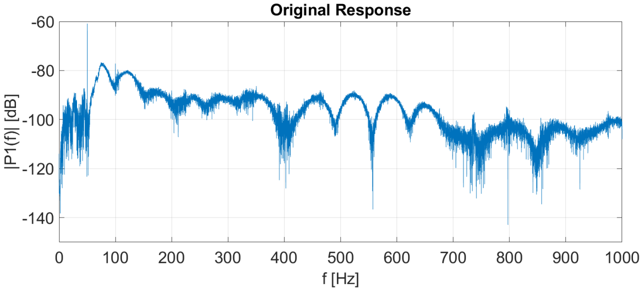


Fig. B.2: Frequency response of the Tickle Tuner.

B.2.3 Filter design

Feeding the Tickle Tuner with unfiltered audio from the smartphone creates a satisfactory haptic feedback that is possibly enhanced by a video source. The device provides robust low frequency vibrations and we expect that it will contribute towards improving the perception of the lowest portion of the sound spectra in CIs users. In order to mitigate the unwanted resonances of our device, we designed 7 biquad filters of the second order. In addition, we developed a second filter that compensates the fingertips' sensibility attenuating the mid-low range of the spectra with a center frequency of 250 Hz [16].

B.3 Methods

In the following section of the paper we will discuss the test of melodic contour identification (MCI) mapping.

B.3.1 Participants

Fifteen normal hearing participants were selected through convenient sampling of which 13 males and 2 females (age M = 28.3, STD = 4.7). The participants self reported an average of 13.5 years of musical experience (STD = 9.7).

B.3.2 Stimuli and mappings

To design the experiment, we initially investigated different mappings developing a test bench in Pure Data[6] coupled with a Graphical User Interface (GUI) using the platform MobMuPlat[7] to have a comfortable access to the parameters from the touchscreen. The software currently features several functionalities such as waves generators (sine, square, triangle, saw-tooth), subtractive synthesizer, ADSR control and sequencer, audio player, pitch tracking algorithm (with 4 controllable partials), filters for frequency response compensation and fingertips sensitivity compensation. Every generator has its own knob for controlling the fundamental frequency. The filter for the subtractive synthesis features also a knob for the quality (Q) factor. All the audio sources have

---

[6] `https://puredata.info/`, last access December 17, 2025

[7] `https://danieliglesia.com/mobmuplat/`, last access December 17, 2025

separate volume control and independent routing to the audio and the haptics' channels. The filters are the same implemented in MATLAB (B.2.3) and are applied only to the haptics channel to leave the audio signal untouched. One specific mapping born after some iterations and with the test bench is the amplitude modulation (AM) of the first partial estimated through the *sigmund* $\sim$ object. This approach has been inspired by Park and Choi [14] work where they investigate the perceptual relations between amplitude modulated signals conveyed through haptics. The algorithm implemented applies AM to the fundamental frequency f_0 with the modulating frequency f_m directly dependent from the f_0. The minimum and maximum frequencies can be set accordingly to the user's needs and the final f_m is thus scaled proportionally in a range between 0 and 20 Hz. The AM method used is called *double-sideband suppressed carrier (DBSC)* since the modulated signal presents two sidebands with half the energy of the carrier.

To test this mapping, We chose four different sound sources to cover different ranges of the spectrum and timbre combinations. The clarinet and viola sound stimuli were generated using the Audio Modeling SWAM plug-in suite[8] that uses a combination of physical models and recorded samples to obtain an extremely convincing simulation of real instruments. The piano was recorded using the Addictive Keys Grand Piano plug-in[9] that uses recorded samples and a large variety of microphones. The last sound source is a simple sine wave generated from a MATLAB script. Each MIDI melodic contour was played through the above mentioned plug-ins using a sample rate of 44.1 kHz and 16 bit of depth. Moreover, we used the Cochlear Implant Simulator AngelSim™[10] with a 8-channel noise vocoded speech simulation to modify the audio stimuli and simulate how a CI might sound. We used three different mappings: the AM, full audio through the haptic actuators and no haptics.

---

[8] `https://audiomodeling.com/swam-engine/`, last access December 17, 2025

[9] `https://www.xlnaudio.com/products/addictivekeys`, last access December 17, 2025

[10] `http://www.tigerspeech.com/angelsim/angelsimabout`, last access December 17, 2025

B.3.3 MCI test variant

This test was developed to investigate the effect of the AM on the perceived pitch using different sound sources. We developed a *MCI* task inspired by Galvin et al. [11, 12] and Omran et al. [22]. The test was designed with a similar interface to the testbench using PureData and MobMuPlat. In this way, the participants interacted directly on the Tickle Tuner's for the whole duration of the test. For this setup, we selected six different melodic contours. The played notes ranged from C3 (130.81 Hz) to G3# (207.65 Hz). The different combinations were written into MIDI files using Reaper[11] to later feed different audio plug-ins. Each note lasted one quarter at 120 beats per minute (0.5 seconds) and the contours were generated with one or two semitones of distance between each note. Thus, the test used twelve different melodic contours (6 contour types $\times$ 2 distances between each note).

B.3.4 Procedures

The test was conducted in an anechoic room of the Aalborg University Campus in Copenhagen. Before the beginning of the test, each participant experienced a mock session to get acquainted with the interface and the mechanics of the test. The training page, with the same aspect of the test page, showed to the participants the different combination of contours. Once the next contour was generated, it was possible to listen to the correspondent melody played by the piano (without CI simulation) as well as feel the AM mapping on the haptic actuators. For every contour, the GUI highlighted the correct answer with a green label. The training session lasted approximately 2 minutes and no information regarding how the mappings work or which sound sources has been used were shared with the participants before the end of the test.

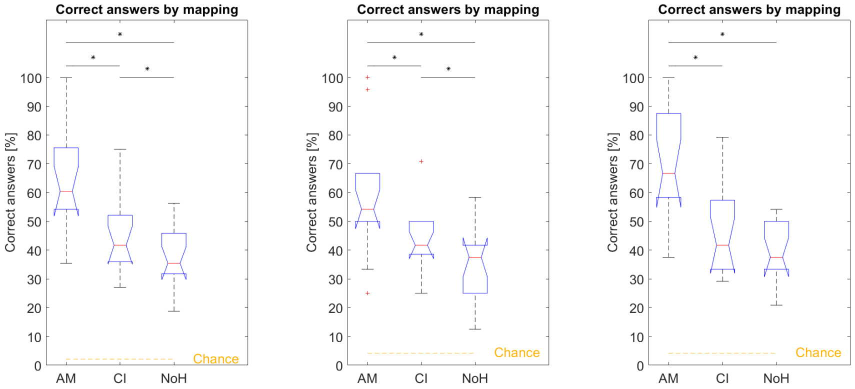
B.4 Results and discussion

*The collected data present no correlation between the musical experience and the scores obtained during the MCI task with CI simulation (correlation coefficient = 0.37). This might be explained because the contours are extremely difficult to recognize when there is a small distance between each note (semitone and tone) and a CI simulation is used [12]. Figure B.3a shows the pooled

---

[11] `https://www.reaper.fm/`, last access December 17, 2025

results for all the instruments and the two possible distances between the notes (1 or 2 semitones) combined. AM mapping has a mean score of 65% correct answers while CI (only audio) is 45% and no haptic 38.33%. Running an one-way ANOVA test and t-tests on pairs of data sets, it is noticeable a significant difference between all the data, with a p-value $\ll 0.05$. This initial analysis shows a good performance of the AM mapping, remembering that the participants did not receive any information about how the mappings work. In figure B.3b and B.3c scores for all the instruments are depicted with one and two semitones respectively. Comparing the two graphs, it is possible to notice that the two semitones are easier to recognize (as expected), and in the two semitones answers there is no significant difference between the full audio and the no haptic feedback mappings. Moreover, the users scored an average of approximately 10% better results with two semitones than with one semitone on the AM mapping. Running a t-test analysis on each mapping comparing the percentage of correct answer for one against two semitones, significant difference (p-value < 0.05) has been found in both AM and NoH mappings while for CI the semitone change does not have any effect.

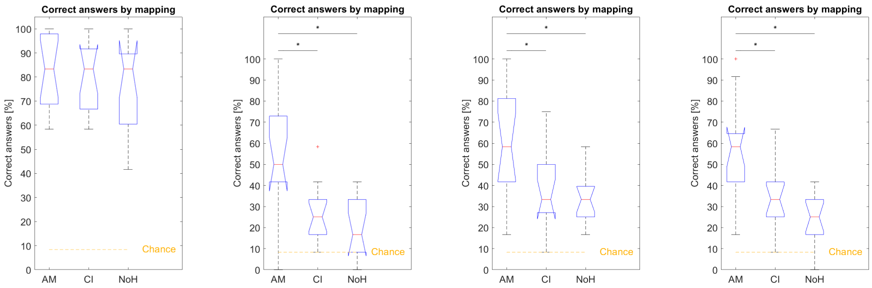


(a) *One and two semitones combined.* **(b)** *One semitone.* **(c)** *Two semitones.*

Fig. B.3: Percentage of correct answers, where AM is the Amplitude Modulation mapping, CI is the full audio through haptics and NoH is no haptic feedback.

In figure B.4 we present the answers filtered by instrument to show how the mappings affected the percentage of correct answer in each condition. The sine wave scored close to 80% for each mapping. This can be explained thanks to the fact that the CIs simulation can adequately reproduce a simple sine wave and most probably the users used mainly the sound stimuli to recognize the correct contour. Removing the sine wave from the data set, a bigger interaction between instrument and mappings is noticeable since AM mappings achieved a mean of 59.07% and audio only mapping 32.78% (gap of 26.29%). Looking at the other instruments, a common trait is the non-significant difference between the full audio (CI) and the no haptics mappings that scored both a low median value of correct answers. The viola (figure B.4b) has the most evident gap between the AM and the other two mappings. Moreover, full audio and no haptics mapping were very difficult to recognize, most probably due to the specific spectral characteristics of the instrument (the fundamental frequency is not very pronounced).

The test demonstrated the validity of the AM mapping to better convey pitch information without any user training compared to no haptic feedback. On the other hand, some limitations should be also highlighted. In fact, the pool of subjects is modest and the CIs simulation is a mere hardware simulation that cannot take into account the very personal hearing panorama that each person with CIs experiences.



(a) *Sine wave.* **(b)** *Viola.* **(c)** *Clarinet.* **(d)** *Piano.*

Fig. B.4: Correct answers percentage for each instrument (1-2 semitones combined).

B.5 Conclusions

In this paper we presented a stand-alone working prototype coupled with a smartphone for CIs users' musical training. The prototype features high quality haptic feedback with an ergonomic interaction. We recorded its frequency response and created a multi-band equalizer to compensate internal resonances and spectral irregularities. Furthermore, we designed a test bench to facilitate the process of audio-haptic mappings creation and designed an audio-haptic mapping suitable for melodic contour identification tasks. We plan to upgrade the prototype with both hardware and software improvements and to conduct further tests to assess its effectiveness in musical training for CI.

References

[1] D. M. Birnbaum and M. M. Wanderley, "A SYSTEMATIC APPROACH TO MUSICAL VIBROTACTILE FEEDBACK," p. 8, 2011.

[2] C. Chafe, "Tactile audio feedback," in *Proceedings of the International Computer Music Conference.* INTERNATIONAL COMPUTER MUSIC ACCOCIATION, 1993, pp. 76–76.

[3] R. S. Cowan, J. Z. Sarant, K. L. Galvin, J. I. Alcantara, P. J. Blamey, and G. M. Clark, "The Tickle Talker: A speech perception aid for profoundly hearing impaired children," *Scientific publications, vol. 5, 1989-1990, no. 300*, 1990.

[4] A. Dementyev, P. Getreuer, D. Kanevsky, M. Slaney, and R. Lyon, "VHP: Vibrotactile Haptics Platform for On-body Applications," p. 15, 2021.

[5] Ward. 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, no. 2, pp. 114–123, Feb. 2015.

[6] Fan-Gang Zeng, S. Rebscher, W. Harrison, Xiaoan Sun, and Haihong Feng, "Cochlear Implants: System Design, Integration, and Evaluation," *IEEE Reviews in Biomedical Engineering*, vol. 1, pp. 115–142, 2008.

[7] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," p. 24, 2000.

[8] M. D. Fletcher, "Can Haptic Stimulation Enhance Music Perception in Hearing-Impaired Listeners?" *Frontiers in Neuroscience*, vol. 15, p. 723877, Aug. 2021.

[9] M. D. Fletcher, R. O. Cunningham, and S. R. Mills, "Electro-Haptic Enhancement of Spatial Hearing in Cochlear Implant Users," *Scientific Reports*, vol. 10, no. 1, p. 1621, Dec. 2020.

[10] M. D. Fletcher, A. Hadeedi, T. Goehring, and S. R. Mills, “Electro-haptic enhancement of speech-in-noise performance in cochlear implant users.” *Scientific reports*, vol. 9, no. 1, p. 11428, Aug. 2019.

[11] J. J. Galvin, Q.-J. Fu, and G. Nogaki, “Melodic Contour Identification by Cochlear Implant Listeners,” *Ear & Hearing*, vol. 28, no. 3, pp. 302–319, Jun. 2007.

[12] J. J. Galvin, Q.-J. Fu, and R. V. Shannon, “Melodic Contour Identification and Music Perception by Cochlear Implant Users,” *Annals of the New York Academy of Sciences*, vol. 1169, no. 1, pp. 518–533, Jul. 2009.

[13] G. Garcia-Valle, M. Ferre, J. Brenosa, and D. Vargas, “Evaluation of Presence in Virtual Environments: Haptic Vest and User’s Haptic Skills,” *IEEE Access*, vol. 6, pp. 7224–7233, 2018.

[14] Gunhyuk Park and Seungmoon Choi, “Perceptual space of amplitude-modulated vibrotactile stimuli,” in *2011 IEEE World Haptics Conference*. Istanbul: IEEE, Jun. 2011, pp. 59–64.

[15] N. T. Jiam, M. T. Caldwell, and C. J. Limb, “What Does Music Sound Like for a Cochlear Implant User?” *Otology & Neurotology*, vol. 38, no. 8, pp. e240–e247, Sep. 2017.

[16] L. A. Jones and N. B. Sarter, “Tactile Displays: Guidance for Their Design and Application,” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 50, no. 1, pp. 90–111, Feb. 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, no. 3, pp. 160–169, Jul. 2009.

[18] C. J. Limb and A. T. Roy, “Technological, biological, and acoustical constraints to music perception in cochlear implant users,” *Hearing Research*, vol. 308, pp. 13–26, Feb. 2014.

[19] V. Looi, H. Mcdermott, C. M. McKay, and L. Hickson, “The effect of cochlear implantation on music perception by adults with usable preoperative acoustic hearing,” *International Journal of Audiology*, vol. 47, pp. 257–268, 2008.

[20] 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, p. 1145, Oct. 2019.

[21] S. Merchel and M. E. Altinsoy, “Psychophysical comparison of the au-

ditory and tactile perception: A survey," *Journal on Multimodal User Interfaces*, vol. 14, pp. 271–283, 2020.

[22] S. Omran, W. Lai, and N. Dillier, "Pitch Ranking, Melody Contour and Instrument Recognition Tests Using Two Semitone Frequency Maps for Nucleus Cochlear Implants," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2010, no. 1, p. 948565, 2010.

[23] F. A. Russo, P. Ammirante, and D. I. Fels, "Vibrotactile discrimination of musical timbre." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 38, no. 4, pp. 822–826, 2012.

[24] T. Singhal and O. Schneider, "Juicy haptic design: Vibrotactile embellishments can improve player experience in games," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–11.

[25] H. Z. Tan, R. Gray, J. J. Young, and R. Traylor, "A Haptic Back Display for Attentional and Directional Cueing," p. 20, 2003.

[26] R. T. Verrillo, "Vibration Sensation in Humans," *Music Perception*, vol. 9, no. 3, pp. 281–302, Apr. 1992.

[27] M. Weber and C. Saitis, "Towards a framework for ubiquitous audio-tactile design," p. 9, 2020.

[28] J. Wu, J. Zhang, J. Yan, W. Liu, and G. Song, "Design of a Vibrotactile Vest for Contour Perception," *International Journal of Advanced Robotic Systems*, vol. 9, no. 5, p. 166, Nov. 2012.

Paper C

Vibrotactile Memory: A Case Study of Timbre Perception Training in Children with Cochlear Implants Using a Video Game

F. Ganis A. Adjorlu, L. Percy-Smith, C. F. Samar, and S. Serafin

The article has been submitted to
The 21th Sound and Music Computing Conference Conference Proceedings,
2024.

© 2024, The Authors, under CC 3.0
The layout has been revised.

Original Publication

F. Ganis, A. Adjorlu, L. Percy-Smith, C. F. Samar and S. Serafin, “Vibrotactile Memory: a Case Study of Timbre Perception Training in Children with Cochlear Implants Using a Video Game”, in *Proceedings of the 21st Sound and Music Computing Conference*, ESMAE, Porto, Portugal, Jul. 2024, pp. 134–141. doi: 10.5281/zenodo.14336923.

Abstract

This paper presents a preliminary study investigating the effectiveness of a computer game designed for enhancing timbre perception in children with cochlear implants (CIs). The game focuses on instrument recognition and incorporates vibrotactile feedback through the PlayStation 5 DualSense controller. A case study was conducted involving a 7-year-old participant diagnosed with hearing loss due to auditory neuropathy who uses both a CI and a hearing aid. Training sessions were conducted over a two-week period, with assessments performed before and after the training thanks to a modified version of the Timbre Perception Test (TPT). Despite incorporating vibrotactile feedback and tracking participant progress, the study did not observe significant improvements in timbre perception. Factors such as the short duration of the training period, the difficulty of the game, and potential boredom may have influenced the outcomes. The study highlights the need for further research to refine the training intervention, including extending the duration, enhancing engagement, and conducting larger-scale studies with diverse participant groups.

C.1 Introduction

Cochlear implants (CIs) are surgically implanted neural prostheses designed to restore hearing in people with profound hearing loss (HL) [23]. Until recently, researchers and companies focused mainly on optimizing these devices for speech, obtaining good results in reestablishing language comprehension [8]. These devices, however, still present challenges for music appreciation. Music, in fact, is a highly complex phenomenon with many factors that can influence its enjoyment and access for hearing aids (HAs) and CI users [6, 20, 22]. The hardware, software, and physical limitations of CIs result in a reduced frequency range, incapability of distinguishing sounds' consonance and dissonance, limited dynamic range, and difficulty in recognizing instruments (timbre) [18]. Consequently, music enjoyment and the emotions that it can elicit are directly affected, most often resulting in limited fruition. In addition, musical background, degree of success of the surgery, as well as the presence of other pathologies, have an effect on the hearing experience, making every user's sound experience unique.

These general considerations apply to young recipients too, but some additional aspects have to be considered when children are taken into account. Early cochlear implantation has been shown to yield long-term positive effects, providing children with speech capabilities comparable to their normal hearing (NH) peers before the age of four [31]. After age four, young recipients of CIs develop speech slower than NH individuals, making personalization of learning goals a key aspect of training [14]. Ad-hoc training is even more important when the child suffers from auditory neuropathy spectrum disorder (ANSD), a pathology that implies disrupted neural activity that impairs timing perception, low frequencies' perception, temporal integration, gap detection, sound localization among other things [32]. However, children with HL diagnosed in recent years that are fitted with HA and CI before one year of age, may perform differently from previous generations.

Emerging evidence shows that music training improves music perception skills and experience [11, 18, 28], together with psychosocial well being and quality of life [21]. One of the novel techniques to possibly improve musical listening performance is the use of vibrotactile feedback [10]. Vibrotactile feedback refers to the mappings between audio and vibrations that can help to emphasize specific features such as pitch [13, 15]. Another common approach that improves learning experiences for children is the gamification of the training activities [7, 17]. This strategy can guide youngsters with HL towards improving their musical skills [16], especially when HL is due to ANSD.

Gamification has demonstrated its effectiveness as a tool for enhancing engagement and motivation across various domains, ranging from business to marketing [2]. It has also shown positive impacts on the learning process, improving engagement and knowledge acquisition [1]. Consequently, gamification appears to be a promising approach for boosting participation in training activities among individuals with hearing impairments. However, it remains a relatively new area of inquiry that has yet to be thoroughly explored [12].

In this paper, we introduce a computer game for musical training of children with cochlear implants, with the aim to improve timbre perception through instrument recognition. The game presents different levels of difficulty and features vibrotactile feedback conveyed through the PlayStation 5 controller DualSense[1]. We report a preliminary case study [4] to collect information

---

[1] Sony Playstation DualSense Controller `https://www.playstation.com/en-dk/accessories/dualsense-wireless-controller/` — Last access

about user interaction and efficacy to train the timbre perception.

In Section C.2 we present the methodology applied for this case study, including the process and design choices made to create the training video game; in Section C.3 we showcase the data retrieved from this experiment and finally in Section C.4 we analyze the results and draw some considerations related to the whole case study.

C.2 Methodology

C.2.1 Participant

The participant was 7 years old at the time of the experiment. At the age of 4, she received a cochlear implant in her left ear, and she uses a hearing aid in her right ear (bi-modal). She has been diagnosed with ANSD, a condition that can pose challenges in the transmission of signals from the auditory nerve to the brain's auditory cortex. No assessment of the actual hearing capabilities of the subject has been performed for the scope of this project. We did not characterize the user's perception of vibrotactile stimulation; given the participant's young age, we assumed they have ideal sensitivity capabilities, as most degradation occurs with aging or due to traumas [24] (e.g., heavy-machinery use). The participant has been recruited through the Center for Hearing and Balance from the Rigshospitalet of Copenhagen with approval from the parents and the clinicians.

C.2.2 Design

In the following sections, we present the project's design process and the experiment.

Participatory Design

We applied the principles of *participatory design* [27, 29] involving the end user in the process to open a dialogue to gather ideas and feedback throughout the development. The child, together with the parents and the audio-verbal therapist contributed to the process through an official meeting and some more informal gatherings. During the first meeting, we collected information about the preferred video games from the child as well as clinical practices and training tools already available. In the follow-ups, we mainly included the clinician to receive feedback about the experiment as well as the interface.

Experiment Design

We designed an experiment that features a video game aimed at training children to improve their timbre perception through listening to musical instruments. The experiment includes an initial assessment of timbre perception, a training period using the video game, and a final assessment of timbre perception. In addition to sound feedback, vibrotactile feedback was incorporated to convey musical information through the sense of touch. The hardware setup comprised a laptop equipped with loudspeakers and a controller to provide both vibrotactile feedback and input capabilities. In Figure C.1 the simple setup is depicted.

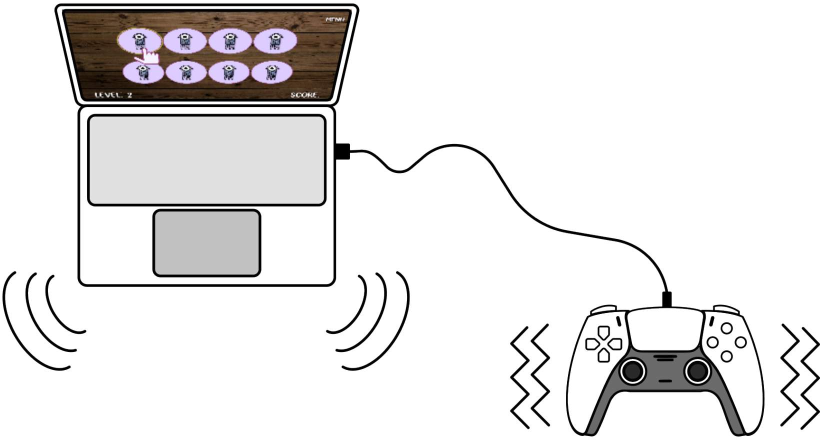


Fig. C.1: Setup for the training experience.

The experiment extended over a two-week period, with the participant conducting the training activity at home three times a week under the supervision of her parents. The training tool systematically tracked the progress of each participant, facilitating a comprehensive analysis of their performance. Ultimately, we conducted interviews with both the child and their parents at the conclusion of the training period to gain insights into their overall experience and gather feedback on potential enhancements to the prototype.

To assess the efficacy of our intervention, a modified version of the Timbre

Perception Test (TPT) [19] was administered both before and after the training in the clinic together with the researchers and the audiologist. The modifications involved scaling the parameters to create a more pronounced difference in the controls, ensuring they were easily audible to the participant.

C.2.3 Assumptions

Given a single participant and the specific design of the test, we did not formulate hypotheses but instead made some assumptions regarding our expectations concerning the training activity that involved the child with ANSD:

1. Increasing the training duration will lead to a reduction in the required interaction, measured by the number of clicks, with the cards.
2. Incorporating vibrotactile feedback will improve the ability to discriminate between the cards.

C.2.4 Training Environment

While creating the training game, we drew inspiration from the classic memory card game, where players take turns flipping pairs of cards to find matching pairs. In our design, we incorporated the same mechanics, but rather than requiring the player to focus on images, we asked them to concentrate on sounds, with particular attention to the timbre of the instrument being played. Once a card is selected, the corresponding sound (and eventually vibration) is played through the laptop's loudspeakers. Only when a matching pair is found, the cards turn, revealing the image of the instrument playing. This approach is borrowed from auditory verbal practice, where sound is provided before vision; thus, the patient is induced to focus first on sound, while visual feedback is given only at a later stage to reinforce the association [14] aiming to enhance both listening capabilities in terms of timbre quality and auditory memory.

We developed the game using the Unity game engine (*ver. 2021.3.11f1*) [30] in a 2D environment. User interaction is made possible through the use of the DualSense controller from Sony. This choice led us to face the problem of feeding the user with a synchronized and coherent vibrotactile signal with the audio. The controller, when connected with a USB cable, behaves as both an input device and vibrotactile feedback. To provide both sound and vibrotactile feedback on independent streams, we opted for controlling the audio and haptic stream through FMOD for Unity [9]. This allows to simultaneously use the integrated sound card of the host laptop and the controller's internal sound card. In Figure C.2 we can observe a screenshot of the video game. In the bottom side of the screen, level and scores are reported. From the main menu of the game it is also possible to check the progress of the game and play again the completed levels for further training.

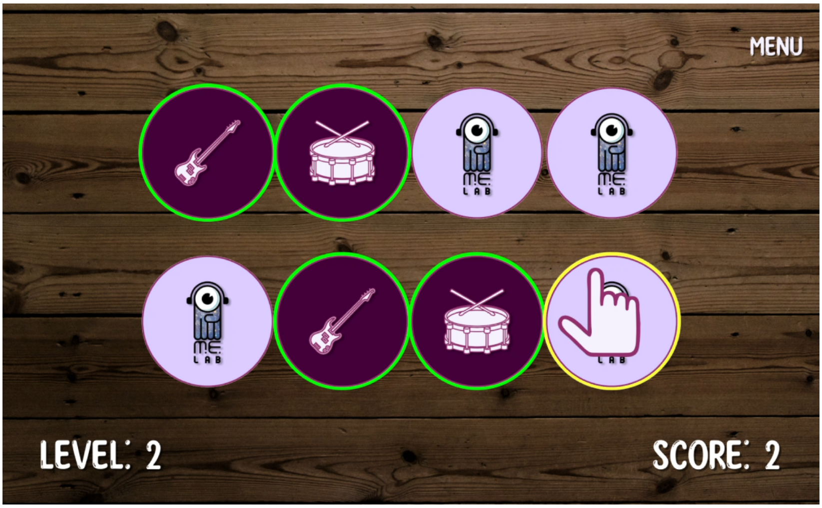


Fig. C.2: View of the memory game.

Level Generation

The game consists of 16 levels with increasing difficulty. The parameters that vary between the levels include the number and similarity of instruments, as well as the melodies played. Instrument similarity is defined in terms of timbre, where, for instance, a violin is very similar to a viola but remarkably different from a trumpet. Table C.1 illustrates how the levels are structured in the game. It's important to note that each level has two different versions—one with haptic feedback and one without. This explains why we report only 8 levels in Table C.1 while in the game there are 16. These versions run consecutively, though this detail is not included in the table for the sake of readability.

In order to prevent any bias, the instrument position and type, and the melodies are randomly assigned to each level and re-calculated at every run.

Level	# Instr.	Similarity	Melody
1.1	4	Different	Different
1.2	5	Different	Different
2.1	4	Different	Same
2.2	5	Different	Same
3.1	4	Similar	Different
3.2	5	Similar	Different
4.1	4	Similar	Same
4.2	5	Similar	Same

Table C.1: Level's difficulty structure

Data Acquisition

With the aim in mind of tracking the progress of the participant, we implemented a saving system that records the content of each level (i.e., instruments and melodies), the number of clicks per card, and the time it took to complete the level. All this information is stored in separate `.json` files saved in a hidden folder of the host computer's file system to prevent data manipulation from the end user.

C.2.5 Stimuli and Instruments

To exert the highest degree of control over the musical stimuli present in the experiment, we chose to generate all the recordings from MIDI files. This involved using a set of four AI-generated melodies[2], one ascending and one descending scale, and two famous excerpts (Eine Kleine Nachtmusik and Piano Sonata No. 16 "Sonata Semplice," both composed by Mozart), totaling eight melodies. All recordings are played at 100 BPM and last for two bars, covering a range from A3 to F5 (220 - 698.5 Hz). The MIDI files were employed to feed a set of audio plug-ins based on either sampled or physically modeled instruments to achieve high-quality audio and fidelity. The instruments selected for the experiment are listed in Table C.2, grouped by timbre similarity.

The sound stimuli were generated by loading the MIDI files into Reaper [3], a digital audio workstation (DAW), along with the virtual instruments shown

---

[2] Magenta Studio Generator — `https://magenta.tensorflow.org/studio` — Last access May 22, 2024

Instrument	Category	Instrument	Category
Xylophone	1	Violin	3
Piano	1	Trumpet	4
Guitar	1	Trombone	4
Bass	2	Flute	5
Cello	3	Sax	5
Viola	3	Clarinet	5

Table C.2: Instruments categories grouped by timbre similarity

in Table C.2. The individual audio files were then exported at 44.1 kHz and 16 bits. All the stimuli were normalized at -12 dB to maintain an equal perceived loudness. These same files were used for both the audio and vibrotactile feedback. The experiment was conducted feeding the vibrotactile actuators with unprocessed audio signal. This choice was driven by the type of training task: since we focus on timbre and musical instrument recognition, we wanted to preserve the frequency spectrum and the envelope of the instruments, following the approach used by Russo *et al.* [26]. In future iterations we will develop other training activities that will include melodic contour identification, and we will use different mapping techniques such as the one used in [13].

C.2.6 Vibrotactile Input

Humans' mechanoreceptive system is capable of perceiving vibrations up to approximately 1 kHz ([24]) and therefore choosing actuators able to cover this frequency range is crucial. In literature, it is possible to find several prototypes available such as vests, gloves, and furniture ([25]). These solutions are not suitable for our project, since the target group requires rugged devices that are easy to use and feature plug-and-play behavior. Consequently, we opted for a commercially available gaming console controller that children might be familiar with. In addition, these devices can be brought at home and connected to any laptop, giving a good degree of flexibility. Recently, Sony released the PlayStation 5 with the DualSense controller (2020) that features high quality vibrotactile feedback conveyed through voice-coil actuators. The working principle of this technology is similar to the loudspeaker, and thus shares a comparable performance in frequency range.

C.3 Results

C.3.1 Video Game's Data

As introduced in Section C.2.4, we tracked how many times the user interacted with the cards by clicking on them and listening to the sound. In Figure C.3, the total number of clicks per session can be seen. The number shown on top of each bar indicates the level in the video game. The bars with the green outline represent the levels with 5 pairs of cards, while the ones without represent 4 pairs of cards. Additionally, there are green and white lines representing the optimal number of moves for 5 and 4 pairs of cards, respectively, where the optimal number represents the expected number of flips (clicks in our case) with a player presenting a perfect memory as demonstrated by Velleman *et al.* [5]. The optimal moves for n pairs of cards are calculated as following:

$$(3 - 2\,ln\,2)\,n + 7/8 - 2\,ln\,2 \approx 1.61 \times n$$
$$n = 4 \rightarrow opt.moves \approx 6.44$$
$$n = 5 \rightarrow opt.moves \approx 8.05$$

In figure C.3 we can notice that the interaction between the cards for both levels with 4 and 5 pairs of cards are approximately twice the optimal moves reported in the formulas above.

Observing Figure C.4, there is no recognizable patter or trend that could suggest an effect of the vibrotactile feedback or total training time on the average number of clicks per card. The average number of clicks per card among all sessions is 2.42.

In Figure C.5 the amount of time elapsed to complete each level is reported. We can observe that the average duration for levels 11 to 16 (excluding the 14th session) is 6 minutes and 8 seconds. This means that participants took between 1 minute and 13 seconds and 1 minute and 32 seconds to find a correct pair, depending on whether the level presented 4 or 5 pairs of cards.

Figure C.6 shows the score deviation between the target sound and the answer from the participant. The values represent the number of steps from the slider used in the TPT interface. Running a t-test on the difference between the answer and the target values, we found a p-value of 0.2466 showing no

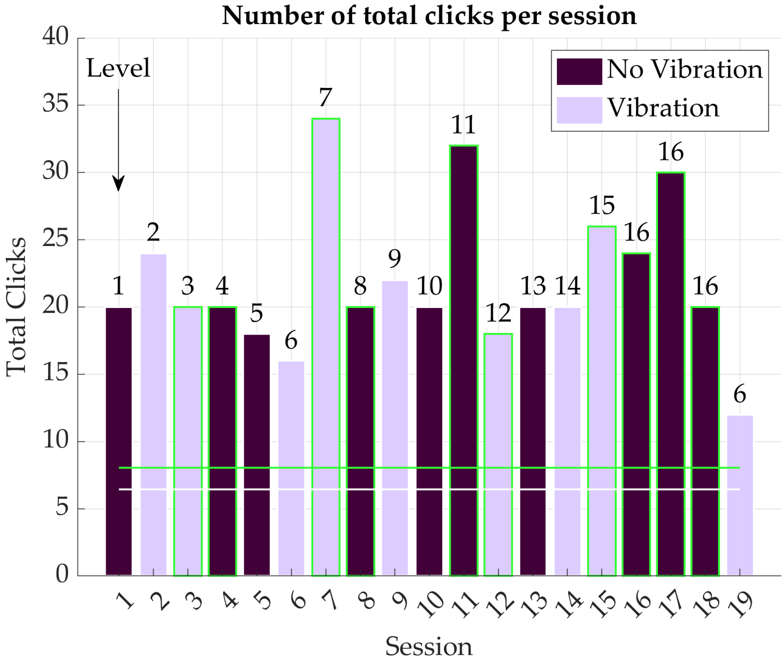


Fig. C.3: Total amount of clicks on the cards per every level. The levels with green outline are the ones with two extra cards. The white and green lines indicate the optimal moves for 8 and 10 cards respectively.

significant difference between the pre- and post-training. For more information we suggest referring to the manuscript [19].

C.3.2 Interview

To prevent any possible bias, we requested the auditory-verbal therapist to hold a brief interview with both parents and the child in the following session after the training. This approach capitalizes on the established trust between the parties and the familiarity between the child and the therapist. The therapist was given prior instructions to pose targeted questions and guide the discussion to gather relevant information. The conversation was conducted in Danish.

The child reported facing boredom during the training and would not like to repeat the training experience. The parents agree that the game "[...] seemed to be another training". On the other hand, they stated that the game made

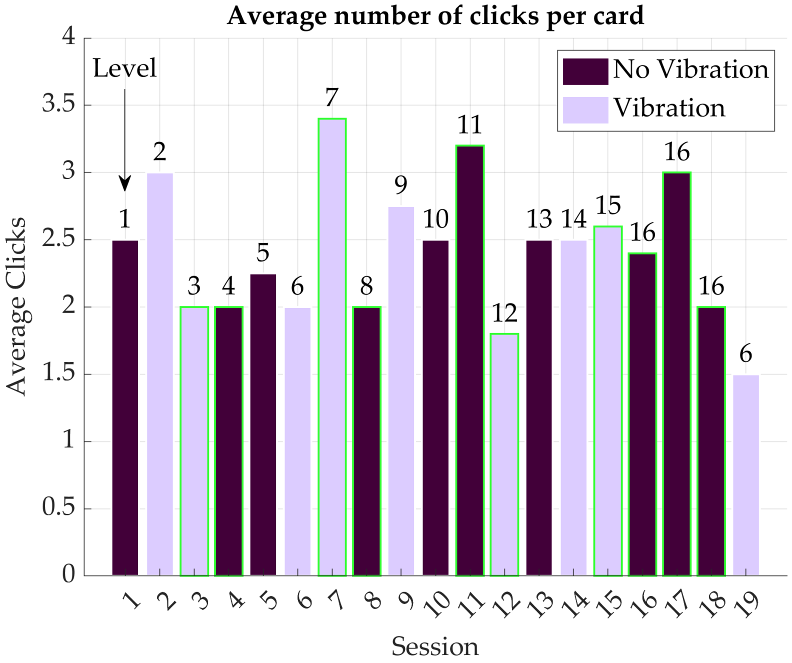


Fig. C.4: Average amount of clicks on the cards per every level. The levels with green outline are the ones with two extra cards.

their child think about the sounds, and she might have learned about the musical instruments. They had to supervise her during all the sessions also because she found a way to cheat, pressing the cards randomly until finding the matching pair. They think it could be suitable for younger recipients too. Finally, they suggest including a more effective rewarding system to improve engagement, and they also think it would be meaningful to be able to listen again to the instruments even after finding the correct pair.

C.4 Discussion

The findings of this study suggest that the video game intervention designed to enhance timbre perception in children did not produce the expected improvements in participant's performance on the TPT. Despite the incorporation of vibrotactile feedback, which was intended to provide additional sensory information and enhance participant's ability to discriminate between sounds,

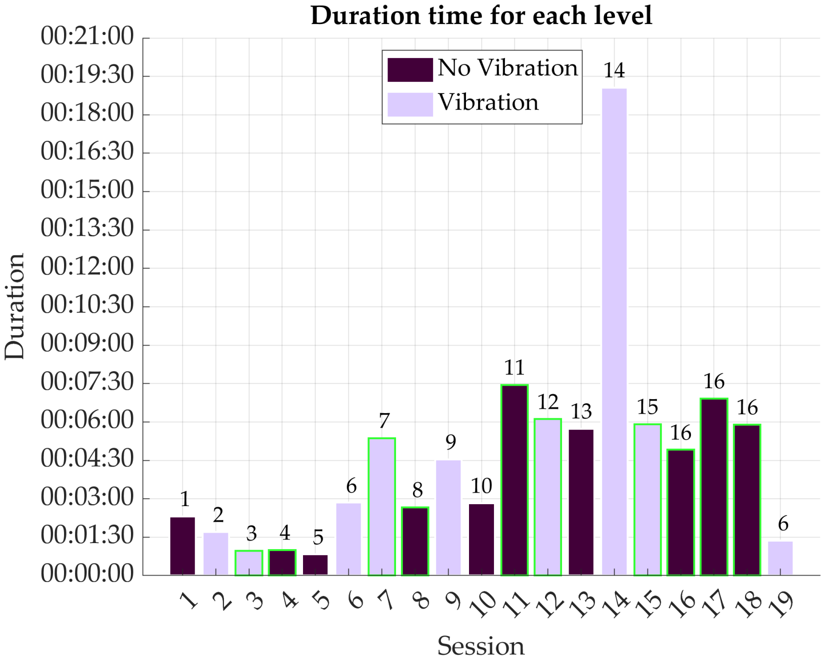


Fig. C.5: Time elapsed to complete each level.

the intervention did not yield significant results. This raises questions about the effectiveness of the training approach and potential factors influencing its outcomes.

Several factors may have contributed to the lack of observed improvement in timbre perception following the training intervention. Firstly, the relatively short duration of the training period, conducted over a two-week period with sessions held three times a week, may not have been sufficient for the participants to develop meaningful improvements in their auditory skills. Secondly, this duration did not allow for gathering enough data; we aimed to have more repetitions of the same levels to perform significant data analysis. The difficulty of the game and potential boredom with its repetitive nature could have impacted engagement and motivation, thereby limiting the effectiveness of the intervention as well as the production of data. This is a crucial aspect in the case study as it demonstrates how the current design includes gamification elements in the training activity without being successful. The child, in fact, reported that she usually plays mainstream video games with her

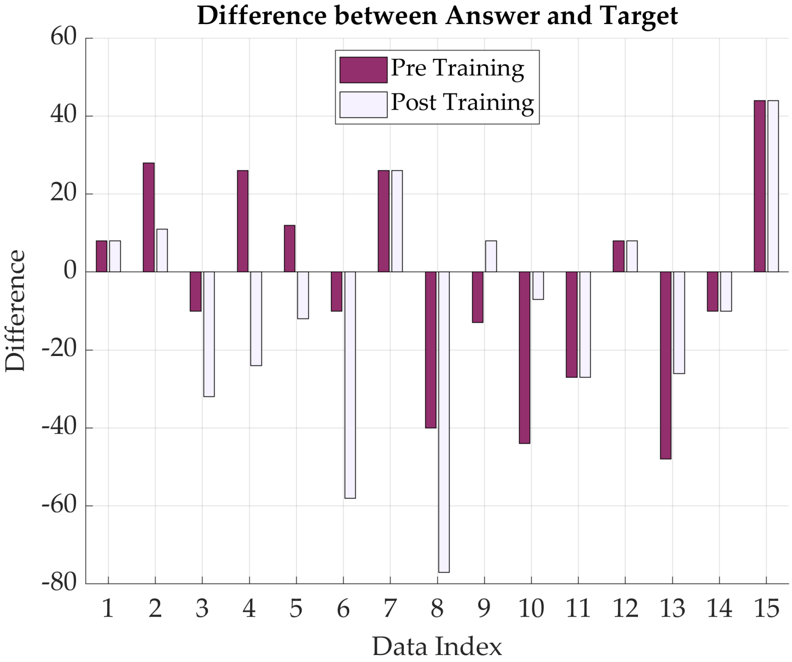


Fig. C.6: TPT before and after training comparison.

friends or alone, and thus we can assume that these games are her comparison term when talking about entertaining activities. The current design features basic gamification mechanisms and, as stated by the parents, could maybe be suitable for younger children in the age range between 4 and 5.

Another point of discussion is the difficulty of the task: from Figure C.5, we can observe that the average duration for levels 11 to 16 (excluding the 14th session) is 6 minutes and 8 seconds. This means that participants took between 1 minute and 13 seconds and 1 minute and 32 seconds to find a correct pair, depending on whether the level presented 4 or 5 pairs of cards. Thus, we can conclude that these levels were significantly more difficult than the first ten, indicating a greater challenge in finding similar instruments. This is congruent with our expectations since the harder level presents instruments with similar timbre, thus more challenging to distinguish. In the future, we will reconsider the complexity of the task, making the difficulty increase more gradually and extending the number of levels to follow a gentler learning curve. Level 14 can be considered an outlier, as the completion time is three times longer than the closest levels in difficulty. It can be assumed that the session was not closed before a break, leaving the timer running until the resumption of the session.

From the graphs in the Results Section C.3, no effect of vibrotactile feedback on both the average number of clicks and duration time is apparent. One hypothesis could be that the activity was extremely demanding for the auditory channel, leaving no space for attention on the tactile one. Another possibility is that the child did not fully grasp the connection between the feedback and the sound due to their young age. Alternatively, it could be interpreted that the vibrotactile feedback is not capable of conveying relevant information for musical timbre. This interpretation contrasts with the findings of Russo *et al.* [26]. However, it's worth noting that their study utilized a completely different setup (two voice-coils on a chair), presented the vibrotactile stimuli without sound and involved an older group of participants.

The TPT allowed us to measure the timbre perception of the participant but did not test the learning experience in terms of recognition of the specific instruments presented in the study. In fact, the parents of the child reported that the child might have learned about the instruments thanks to the training video game, but this has not been proven.

The findings of this study underscore the need for further research to refine and optimize the training intervention for timbre perception. Future iterations could focus on extending the duration of the training period to allow for more comprehensive skill development, but only with variations in the game design to increase engagement and motivation. For instance, the game could be part of a suite of different mini-games aimed at training musical skills of children with ANSD. The design will be put under revision to improve both the dynamics and the interaction with the aim of making the game more appealing to the target group.

It is important to acknowledge the limitations of this study, including the small sample size and the lack of a control group for comparison. As mentioned at the beginning of this paper, the study has to be considered as a case study that requires further validation. Allowing the participant to train at home ensured that the activity could be performed in an ecologically valid environment. However, this approach involved some compromises, such as the absence of researcher supervision, leading to uncertainty about the participant's interaction with the controller. Although the controller's design suggests a specific method of interaction and handhold position, we could not guarantee that the participant held the device optimally. Therefore, the participant's grip on the device might have affected the vibrotactile feedback contribution. Additionally, the use of a single participant group and the specific design of the training intervention may limit the generalizability of the findings. Future research should aim to address these limitations by conducting larger-scale studies with diverse participant groups and incorporating control conditions to more rigorously evaluate the effectiveness of the intervention.

C.5 Conclusions

In conclusion, while the initial findings of this study did not demonstrate significant improvements in timbre perception following the training intervention, they provide valuable insights into the challenges and considerations involved in developing effective auditory training interventions for children. By addressing the limitations and exploring alternative approaches, future research has the potential to enhance our understanding of auditory perception and improve the effectiveness of interventions aimed at developing auditory skills in children with CI and ANSD.

References

[1] R. I. Busarello, V. R. Ulbricht, L. M. Fadel, and A. V. de Freitas e Lopes, "Gamification approaches to learning and knowledge development: A theorical review," in *New Advances in Information Systems and Technologies*. Springer, 2016, pp. 1107–1116.

[2] A. Christopoulos and S. Mystakidis, "Gamification in Education," *Encyclopedia*, vol. 3, no. 4, pp. 1223–1243, Oct. 2023.

[3] Cockos, Justin Frankel, "Reaper," Cockos.

[4] G. Cousin, "Case Study Research," *Journal of Geography in Higher Education*, vol. 29, no. 3, pp. 421–427, Nov. 2005.

[5] Daniel J. Velleman and Gregory S. Warrington, "What to Expect in a Game of Memory," *The American Mathematical Monthly*, vol. 120, no. 9, p. 787, 2013.

[6] Ward. 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, no. 2, pp. 114–123, Feb. 2015.

[7] Z. Duan, C. Gupta, G. Percival, D. Grunberg, and Y. Wang, “SECCIMA: Singing and ear training for children with cochlear implants via a mobile application,” in *Proceedings of the 14th Sound and Music Computing Conference, Espoo, Finland*, 2017, pp. 5–8.

[8] Fan-Gang Zeng, S. Rebscher, W. Harrison, Xiaoan Sun, and Haihong Feng, “Cochlear Implants: System Design, Integration, and Evaluation,” *IEEE Reviews in Biomedical Engineering*, vol. 1, pp. 115–142, 2008.

[9] Firelight Technologies, “FMOD,” Firelight Technologies.

[10] M. D. Fletcher, “Can Haptic Stimulation Enhance Music Perception in Hearing-Impaired Listeners?” *Frontiers in Neuroscience*, vol. 15, p. 723877, Aug. 2021.

[11] C. D. Fuller, J. 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, p. 233121651876537, Jan. 2018.

[12] F. Ganis, A. Gulli, F. Fontana, and S. Serafin, “The Role of Haptics in Training and Games for Hearing-Impaired Individuals: A Systematic Review,” *Multimodal Technologies and Interaction*, vol. 8, no. 1, p. 1, Dec. 2023.

[13] F. Ganis, M. Vatti, and S. Serafin, “Tickle Tuner - Haptic Smartphone Cover for Cochlear Implant Users’ Musical Training,” in *Haptic and Audio Interaction Design*, C. Saitis, I. Farkhatdinov, and S. Papetti, Eds. Cham: Springer International Publishing, 2022, vol. 13417, pp. 14–24.

[14] K. Gfeller, V. Driscoll, M. Kenworthy, and T. Voorst Van, “Music Therapy for Preschool Cochlear Implant Recipients,” *Music Therapy Perspectives*, vol. 29, no. 1, pp. 39–49, Jan. 2011.

[15] Gunhyuk Park and Seungmoon Choi, “Perceptual space of amplitude-modulated vibrotactile stimuli,” in *2011 IEEE World Haptics Conference*. Istanbul: IEEE, Jun. 2011, pp. 59–64.

[16] K. Hansen and R. Hiraga, “The Effects of Musical Experience and Hearing Loss on Solving an Audio-Based Gaming Task,” *Applied Sciences*, vol. 7, no. 12, p. 1278, Dec. 2017.

[17] J. Jaime, I. Barbancho, C. Urdiales, L. J. Tardón, and A. M. Barbancho, "A new multiformat rhythm game for music tutoring," *Multimedia Tools and Applications*, vol. 75, no. 8, pp. 4349–4362, Apr. 2016.

[18] N. T. Jiam, M. T. Caldwell, and C. J. Limb, "What Does Music Sound Like for a Cochlear Implant User?" *Otology & Neurotology*, vol. 38, no. 8, pp. e240–e247, Sep. 2017.

[19] H. Lee and D. Müllensiefen, "The Timbre Perception Test (TPT): A new interactive musical assessment tool to measure timbre perception ability," *Attention, Perception, & Psychophysics*, vol. 82, no. 7, pp. 3658–3675, Oct. 2020.

[20] C. J. Limb and A. T. Roy, "Technological, biological, and acoustical constraints to music perception in cochlear implant users," *Hearing Research*, vol. 308, pp. 13–26, Feb. 2014.

[21] C. Y. Lo, V. Looi, W. F. Thompson, and C. M. McMahon, "Beyond Audition: Psychosocial Benefits of Music Training for Children With Hearing Loss," *Ear & Hearing*, vol. 43, no. 1, pp. 128–142, Jan. 2022.

[22] V. Looi, H. Mcdermott, C. M. McKay, and L. Hickson, "The effect of cochlear implantation on music perception by adults with usable preoperative acoustic hearing," *International Journal of Audiology*, vol. 47, pp. 257–268, 2008.

[23] O. Macherey and R. P. Carlyon, "Cochlear implants," *Current Biology*, vol. 24, no. 18, pp. R878–R884, Sep. 2014.

[24] S. Merchel and M. E. Altinsoy, "Psychophysical comparison of the auditory and tactile perception: A survey," *Journal on Multimodal User Interfaces*, vol. 14, pp. 271–283, 2020.

[25] 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, no. 19, p. 6575, Sep. 2021.

[26] F. A. Russo, P. Ammirante, and D. I. Fels, "Vibrotactile discrimination of musical timbre." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 38, no. 4, pp. 822–826, 2012.

[27] H. Sharp, Y. Rogers, and J. Preece, *Interaction Design : Beyond Human-Computer Interaction*, fifth edition ed. Indianapolis, IN: John Wiley & Sons, 2019.

[28] 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, no. 1, pp. 15–28, Feb. 2021.

[29] C. Spinuzzi, "The Methodology of Participatory Design," *Technical communication (Washington)*, vol. 52, no. 2, pp. 163–174, 2005.

[30] Unity Technologies, "Unity," Unity Technologies, 2021.

[31] O. B. Wie, J. V. K. Torkildsen, S. Schauber, T. Busch, and R. Litovsky, "Long-Term Language Development in Children With Early Simultaneous Bilateral Cochlear Implants," *Ear & Hearing*, vol. 41, no. 5, pp. 1294–1305, Sep. 2020.

[32] F.-G. Zeng, Y.-Y. Kong, H. J. Michalewski, and A. Starr, "Perceptual Consequences of Disrupted Auditory Nerve Activity," *Journal of Neurophysiology*, vol. 93, no. 6, pp. 3050–3063, Jun. 2005.

Paper D

Vibrotactile Teddy Bear: Enhancing Musical Experiences for Children with Cochlear Nerve Deficiency through a Vibrotactile Soft Toy

F. Ganis A. Adjorlu, K. L. Roslyng, L. M. Percy-Smith, S. Serafin

The article has been submitted to
The 22th Sound and Music Computing Conference Conference Proceedings,
2025.

© 2025, The Authors, under CC 4.0
The layout has been revised.

Original Publication

F. Ganis, A. Adjorlu, K. L. Roslyng, L. M. Percy-Smith and S. Serafin, “Vibrotactile Teddy Bear: Enhancing Musical Experiences for Children with Cochlear Nerve Deficiency through a Vibrotactile Soft Toy”, in *Proceedings of the 22nd Sound and Music Computing Conference*, Graz, July 2025, Graz, Austria, Jul. 2025, pp. 115-120. doi: 10.5281/zenodo.15841353.

Abstract

This study investigates the potential of a soft toy prototype that integrates vibrotactile feedback and sound to enhance musical experiences for children with hearing impairments, particularly those with auditory neuropathy spectrum disorder (ANSD) and cochlear nerve deficiency (CND). Traditional hearing aids (HAs) and cochlear implants (CIs) often fail to deliver the full spectrum of musical experiences, impacting children's development and quality of life. The soft toy provides vibrotactile feedback synchronized to music, allowing children to feel rhythm and harmonic content, creating an holistic musical experience.

The prototype, designed as a soft animal toy, includes a speaker and a vibrotactile motor controlled via Bluetooth. Two children with CND interacted with the toy during auditory-verbal therapy (AVT) sessions. Observations showed varying engagement levels based on age and developmental stage. The younger child required redirection, while the older child was more focused.

Findings suggest that vibrotactile feedback in AVT can enhance auditory and emotional experience, highlighting the need for personalized, age-appropriate approaches and further research.

D.1 Introduction

For some young children with hearing challenges (HC), significant barriers exist in accessing and enjoying music, affecting their overall development and quality of life. Traditional hearing aids (HAs) and cochlear implants (CIs) have made strides in improving speech comprehension, but they still fall short in delivering the full spectrum of musical experiences. Music, with its complex tones, rhythms, and emotional nuances, may be difficult for many HA and CI users to fully appreciate due to limitations in frequency range, dynamic range, and timbre recognition [6].

Research has shown that music training can significantly improve music perception skills and overall well-being for children with hearing impairments [3, 6, 10]. This is particularly relevant for children with auditory neuropathy spectrum disorder (ANSD) and cochlear nerve deficiency (CND), conditions characterized by an underdeveloped or absent cochlear nerve, which can lead to profound hearing loss. In particular, CND is often associated with very poor hearing performance, and cochlear implantation has been proven to be a viable improvement for children affected [1, 12]. Regardless of how auditory skills develop, the auditory signal provided is significantly weaker compared to a hearing system with a normal cochlear nerve. Consequently, the improvement must stem from the brain's ability to process this limited auditory signal from the CI. This explains why a longer period is needed, particularly for the development of speech perception [12]. Early intervention and personalized training are crucial for children with hearing loss, especially those with ANSD or CND. Thus, it is interesting to study how music can be enjoyed by children with conditions that disrupt neural activity and affect sound perception [13].

One technique that might lead to positive effects is the use of vibrotactile feedback to aid the hearing sense, which involves mapping audio signals to tactile vibrations to emphasize specific features such as pitch [4, 5]. By combining auditory and tactile stimuli, vibrotactile feedback can enhance sound awareness and music perception for children with hearing loss [9].

Auditory-verbal therapy (AVT) is an established tool for enhancing auditory skills and emotional well-being in children with hearing challenges. It has been shown that children who undergo a 3-year program improve emotional and behavioral problems, as well as social strengths, to the extent of being comparable with typically hearing peers [11]. Thus, by engaging multiple senses during AVT sessions, there is a possibility of obtaining a more holistic and enriching experience for individuals facing the most difficult hearing challenges. The choice of haptics over vision is driven by the AVT practice to prioritize sound: literature shows that vision tends to overpower other senses in CI recipients (the McGurk effect) [8]. By focusing on vibrotactile aids, we aimed to provide additional sensory input that supports auditory processing without taking over other senses.

To address these hurdles and facilitate the learning path of children with specific aetiologies, we propose a soft toy prototype that integrates haptic feedback and sound to enhance musical experiences. This toy is designed to provide vibrotactile feedback synchronized to music, allowing children to feel the rhythm and harmonic content of the music. By combining auditory and tactile stimuli, the toy aims to create a more immersive, multisensory, and enjoyable musical experience. We believe that technologies embedded in friendly and familiar objects such as toys can make the rehabilitation process more pleasant.

In this paper, we present the design of the prototype, the methodology of the study, the observations, and the discussion of the findings.

D.2 Prototype Description

The prototype is a soft toy shaped like a small animal, made of soft fabric and synthetic wool-like filling. The soft toy was modified by adding two *hook-and-loop fasteners* strips on the sides of a cut in the back, allowing access to the electronic circuit inside.

The electronic system provides sound and vibrotactile feedback, controllable via any music player capable of streaming sound through Bluetooth. The toy features a speaker located in the middle of the torso and a vibrotactile motor positioned in the front center of its body. The vibrotactile motor is a voice-coil (Tactuator BM1C from TactileLabs) that provides strong and broadband vibrations. It is powered by a rechargeable battery, which can be charged through a USB port. The toy can be turned on and off using a switch located on its back.

D.2.1 First Design Iteration

In the first iteration, we integrated a stereo low-power Bluetooth receiver, the Sunrom M18[1], with a 3W stereo amplifier. The amplifier's outputs were connected to both a vibrotactil actuator and a mini 8 Ohm loudspeaker. Power was supplied by three 1.5V AA batteries in series configuration, providing a total of 4.5V. With this configuration, the loudspeaker and the actuator each received one of the two stereo output channels. Without proper configuration of the streaming device, the system would provide separate audio streams to the two actuators, according to how the music was mixed. Figure D.1 illustrates the electronic circuit. The circuitry and loudspeaker were housed inside a cardboard box to protect the fragile components from impacts, while the vibrotactile actuator, enclosed in a 3D-printed case, was positioned in the front area of the stuffed animal's belly to deliver strong vibrations when the toy is hugged.

---

[1] Sunrom Bluetooth Transfer Module MX8, last accessed December 17, 2025

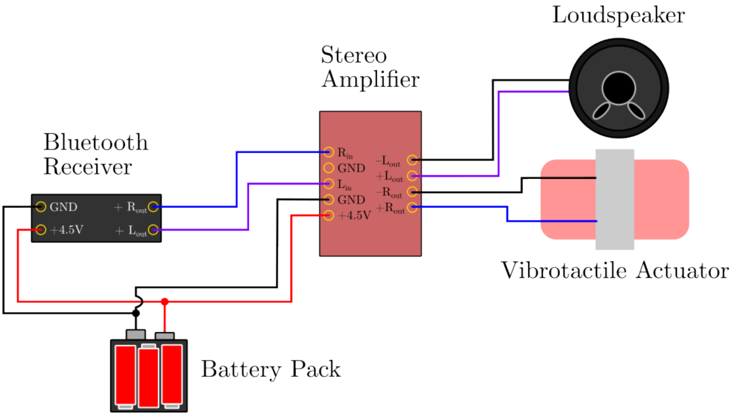


Fig. D.1: First iteration of the soft toy's circuitry

D.2.2 Second Design Iteration

In the second iteration, we integrated the Bluetooth receiver and the amplifier into a single module that also includes a charging circuit for handling a rechargeable LiPo battery. The board used is an ESP-32 A1S AudioKit, which we programmed to behave like a Bluetooth loudspeaker. It receives, decodes, and plays audio to both the loudspeaker and the haptic actuator. The new circuitry is shown in Figure D.2. Additionally, we designed a simple 3D-printed enclosure to accommodate the entire system within the stuffed toy and protect the components from impacts. The software also allows for a dual-mono configuration, providing the exact same signal to both outputs, if needed.

The enclosure ensures the possibility to place the components in any medium-sized soft toy. Thus, the toy can be easily customized to fit the child's preferences. In addition, the programmable board allows for further customization such as real-time digital signal processing techniques to process sound and vibrations.

D.3 Methodology

In this section, we will describe how the study was conducted, including the participants, the procedure, and the materials used.

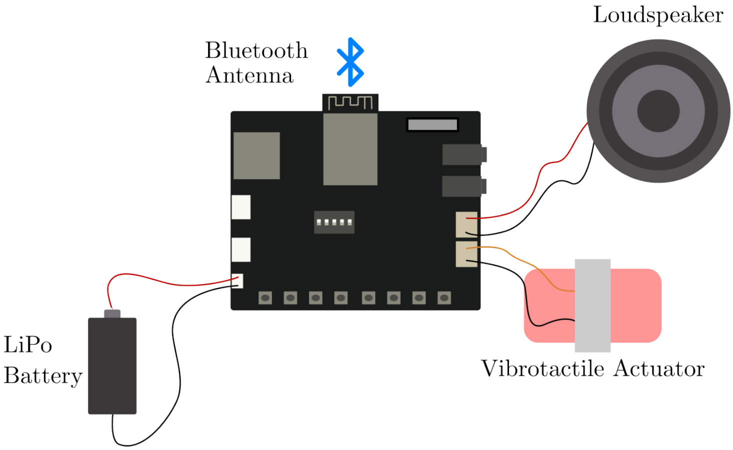


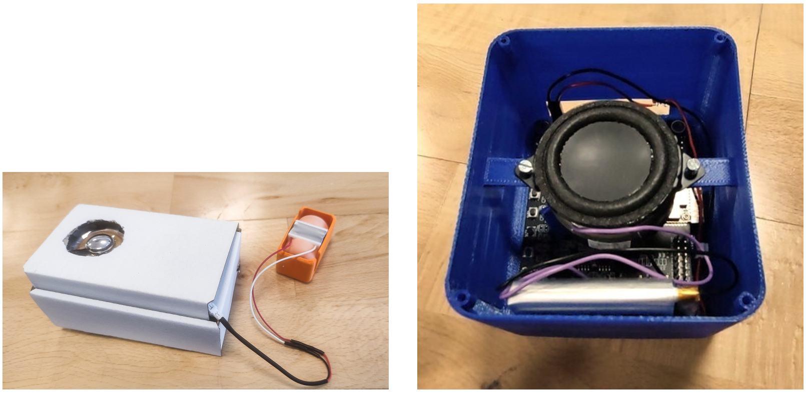
Fig. D.2: Second iteration of the soft toy's circuitry.

D.3.1 Participants

We included two young children aged 2 years and 10 months, and 4 years and 5 months. From now on, the younger participant will be referred to as *Little Star*, and the older as *Bright Star*. Both children were diagnosed with hearing loss and cochlear nerve deficiency and were recipients of bilateral CIs. In the case of Little Star, her parents struggled to keep the CIs on, and as a result, they were not used full-time on a daily basis. At the time of the study, the two children were in their second year of AV therapy. They were recruited through the Copenhagen Hearing and Balance Center (CHBC) at Rigshospitalet, the largest hospital in Denmark. The study was approved by the ethical board of Rigshospitalet through the speech and language pathologists involved in the study. The participants were accompanied by their parents, who were informed about the study's purpose and procedures. The parents provided verbal consent, allowing their children to play and interact with the prototype during the AVT session.

D.3.2 Procedure

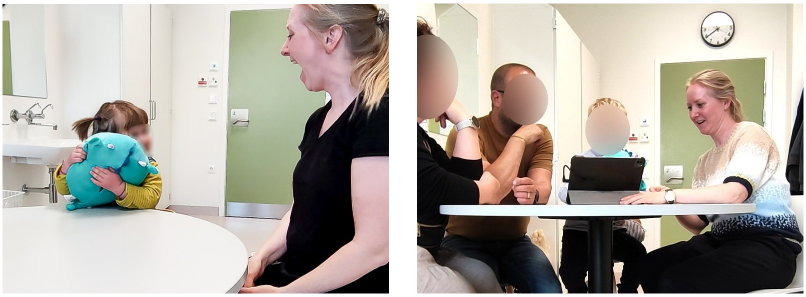
We arranged two sessions at the CHBC during their individual AVT program. This choice was made to include the AV therapist in the session, in order to provide a guided experience for the child. Moreover, the children had the chance to be introduced to new technology with a trusted and familiar person.



(a) (b)

Fig. D.3: The first (a) and second (b) iterations of the soft toy's protective enclosure. In figure (a), it is possible to see the 3D-printed case containing the vibrotactile actuator.

Additionally, they were in the presence of their parents, as is typical for all AV sessions. The sessions were recorded using a video camera, and the children's reactions were noted by the researchers.



(a) (b)

Fig. D.4: Children interact with the prototype under the supervision of the Auditory-Verbal therapist.

Participant Little Star

In the guided exploration at the hospital, the child was introduced to the soft toy and allowed to play with it for 15 minutes during the beginning of the AV session. During this time, the child was encouraged to interact with the toy while music, typically used for therapeutic activities such as *"Fem Små Aber"*, played through its loudspeaker and vibrotactile actuator. The child was invited to listen to the song, touch the soft toy, and feel the vibrations.

In a later iteration, we provided the child with the soft toy to take home for one and a half months. The child was encouraged to play with the toy and listen to music through it. The parents were instructed to use the teddy bear as a tool to engage the child in musical activities and to observe the child's reactions to the toy. At the end of this period, the parents had an informal conversation with the researchers to report their observations and experiences with the prototype.

Participant Bright Star

A similar approach to the one used with Little Star was adopted with Bright Star. The child was introduced to the soft toy during the AV session. This time, due to the older age of the participant, the AV therapist instructed the child to play a simple video game based on the 6 Ling sound identification task. The game, developed by one of the authors, was designed to train the child's listening skills in distinguishing sounds like *"sh"* and *"ssss"*. The game was controlled through an iPad placed in front of Bright Star, while the child was hugging the soft toy. Bright Star was encouraged to listen to the sounds, touch the soft toy, and feel the vibrations to help him distinguish the sounds. Once the sound was assessed, the participant had to report the answer by touching the corresponding image on the iPad.

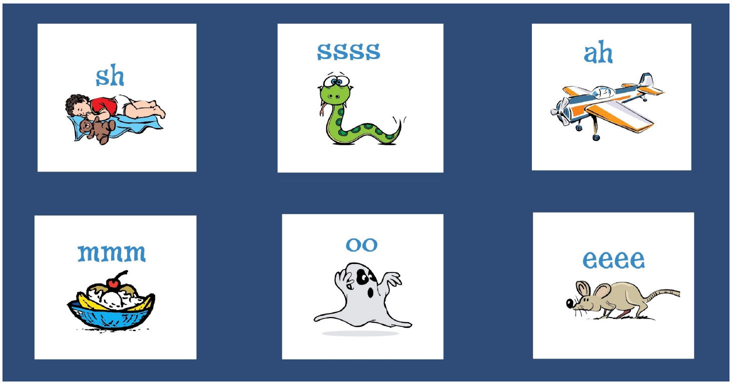


Fig. D.5: Interface of the video game for 6 Ling sound identification task.

D.4 Observations and Results

In this section, we present the observations and results of the study, including the children's reactions to the prototype and the parents' feedback. During the first interaction, Little Star showed curiosity and interest in the soft toy. She was eager to touch and hug it, and seemed to enjoy the vibrations. She was particularly drawn to the toy's belly, where the vibrotactile actuator was located. She laid her head on the toy's belly and showed surprise. Over the session, her interest expanded to the entire toy, and she began to play with it as if it were a regular stuffed animal.

The therapist, in accordance with the session's tasks, used the toy to engage the child in musical activities, such as listening to songs and playing simple musical games involving wooden animals. In this specific activity, the soft toy served merely as a loudspeaker, as Little Star was supposed to interact with the wooden toys. Overall, she seemed to enjoy the activities and showed a positive reaction to the toy, even though her focus was not always on it. After the initial interest and surprise for the novel object, she seemed to lose interest in the soft toy, and the therapist had to redirect her attention to it. Due to her young age and possibly limited daily use of her CIs, Little Star required guidance during the interaction, and the therapist had to redirect her attention several times.

In contrast, Bright Star, the older participant, showed a different reaction to the prototype. Bright Star was more focused and diligently held the toy on the chest while interacting with the iPad. This allowed him to feel the vibrations and hear the sounds while playing the game. Bright Star seemed engaged in the task, and the therapist did not have to redirect attention to the toy, as it was an integral part of the activity. The child showed a positive reaction to the entire experience.

The parents provided positive feedback about the prototype, noting that their children enjoyed interacting with it. Little Star's parents agreed to take the teddy bear home to allow her more opportunities to play with it and engage with the sound features. In particular, Little Star's parents reported that their child played a lot with the teddy bear during the period at home, also sharing it with her brother. At the time of return, the toy had been broken due to extensive and intensive use.

D.5 Discussion

In this small study, we observed a significant difference in the reactions to the soft toy from the two children with CND. Little Star, the younger participant, showed initial interest and curiosity in the toy but quickly lost focus and needed redirection from the therapist. Bright Star, the older participant, was more focused and engaged with it, using the teddy bear as a tool to hear and feel the sound stimuli. These differences can be attributed to the children's age and developmental stage, as well as their familiarity with technology and toys. Additionally, age highlights differences in attention span, comprehension, and ability to follow instructions between the two children.

These observations suggest that toy design, as well as the activities connected to it, should be tailored to different developmental stages to ensure engagement and effectiveness in AVT sessions for children with CND. The vibrotactile feedback might assume different roles and meanings according to the participants' age. For instance, the vibrotactile feedback might be more engaging for older children who can understand the connection between sound and vibration, perceiving details and nuances, while younger children might benefit from it as a confirmation of sound presence. Furthermore, familiarity with the toy might change the engagement over time, as the child becomes more accustomed to it and its features. This suggests that the toy should be introduced gradually, allowing the child time to explore and interact with it at their own pace. Additionally, due to the rapid development in children, the design needs to be adapted to their specific needs at the moment and be capable of modification and upgrades over time. This might become an expensive issue for toys that require sophisticated technology.

From the literature, we know that multisensory integration between sound and vibrotactile feedback might be beneficial for CI recipients, as it conveys additional information that can support the listening experience [2, 7]. Further research should explore different ways to integrate and map vibrotactile feedback in AVT practices. Mappings have been shown to be crucial in the effectiveness of vibrotactile feedback, as they can highlight specific features of the sound [4, 5].

Another consideration is the integration of the prototype into AVT practices. The therapist's role is crucial in facilitating effective use of the device, as they can guide the child in the interaction and ensure that the toy is used correctly. The therapist can also provide feedback to the parents on how to use the toy at home and integrate it into the child's daily play routine. This requires a plan from the therapist on how to design the training interaction and find suitable activities that can easily include the use of a soft toy that provides vibrotactile feedback. The clinical integration of such a device is an interesting topic as well. The design process should take into account the challenges brought by the use of equipment in a clinical environment, considering safety, sanitization, and ease of use as core characteristics. Ultimately, one of the core strengths of AVT lies in its emphasis on guiding and coaching parents. Every session should include strategies for transferring activities into the home environment. The overarching goal of AVT is to ensure that all training is functional and applicable in everyday listening situations.

From a clinical perspective, engaging children with CND in AVT sessions can be challenging due to a limited auditory access and limited attention span compared to children with a regular sensorineural hearing loss. Integrating both the auditory and tactile sense in the therapy showed to be very positive and more engaging for the children. When demonstrating the toy we experienced that both children showed interest in the listening task and maintained focus when we kept the activity short and implemented toys relevant to their developmental stage in language and play level.

Conducting studies with a larger and more diverse group of participants would help validate the findings and explore the toy's effectiveness across different age groups and hearing conditions. Implementing longitudinal studies would allow researchers to assess the long-term impact of the toy on auditory and multisensory development, as well as its integration into daily routines. However, the group of children with CND is small and very heterogeneous, and vibrotactile feedback is not necessary for the vast majority of children with hearing challenges who undergo the AVT program.

The role of vibrotactile feedback in AVT is still underexplored, as most newer generations of children with CIs acquire hearing capabilities comparable to their typically hearing peers. However, the contribution of vibrotactile feedback might be beneficial for children who present more hearing challenges, as seen with our two participants.

D.5.1 Limitations of the Study

This study has several limitations that should be considered when interpreting the results. Firstly, the small sample size of only two participants limits the generalizability of the findings. A larger sample size would provide more robust data and allow for more comprehensive conclusions. Secondly, the study was conducted over a relatively short period, with limited interaction time with the prototype. Longer-term studies are needed to assess the sustained impact and effectiveness of the toy.

Additionally, the two participants were at different developmental stages, which may have influenced their interactions with the toy and made comparisons difficult. This is mainly due to the rarity of the hearing aetologies we took into consideration during the recruitment process. Future studies should consider a more homogeneous group or cover a larger sample for developmental variability. The activities conducted during the study were specific to the AVT sessions and may not encompass the full range of potential uses for the toy. Exploring a wider variety of activities could provide a more comprehensive understanding of the toy's effectiveness.

D.5.2 Future Work

Based on the findings and limitations of this study, directions for future research are proposed.

Investigating ways to customize the toy to better suit individual preferences and needs, such as independently adjustable vibration intensity and different types of vibrotactile feedback, could enhance its effectiveness. Technological advancements, such as integrating more sophisticated sensors, enhancing the quality of vibrotactile feedback, and developing interactive features that respond to the child's actions, should also be explored.

Studying the practical aspects of integrating the toy into clinical settings, including safety, sanitization, and ease of use, is crucial to ensure it meets the standards required for therapeutic devices. Developing training programs for parents and therapists to effectively use the toy and integrate it into AVT practices would ensure consistent and beneficial use.

D.6 Conclusion

In this study, we explored the potential of a soft toy prototype that integrates haptic feedback and sound to enhance musical experiences for children with hearing challenges. The prototype was designed to provide vibrotactile feedback synchronized to music, allowing children to feel the rhythm and harmonic content of the music. Our observations indicated that the toy was well-received by the children, with varying levels of engagement and interaction based on their age and developmental stage.

The findings suggest that the integration of vibrotactile feedback in AVT sessions is a viable path for making the sessions more holistic and engaging for children with CND. However, the study also highlighted the need for personalized and age-appropriate approaches to maximize the effectiveness of such interventions.

Future research should focus on a wider range of activities and contexts. Additionally, technological advancements and customization options should be investigated to better suit individual preferences and needs. The practical aspects of integrating the toy into clinical settings and developing training programs for parents and therapists are also important areas for further exploration.

Overall, this study provides a preliminary foundation for the development of multisensory tools that can support the auditory and emotional development of children with specific hearing aetiologies.

References

[1] A. Alahmadi, Y. Abdelsamad, N. A. AlAmari, M. Y. Alyousef, M. Al-Momani, F. N. Altamimi, S. F. Alhabib, and A. Hagr, “Hearing implants in pediatrics with cochlear nerve deficiency: An updated systematic review,” *European Archives of Oto-Rhino-Laryngology*, Dec. 2024.

[2] M. D. Fletcher, “Can Haptic Stimulation Enhance Music Perception in Hearing-Impaired Listeners?” *Frontiers in Neuroscience*, vol. 15, p. 723877, Aug. 2021.

[3] C. D. Fuller, J. 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, p. 233121651876537, Jan. 2018.

[4] F. Ganis, M. Vatti, and S. Serafin, “Tickle Tuner - Haptic Smartphone Cover for Cochlear Implant Users’ Musical Training,” in *Haptic and Audio Interaction Design*, C. Saitis, I. Farkhatdinov, and S. Papetti, Eds. Cham: Springer International Publishing, 2022, vol. 13417, pp. 14–24.

[5] Gunhyuk Park and Seungmoon Choi, “Perceptual space of amplitude-modulated vibrotactile stimuli,” in *2011 IEEE World Haptics Conference*. Istanbul: IEEE, Jun. 2011, pp. 59–64.

[6] N. T. Jiam, M. T. Caldwell, and C. J. Limb, “What Does Music Sound Like for a Cochlear Implant User?” *Otology & Neurotology*, vol. 38, no. 8, pp. e240–e247, Sep. 2017.

[7] 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, p. 1145, Oct. 2019.

[8] J. Rouger, B. Fraysse, O. Deguine, and P. Barone, “McGurk effects in cochlear-implanted deaf subjects,” *Brain Research*, vol. 1188, pp. 87–99, Jan. 2008.

[9] F. A. Russo, P. Ammirante, and D. I. Fels, “Vibrotactile discrimination of musical timbre.” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 38, no. 4, pp. 822–826, 2012.

[10] 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, no. 1, pp. 15–28, Feb. 2021.

[11] S. Wischmann, C. F. Samar, M. K. Hestbæk, S. Serafin, P. Cayé-Thomasen, and L. Percy-Smith, “Quality Assurance of a Hospital-Based Auditory Verbal Intervention for Children with Hearing Loss,” *Journal of Clinical Medicine*, vol. 14, no. 4, p. 1098, Feb. 2025.

[12] N. M. Young and K. Iler Kirk, Eds., *Pediatric Cochlear Implantation: Learning and the Brain*. Cham: Springer Nature Switzerland, 2024.

[13] F.-G. Zeng, Y.-Y. Kong, H. J. Michalewski, and A. Starr, “Perceptual Consequences of Disrupted Auditory Nerve Activity,” *Journal of Neurophysi-*

ology, vol. 93, no. 6, pp. 3050–3063, Jun. 2005.

Paper E

SoundCubes - Preliminary Evaluation of a Spatial Hearing Training Tool and a Sound Localization Test for Children with Hearing Loss

F. Ganis, A. A. Kressner, L. Percy-Smith, A. Adjorlu, and S. Serafin

The article has been submitted to
International Journal of Human-Computer Interaction, 2025.

The article is under revision.

© 2025, The Authors, under CC 4.0

The layout has been revised.

Original Publication

–

Paper F

EmotiCubes - A Training Tool for Vocal Emotion Recognition in Children with Hearing Loss

F. Ganis, L. Rachman, D. Başkent, and S. Serafin

The article has been submitted to
Journal of Clinical Medicine - Special Issue: *Pediatric Hearing Loss: Advances in Early Detection, Intervention, and Family-Centered Care*, 2025.

The article is under revision.

© 2025, The Authors, under license to MDPI, Basel, Switzerland - CC 4.0
The layout has been revised.

Original Publication

–

Part III

Additional Papers

Paper G

Enhanced Neural Phase Locking Through Audio-Tactile Stimulation

M. Jagt, F. Ganis, and S. Serafin

The article has been submitted to
Frontiers in Neuroscience, 2024.

© 2024, The Authors, under CC 4.0
The layout has been revised.

Original Publication

M. Jagt, F. Ganis, and S. Serafin, “Enhanced neural phase locking through audio-tactile stimulation,” Front. Neurosci., vol. 18, p. 1425398, Oct. 2024, doi: 10.3389/fnins.2024.1425398.

Abstract

Numerous studies have underscored the close relationship between the auditory and vibrotactile modality. For instance, in the peripheral structures of both modalities, afferent nerve fibers synchronize their activity to the external sensory stimulus, thereby providing a temporal code linked to pitch processing. The Frequency Following Response is a neurological measure that captures this phase locking activity in response to auditory stimuli. In our study, we investigated whether this neural signal is influenced by the simultaneous presentation of a vibrotactile stimulus. Accordingly, our findings revealed a significant increase in phase locking to the fundamental frequency of a speech stimulus, while no such effects were observed at harmonic frequencies. Since phase locking to the fundamental frequency has been associated with pitch perceptual capabilities, our results suggests that audio-tactile stimulation might improve pitch perception in human subjects.

G.1 Introduction

The combination of both psychophysical and neurophysiological methodologies have aimed to provide a comprehensive understanding of how physical stimuli are transformed into perceptual experiences through neural processes. This paper focuses on the perception of pitch. Initially, we will outline how pitch is manifested within the auditory system, followed by a discussion on its analogous representation via vibrotactile stimulation. Subsequently, existing interactions between auditory and vibrotactile pitch processing are explored, establishing a solid foundation prior to introducing our research approach, which is presented in the end of this section.

The psychophysical study of pitch perception has a long history, stretching back to the period where Pythagoras investigated the connection between the length of a plucked string and its excitation frequency. Meanwhile it is understood that sound essentially involves airborne pressure waves that are commonly characterized by three physical attributes: frequency, amplitude and time/phase [30, 61]. Pitch, the degree to which a sound is perceived as ”high” or ”low”, is arguably the most critical perceptual feature of sound. This attribute plays an essential role in musical appraisal, speech and the identification of sources [17]. In the case of simple pure tones representing sinusoidal waves with a singular frequency, pitch directly correlates to its frequency [61]. The human auditory system is capable of detecting these pure tones across a frequency spectrum from 20 Hz to 20,000 Hz, albeit with varying degrees of sensitivity across this range. As illustrated in Fig. G.2A, the hearing has lowest detection thresholds to frequencies between approximately 300 Hz - 7000 Hz. For more complex stimuli, the perception of pitch is not always as straightforward as with pure tones. Many everyday sounds involve harmonic complexes which are characterized by a fundamental frequency, the first harmonic, alongside additional frequency components that occur at integer multiples of this fundamental frequency, known as higher harmonics. Generally, the pitch can be directly derived from the fundamental frequency, however this is not always the case as illustrated with the missing-fundamental stimulus [55]. A more elaborate discussion on the different conditions that lead to pitch perception can be found in [61].

Additionally, over the past century, neurophysiological research into pitch perception has progressed our understanding of the underlying neural mechanisms. Today, it is commonly believed that much of the pitch information is contained within the temporal code of spiking neurons [38]. This is well represented in the auditory periphery where the auditory nerve fibers synchronize their firing (i.e., phase lock) to the frequency of the vibration at each place on the basilar membrane, which is sensed through mechanoreceptors (i.e., the inner hair cells) in the cochlea (Fig. G.2B). The upper limit of this phase locking behavior is about 5 kHz, though the exact number is still under debate [53]. Subsequently, pitch could be encoded by synchronization to either the stimulus' temporal fine structure ($\leq$ 5 kHz) or the temporal envelope. Towards higher stages in the ascending auditory pathway, the upper limit of phase locking progressively decreases. The frequency following response (FFR) is a non-invasive electrophysiological measure that captures phase locking activity ($\leq$ 2 kHz) within the brain stem [18] (Fig. G.1). Commonly, a ”vertical” one-channel electrode configuration is used with the active electrode on the top head (Cz), the reference electrode on the earlobe (A1/A2) and the ground electrode on the forehead (Fpz). Although this configuration primarily targets the sum of synchronized neural activity in the inferior colliculus, it is believed that the FFR in reality represents multiple neural sources that all operate in concert with each other [5]. Interestingly, differences in FFR strength (i.e., phase locking efficacy to the fundamental frequency) has been correlated with differences in pitch perception. For example, the work of Krishnan et al. has demonstrated that tone language speakers (e.g., Chinese, Thai) have an enhanced FFR compared to non-tone language speakers (e.g., English) [19, 20]. Additionally, improved pitch discrimination following short-term training has been shown to strengthen the FFR [4]. These results suggest that the FFR likely bears pitch relevant information, despite not directly depicting the precise mechanism of pitch extraction. Moreover, besides pitch, it has to be emphasized that the FFR also contains information relevant to other aspects of auditory processing. An extensive overview of the latter is presented in [22]. Ascending further towards the cortex, the phase locking ability diminishes drastically. As a consequence, the temporal code for pitch is likely transformed into another type of encoding. Less is understood about these higher order representations, while current research mainly focuses on finding cortical ”pitch centers” using different techniques (e.g., fMRI, EEG) [38].

To this point, the discussion on pitch perception has only focused on its manifestation within the auditory system. However, a comparable phenomenon exist for the somatosensory system. More precisely, it has been demonstrated that vibrotactile stimuli, referring to the detection of vibrations on the skin, can also be perceived as ”high” or ”low”, akin to auditory stimuli [39]. A number of psychophysical analogies can be drawn between these two modalities [30]. That is, most physical properties of airborne vibrations (e.g., frequency, amplitude) find their direct counterpart in skin vibrations [56]. A significant distinction, however, is that vibrotactile stimuli can be perceived at different locations on the body. Here, regions with high (vibro)tactile acuity (i.e., hands or fingertips) have been characterized by a dense concentration of sensory receptors [3, 40]. Many research efforts have been focused on the fingertip, and evidence seems to indicate that the perception of vibrotactile pitch for simple sinusoids also greatly depends on the vibration’s frequency, while possibly being modulated by its amplitude [39]. Moreover, similar to the auditory system, a V-shaped sensitivity curve has been computed for sinusoidal skin vibrations (Fig. G.2A). In comparison, the frequency discrimination of vibrotactile stimuli exhibits a significantly lower resolution, with Weber fractions ranging from 0.2-0.3 in contrast to 0.003 as shown for the auditory system [45]. Furthermore, while the frequency spectrum is somewhat restricted to a maximum of 1000 Hz, it is perceptually classified into two distinct categories: a ’flutter’ range up to 50 Hz and ’smooth vibrations’ for frequencies above 50 Hz, mirroring the auditory system. In both senses, the flutter range is characterized by the absence of pitch where individual cycles are discernible, opposed to higher frequencies that generate the percept of a continuous signal with an identifiable pitch. Early work by Mountcastle et al. has linked these perceptual findings with neurophysiological studies [32, 33]. Accordingly, it has been shown that primarily two types of mechanoreceptors are involved in the detection of vibrations within the somatosensory periphery. Specifically, the Meissner's corpuscles mainly account for the detection of the flutter range vibrations, while the Pacinian corpuscles are predominantly involved in the detection of the smooth vibrations. Intriguingly, these peripheral structures exhibit phase locking behavior analogous to that observed in the auditory periphery (Fig. G.2B). That is, the activity of afferent nerve fibers innervating these mechanoreceptors periodically entrain to the frequency of the vibrating stimulus, enabling synchronization to skin vibrations across the entire spectrum up to 1000 Hz [45]. Hence, this temporal code also carries stimulus information, and therefore potentially bears relevant information to vibrotactile pitch processing.

Besides these psychophysical and neurophysiological parallels in pitch perception across both sensory modalities, there is another empirical argument that further supports the existence of a close relationship. Particularly in the music domain, the sensation of sound can be coupled with vibrations on the skin. This becomes apparent in the setting of a live concert where vibrotactile stimuli complement auditory cues to enhance the musical journey for both the audience as well as the performers [29]. Collectively, these observations motivate the investigation into the possible interactions between the auditory and vibrotactile (i.e., audio-tactile) modality, especially with respect to pitch processing. This view is consistent with the recent trend of the last few decades where the field of sensory processing has shifted from scientific research investigating each sensory modality in isolation, to the perspective of a highly interconnected, interactive multisensory network [52].

Over the past two decades, a considerable body of research has been dedicated to investigating the impact of audio-tactile interactions on human perception. Consequently, it has been shown that audio-tactile stimulation improves reaction speed [51], stimulus detection [13] and enhances the perceived loudness [13, 47]. Interestingly, these interactions seem to depend on the relative frequencies between both modalities [57, 58], with the most pronounced effects observed when the frequencies of auditory and vibrotactile stimuli closely overlap. This observation also extends to audio-tactile influences on pitch perception. In subsequent sections, we primarily concentrate on the examination of audio-tactile interactions within the supra-flutter range of 50 Hz to 1000 Hz because: i) it contains a distinct and identifiable perception of pitch; ii) it represents the overlapping region where frequencies are perceptible to both the auditory and somatosensory systems. Accordingly, the work of Yau et al. has revealed that auditory and vibrotactile stimuli reciprocally bias each other, demonstrating that the concurrent presentation of a vibrotactile distractor significantly influenced the perception of auditory pitch, and vice versa [59, 60]. Concretely, using simple sinusoids, the pitch frequency of a stimulus in one modality was pulled towards the frequency of the distractor modality. Subsequent research following a crossmodal adaptation design further substantiated these observations, demonstrating how vibrotactile pitch perception is influenced by auditory stimuli using band pass noise [7] and sinusoidal sweeps [6] respectively. A key finding was that these results were only obtained when the frequencies of both modalities were sufficiently aligned. Collectively, this further implicates an intimate relationship between both modalities, and suggests shared interactive neural mechanisms regarding frequency processing.

Subsequent research efforts have been devoted to elucidate such neural foundations that facilitate these observed audio-tactile effects. In this regard, a significant number of studies in human subjects have utilized the fMRI neuroimaging technique, while specifically targeting the cerebral cortex. Accordingly, it has been shown that areas traditionally associated with a single sensory modality are susceptible to crossmodal influences. That is, auditory stimulation revealed robust and frequency-specific responses within the traditionally defined somatosensory regions of the parietal lobe [36]. In the same way, vibrotactile stimuli has shown to activate areas typically associated with auditory processing within the temporal lobe [35, 46]. Regarding the latter, it was specifically shown that 100 Hz sinusoidal vibrations selectively impacted the left auditory cortex, an area thought to have a specialized role in detecting fundamental frequencies [35]. Consistent with these results, a more recent study revealed similar overlapping activation regions and highlighted their involvement in frequency specific processing [41]. While these results tentatively indicate the existence of shared neural populations regarding frequency processing, the direct link of these observations to the psychophysical audio-tactile interactions of pitch perception remains rather obscure. This challenge can be attributed in part to the inherent constraints of fMRI. While fMRI has proven effective for spatially identifying brain regions engaged in frequency processing, its relatively limited temporal resolution and the sluggishness of the blood oxygenation level dependent (BOLD) signal complicate the task of directly linking perceptual outcomes with underlying neural processing activities. Moreover, emerging evidence from animal studies has revealed extensive audio-tactile interactions within subcortical regions, involving both ascending and descending projections [26]. Consequently, the activities observed in cortical regions might merely reflect the crossmodal influences originating from these lower-level neural structures. This observation would not come completely unexpected, considering the analogous temporal coding mechanisms present in subcortical regions across both modalities.

To summarize, the pronounced similarities between the auditory and vibrotactile modalities concerning pitch processing, in combination with substantial perceptual evidence of audio-tactile interactions, collectively suggest the existence of shared neural pathways for frequency processing. Efforts to investigate such putative networks in human subjects, particularly focusing on the cerebral cortex, have yet to yield compelling evidence. Meanwhile, research in animal models indicates the presence of significant audio-tactile interactions at subcortical stages. Hence, it may be beneficial to explore analogous regions in human participants. Accordingly, the principal aim of this study is to address the latter proposition by exploring the following questions: i) is it possible to furnish evidence supporting the presence of audio-tactile interactions within subcortical structures in human subjects?; ii) should such evidence emerge, what would be its implications for the temporal coding mechanism? Our hypothesis states that vibrotactile stimuli complement auditory stimuli and improve phase locking acuity. Considering the efficacy of FFR in capturing subcortical phase locking activity to auditory stimuli, our research seeks to build upon this by incorporating a concurrent vibrotactile stimulus. Accordingly, this enables the investigation of how the synchronized neural activity in the brain stem is potentially modulated by the vibrotactile modality.

G.2 Material and Methods

Previous studies have predominantly utilized basic and coarse audio-tactile stimuli, such as simple sinusoidal waves, often delivered through insert earphones and applied to a single finger digit. This study intends to adopt a more natural stimulation paradigm. To achieve this, we choose to utilize real-world speech stimuli, presented through insert earphones, alongside an ergonomic vibrotactile controller that stimulates the entire hand. Most settings concerning the auditory stimulation and FFR recording are directly derived from [21]. Details are presented below.

G.2.1 Participants

The dataset consisted of FFRs recorded from 22 healthy young adults (age: 28 ± 6) of which 11 are female. None of the participants had a history of neurological dysfunction or a reported hearing loss and all gave written consent to participate on voluntary basis.

G.2.2 Stimulus Selection

The selected stimulus was identical for both the auditory and vibrotactile modality, and involved the 40-ms /*da*/ speech syllable. This /*da*/ is a generated speech sound [16] with five formants. The syllable is characterized by an initial noise burst followed by a formant transition between the consonant and the vowel. More specifically, during the 40-ms, the fundamental frequency (F_0) and the first three formants (F1, F2, F3) shift linearly (F_0: 103-125 Hz, F1: 220-720 Hz, F2: 1700-1240 Hz, F3: 2580-2599 Hz). The formants F4 (3600 Hz) and F5 (4500 Hz) however remain constant.

This specific stimulus was chosen for the following reasons. First, /*da*/ is a relatively universal syllable found in the phonetic inventories of most European languages [27]. Second, conventional FFR studies focusing on the auditory modality only, showed that this sound stimulus elicits a clear and replicable response (Fig. G.3B), thereby providing a normative database [49]. Last, the fundamental frequency, as well as the first formant, of /*da*/ lie well within the perceivable frequency range of the vibrotactile modality and are below subcortical phase-locking limits.

G.2.3 Stimulus Presentation

Vibrotactile stimulation was provided through a DualSense Controller (Sony Corporation, Tokyo, Japan) via USB connection. The reason for choosing this device is two-fold. First, the controller has an ergonomic casing that is rigorously tested to provide comfortable vibrations to a broad audience. Second, on both the left and right hand side, it has two built-in voice-coil actuators (ref. 622008, Foster Electric Company, Tokyo, Japan). These type of actuators are frequently employed for transmitting musical information via vibrotactile stimuli [42]. For example, audio signals can be utilized to drive them with minimal or no further signal processing required, where pitch and loudness directly maps to the frequency and amplitude of the vibration respectively [37]. Corollary, it is posited that such actuators are also adequate for conveying speech stimuli, including the /*da*/ used in this study. Nonetheless, given that the DualSense controller is primarily designed for entertainment purposes, further validation was necessary to evaluate its appropriateness for this research (see section G.2.6). After the successful preliminary validation, vibrotactile signals were delivered bimanually and calibrated at 0.85 m/s^2 peak-to-peak.

Auditory stimuli were presented through an IP30 insert earphone (RadioEar, Middelfart, Denmark) via the Steinberg UR44C audio interface (Steinberg Media Technologies GmbH, Hamburg, Germany). These earphones are electromagnetically shielded in order to minimize the stimulus artifact (i.e., the electrical signal produced by a transducer that could contaminate the FFR recording). Subsequently, auditory signals were delivered monoaurally in the right ear at 80 dB SPL. An earplug was deeply inserted in the left ear of the participant to minimize interference from other potential sound sources (e.g., audible vibrations of the DualSense Controller).

The presentation of both sensory stimuli followed a block paradigm (Fig. G.3A). Within each block, 100 /*da*/ trials were presented in an alternating polarity (i.e., 50 positive and 50 negative polarity) at a rate of 10.9 Hz. In total 40 blocks were presented for a single session, equating to $40 \times 100 = 4000$ /*da*/ trials per session. Importantly, the inclusion of a vibrotactile block was randomized with probability 0.5. Hence, only 20 blocks involved the simultaneous presentation of both sensory modalities where temporal synchronization was accounted for by setting adequate latencies. That is, we ensured a simultaneous peripheral stimulation between both modalities. Besides the audio-tactile condition, the remaining 20 blocks comprise only the auditory condition. The random ordering of consecutive conditions aimed to minimize potential confounding factors such as the expectancy effect. After the completion of a single session, a brief intermission of $\sim$15 seconds was placed before proceeding to the next session. Each participant was exposed to a total of 3 repetitions, equating to a cumulative amount of $4000 \times 3 = 12000$ /*da*/ trials.

G.2.4 Data Collection

The FFR was collected with the Eclipse EP15 system (Interacoustics, Middelfart, Denmark) using disposable Ag/AgCl gel electrodes and the EPA preamplifier (Interacoustics, Middelfart, Denmark). The electrodes were applied in a vertical montage with the active, reference and ground electrode located at Cz (top head), A2 (right earlobe) and Fpz (forehead) respectively (Fig. G.3A). The electrode impedance was kept below 5 kΩ. The acquired signal was subsequently recorded at 48 kHz (sampling rate) with the Steinberg UR44C audio interface (Steinberg Media Technologies GmbH, Hamburg, Germany). This audio interface was thus employed for both the delivery of auditory stimuli and the recording of the FFR. This arrangement aided in facilitating precise alignment between stimulation and recording, coordinated through a custom made Python script. Furthermore, during the experiment the participant was instructed to sit relaxed on a comfortable chair while watching a mute movie of choice (e.g., Charlie Chaplin). This visual distractor controls for attention and aimed to minimize head movements of the participant while promoting relaxation. Moreover, since the FFR signal is on the order of nanovolts, the experiments took place in an anechoic chamber that was enclosed in a Faraday cage to reduce (electrical) noise (Fig. S1) as well as possible auditory distractors.

G.2.5 Data Analysis

Initial processing of the FFR was identical to that of [22]. Hence, individual FFRs for each participant were filtered offline from 100 to 2000 Hz with a second order digital Butterworth bandpass filter [54]. After filtering, all trials were averaged over a 75 ms window, starting 15.8 ms prior stimulus onset. The artifact rejection criterion for invalid trials (e.g., myogenic activity) was set at $\pm 23.8\ \mu$V. The averaged response corresponding to each polarity was added together (i.e., $\frac{\text{negative polarity + positive polarity}}{2}$), thereby minimizing the stimulus artifact and the cochlear microphonic [22, 50]. This procedure was followed for both the audio and the audio-tactile condition.

The stereotyped peak landmarks and their distinct timing for the FFR to the short */da/* stimulus are well established in literature for the auditory modality [49], and are termed 'V', 'A', 'C', 'D', 'E', 'F', 'O' respectively (Fig. G.3B). Peak 'V' signifies the positive amplitude deflection associated with the stimulus onset and occurs at $\sim$6-8 ms. Peak 'A' is a negative amplitude deflection directly following 'V'. Peak 'C' reflects the transition from the onset burst to the onset of voicing. Subsequently, the three peaks 'D', 'E', and 'F' are all negative deflections related to the voicing of the speech sound and are spaced $\sim$8 ms apart (i.e., period of the fundamental frequency). Lastly, 'O' constitutes a negative amplitude deflection characterizing the sound offset response. For both the audio and audio-tactile condition, peaks were identified and labelled manually from the averaged response of the participant.

To investigate the neural frequency processing, we were specifically interested in the spectral encoding of the FFR. Therefore, a fast Fourier transform (FFT) was applied to the formant transition of the */da/* sound as suggested in the study of [21]. This transition corresponded to the 19.5-44.2 ms period in the averaged window. Zero-padding was applied to increase the spectral resolution to at least 1 Hz, and a Hanning window was applied to minimize spectral leakage. The obtained spectral encoding was then analyzed to investigate the fundamental frequency (F_0: 75-175 Hz), a neural correlate associated with pitch perception. Additionally, the harmonics of F_0 were examined. These harmonics were categorized into two bins: lower and higher harmonic content. The lower harmonics contained the first formant (F1) and ranged from 175-750 Hz. The higher harmonics, termed high frequency (HF), represented the frequencies between the first formant and the midbrain phase-locking limits (up to 1050 Hz). Accordingly, the magnitude corresponding to the average spectral energy of each frequency bin (i.e., F_0, F1, HF) were computed for comparison. Normative data for the auditory modality of the spectral encoding is visualized in Fig. G.3B.

Subsequent data analysis followed a within-subject design. More specifically, the FFRs under both audio and audio-tactile conditions were compared within each participant. These intra-individual comparisons were then pooled across participants to facilitate the statistical analyses. Accordingly, motivated by the benefits of simulation methods, paired permutation tests were employed to assess statistical significance [14]. Suppose we collect the random variable Y = $X^i_{\text{audio-tactile}}$ - X^i_{audio} for each participant i, where X^i represents any type of FFR measure (e.g., the average spectral energy of F_0). Under the null hypothesis of exchangeability, a test statistic distribution $\overline{Y}$ is created by randomly permuting the condition labels (i.e., audio-tactile and tactile) for each participant, 100.000 times. Subsequently, the p-value with alternative hypothesis H_A: $\overline{y} > 0$ (frequency domain analysis) or H_A: $\overline{y} \neq 0$ (time domain analysis) was determined for the observed unpermuted data. In case of multiple testing, a Bonferroni correction was performed to account for the family-wise type I error rate.

G.2.6 DualSense Validation

Following the protocol outlined by A. Farina [8], the frequency response characteristic of the DualSense controller was determined by employing an exponential sine sweep signal:

$$\mathrm{x(t)} = \sin\left(\frac{2\pi \cdot \mathrm{f}_1 \cdot \mathrm{T}}{\ln\left(\frac{\mathrm{f}_2}{\mathrm{f}_1}\right)} \cdot \left(\mathrm{e}^{\frac{\mathrm{t}}{\mathrm{T}}\ln\left(\frac{\mathrm{f}_2}{\mathrm{f}_1}\right)} - 1\right)\right) \tag{G.1}$$

with start frequency f_1 = 10 Hz, stop frequency f_2 = 1050 Hz and duration T = 50 s. In accordance with updated recommendations by A. Farina [9], this single very long sweep was further processed by applying a fade-in using a one-sided Hanning window of 0.1 s. Measurements were conducted in the same anechoic chamber as the FFR recordings, and a GY-61 ADXL335 analog 3-axis accelerometer (Analog Devices inc, Wilmington, United States) was placed at both the left and right hand side of the controller to record the vibrations (Fig. S2A). Furthermore, to simulate the natural dampening effect while holding the controller, it was placed on top of an ordinary blanket. The analog signal of the accelerometer was captured using the Steinberg UR44C audio interface (Steinberg Media Technologies GmbH, Hamburg, Germany) at a sampling rate of 48 kHz.

Additionally, the controller's ability to handle the high stimulus rate of the /da/ trials was evaluated. This involved stimulating the controller with

4000 /da/ trials at a rate of 10.9 Hz (positive polarity only). The positions of the accelerometers, the recording apparatus, and the sampling rate remained unchanged compared to the acquisition of the frequency response. The response to each /da/ trial was then analyzed by comparing the root mean square (RMS) value during a prestimulus period of 24 ms before onset, to the stimulus period of 40 ms after onset (after correcting the vibrotactile delay, see Fig. S3B). Subsequently, the signal-to-noise ratio (SNR) for each trial was computed according to:

$$\mathrm{SNR} = 20\log_{10}\left(\frac{\mathrm{RMS}_{\mathrm{stimulus}}}{\mathrm{RMS}_{\mathrm{prestimulus}}}\right) \tag{G.2}$$

For statistical analysis a similar paired permutation test was applied as detailed in section G.2.5, now adapted to evaluate $Y = 20\log_{10}\left(\frac{\mathrm{RMS}^{i}_{\mathrm{stimulus}}}{\mathrm{RMS}^{i}_{\mathrm{prestimulus}}}\right)$ for each */da/* trial i. The p-values were determined following the alternative hypothesis H_A: $\overline{y} > 0$.

G.3 Results

This section outlines the results following the methods described in section G.2. Initially, it discusses the preliminary validation of the vibrotactile controller which demonstrated positive results supporting its suitability for the FFR recordings. Following parts of this section will focus on the time domain and frequency domain analyses of the acquired FFRs.

G.3.1 Preliminary Validation of the Vibrotactile Controller

The controller's frequency response was measured at two symmetrical locations adjacent to the vibrotactile actuators and the regions where the participant's hands make contact (Fig. S2A). Both locations produced similar responses as shown in Fig. S2C. The data obtained across both sides were averaged for each dimension (i.e., x, y, z) and are presented in Fig. G.4.

Furthermore, given the high stimulus rate of the transient */da/*, a valid concern arises regarding the inertia of the vibrotactile actuator. That is, if the actuator fails to return to its baseline level in time between two successive trials, the participant may not be able to differentiate between the individual trial instances. For vibrotactile stimuli presented to the hand, literature has shown that the minimum detectable separation between two successive stimuli is on the order of 8-12 ms [30]. Those thresholds were found for noise and clicks at sensation levels of about 35 dB and for sinusoids with sensation levels of about 20 dB. While these stimulus parameters are distinct from the /*da*/ signal employed in this study, they do provide a reference point for estimating the temporal discrimination threshold of the vibrotactile modality. In this study, a wide separation margin of 24 ms was employed. Specifically, a prestimulus period of 24 ms was defined as the baseline activity. Accordingly, the SNR was computed for 4000 consecutive /*da*/ trials, presented at 10.9 Hz. The averaged vibrotactile signals are visualized in Fig. S2B and the statistical results are summarized in Table G.1. Hence, for each dimension the p-value neared 0 and large SNR values were found. Together, this suggested that the signal returned sufficiently to baseline levels between consecutive trials. The latencies observed for the onset of vibrations (Fig. S2B) were the result of the inertia of the vibrotactile actuator. This delay amounted to $\sim$8 ms, a duration which is well below the perceivable threshold of $\sim$40 ms [2].

The next step was to systematically calibrate the amplitude levels. Therefore, the peak-to-peak level was defined at the most dominant axis, orientated parallel to the built-in voice coil actuators of the controller (i.e., the z-axis). Accordingly, the 40-ms /*da*/ stimulus was calibrated to maintain a consistent amplitude of 0.85 m/s^2 (Fig. S2B) for all remaining experiments.

G.3.2 FFR Time Domain

The aim of the FFR time domain analysis was to determine whether the incorporation of the vibrotactile modality exerted any substantial effect on the timing of peaks, which could in turn confound the results of spectral analysis. A summary of the peak picking is presented in Table G.2. No significant effect was found for all of the peak timing differences. The grand average across all participants is visualized in Fig. G.5A. Furthermore, the FFR onset delay of $\sim$6-7 ms was in accordance with the neural transmission time [22, 49].

Additionally, control experiments were performed among 3 of the participants under vibrotactile stimulation only. That is, the procedure for recording the FFRs remained nearly identical, only now the insert earphone transducer in the right ear was replaced by an earplug. This modification facilitated a direct comparison between the vibrotactile condition and a baseline condition absent of both vibrotactile and auditory stimulation (Fig. S3). As a result, the vibrotactile modality only did not replicate the same waveform patterns as observed in Fig. G.5A. Instead, the response rather converges to the baseline level, suggesting that exclusively employing the vibrotactile signal in this study was insufficient to produce the stereotypical FFR.

G.3.3 FFR Frequency Domain

To assess the potential facilitatory role of vibrotactile stimuli in subcortical phase locking activity, it was crucial to examine if such an effect was reflected in the spectral encoding of the FFR. The grand spectral average across all participants is visualized in (Fig. G.5A). The magnitude of the fundamental frequency (i.e., F_0: 75 Hz - 175 Hz) required special attention due to its correlation with pitch perception in earlier FFR studies. Focusing on F_0, a paired permutation test was performed and visualized in Fig. G.5B. A significant difference of 3.30 $\pm$ 8.41 μV (p-value = .033, Cohen's d = 0.39, n = 22) was found, demonstrating enhanced phase locking activity of the fundamental frequency. In contrast, brief inspection of the higher harmonics F1 (175 Hz - 750 Hz) and HF (750 Hz - 1050 Hz) revealed similar spectral amplitudes between both audio and audio-tactile conditions.

Additionally, spectral encoding of the control experiments under exclusive vibrotactile stimulation (Fig. S3) rendered similar amplitude levels as baseline activities. This restates that vibrotactile stimuli in isolation do not produce the observed FFR response.

G.4 Discussion

This study evaluated the hypothesis that vibrotactile stimuli enhance auditory phase locking in subcortical regions. The obtained FFRs from the audio-tactile condition corroborated this proposition, exhibiting increased neural synchronization at the fundamental frequency F_0 specifically. Vibrotactile stimulation in the absence of sound produced a response that was indistinguishable from baseline levels. The latter hints that the observed multisensory effect does not arise from a mere linear summation but rather involves a super-additive interaction. Though, additional control experiments with statistical reporting should provide direct evidence to support this statement.

Previous research has already documented the role of the midbrain's inferior colliculus (i.e., the main neural source of the FFR) as a central processing hub for the integration of multisensory signals [17, 26]. This includes both feedforward projections [15] and corticofugal projections [25] mediated by somatosensory substrates. As mentioned, the temporal coding mechanism in the sensory peripheries of both auditory and vibrotactile modalities are highly similar. It therefore appears plausible that the observed multisensory effect is facilitated through connections in the afferent pathway, exhibiting Hebbian learning. However, the potential contributions of efferent projections from both primary and non-primary cortical areas to the inferior colliculus cannot be discounted. For instance, an analogous study on the impact of visual stimulation on the auditory FFR similarly reported an increased representation of F_0 [34]. As mentioned by the authors, one explanation could be derived from the reverse hierarchy theory [1]. This states that peripheral plasticity can be mediated by top-down corticofugal influences that originate from multisensory training. Likewise, Lakatos et al. demonstrated how a salient nonauditory stimulus (e.g., vibrotactile or visual) can enhance the neural excitability in the auditory cortex by phase resetting the ongoing oscillatory activity [24]. This observation has been linked to the phenomenon of increased perceived loudness of auditory signals under concurrent vibrotactile stimulation [23]. Furthermore, our research does not exclude the possibility that the observed effect may be attributed to a generic "novelty effect" rather than being an exclusive integration of audio-tactile stimuli. Hence, what mechanism exactly applies to the increased phase locking effect observed in the current study remains to be answered.

Our investigation did not detect enhanced phase locking at harmonic frequencies beyond the fundamental frequency. The harmonic content of acoustical stimuli has been linked to the perception of auditory timbre, which distinguishes the sounds of different instruments and voices [45]. Analogously, the perception of tactile texture has been correlated with the harmonic content of skin vibrations [28]. These harmonic frequencies are additionally reflected in the temporal coding of afferent neurons within both sensory peripheries. Hence, the substantial parallels between auditory timbre and tactile texture suggests a close relationship, similar to pitch processing. In this regard, previous research by Russo et al. has shown that auditory timbre could be discerned solely through vibrotactile stimulation [43, 44]. Yet, similar observations were not directly reflected in our findings. This may be attributed to both biological and technological constraints inherent in the present study. First is the limiting nature of perceived vibrations through the hand, exhibiting a relatively small bandwidth with optimal sensitivity around 240 Hz (Fig. G.2A). This sensitivity markedly declines at frequencies extending up to 1000 Hz. As a result, the harmonic content of the employed /*da*/ stimulus might not have been effectively transmitted. Additional low-pass filtering imposed by the vibrotactile controller may have further exacerbated this issue. Specifically, its frequency response peaks around the fundamental frequency of /*da*/ and decreases for higher frequencies. Furthermore, the choice of adding both polarities favors the FFR to the temporal envelope which contains the low frequency content including the fundamental frequency [22]. This is due to the fact that the temporal envelope is relatively phase invariant, thereby showing similar responses to both opposing polarities. Conversely, the temporal fine structure which includes the harmonic content is sensitive to the phase and thus cancels out when adding both polarities. Collectively, our design may have been biased towards the fundamental frequency, potentially obscuring any audio-tactile phase locking effects at higher harmonics. One improvement for future endeavours could be to explore the subtracted polarity since it accentuates the spectral fine structure at the expense of introducing more noise [22, 50].

Besides, both the auditory and vibrotactile /*da*/ stimulus were presented at levels well above the detection threshold. It is, however, widely documented that multisensory interactions are most effective when employing stimuli that, in isolation, are minimally effective in producing neural responses. This relationship between the intensity of unisensory stimuli and the relative strength of the combined multisensory response is denoted as the principle of inverse effectiveness [31]. For example, Fu et al. demonstrated that tactile input fluctuating in-phase with auditory noise amplifies the fluctuations of the noise [10]. Importantly, this multisensory effect was largest when the auditory stimulus was weakest. Hence, it would be interesting to investigate whether the enhanced audio-tactile phase locking effect obeys the same rule. Future research should explore the use of weaker auditory stimuli to determine whether subsequent incorporation of vibrotactile stimuli could yield more pronounced effects with larger effect sizes.

Furthermore, despite utilizing speech stimuli and an ergonomic controller, the ecological validity of this study may still be questioned. The necessity for the high repetitiveness of a single stimulus to record the FFR is not typical in real-life scenarios. Additionally, correlations with behavioral measures have not been performed. Previous studies on the FFR has demonstrated a connection between enhanced neural synchronization at the fundamental frequency and improved pitch discrimination capabilities [4, 19, 20]. It is therefore crucial to correlate the observed neurological effect presented in this study with behavioral outcomes to validate their functional significance. Accordingly, we propose that future studies should include a set of vibrotactile stimuli varying in specific parameters. For instance, it would be intriguing to explore vibrations with shifted pitches (i.e., shifted F_0) compared to the auditory modality. Research by Yau et al. has indicated that the perception of auditory pitch shifts towards a concurrent vibrotactile pitch when the frequencies of both modalities are closely aligned. [60]. Investigating whether a similar pattern is reflected in the encoding of F_0 in the FFR would be valuable. It would also be worthwhile to examine whether audio-tactile effects are restricted to vibrotactile stimulation of the hand only, or if other locations exhibit similar phenomena. Thus, future research should also consider varying the stimulation sites to assess its impact.

Additionally, the inclusion criteria for the participants in this study simply considered individuals with no self-reported hearing impairments. The latter statement could be strengthened by measuring actual audiograms. It would further be beneficial to gather a more comprehensive range of demographic data. Such data would enable the comparison of different populations to identify potential confounding variables that may influence the effectiveness of integrating the vibrotactile modality. An initial area of interest could be the role of musicianship. Musicians have shown to exhibit superior auditory processing abilities compared to non-musicians, including enhanced FFRs [17]. For example, the enhanced F_0 representation in the audio-visual paradigm was more pronounced among musicians [34]. Given that they also exhibit improved tactile frequency discrimination capabilities [48], it would be logical to hypothesize a similar positive correlation between musicianship and the effectiveness of audio-tactile stimulation on the FFR. Additionally, extending this research to clinical populations, particularly individuals with hearing impairments such as cochlear implant (CI) users, offers an interesting avenue for research explorations. Little is understood regarding the neural plasticity occurring within this population, and empirical observations revealed distinct patterns of musical engagement compared to normal hearing people who more dominantly rely on the hearing senses [11]. Possibly, inclusion of vibrotactile stimuli in the study of the FFR may also exhibit enhanced effectiveness in such individuals.

In conclusion, using speech stimuli, our data show elevated phase locking activity at the fundamental frequency in human subjects under audio-tactile stimulation. Given that prior research has linked enhanced F_0 encoding with augmented pitch processing capabilities, these findings hold promising practical implications. For example, implementing audio-tactile training to enhance pitch intelligibility may offer a practical approach for individuals struggling with pitch deficits, such as those with tone deafness or CI users. While this study represents an initial step towards exploring the potential benefits of audio-tactile stimulation, future studies are essential to further investigate its efficacy.

Figure captions

Dimension	*M*	*SD*	*p*	Cohen's *d*
x	15.49	0.68	$< .001$	22.83
y	14.74	1.56	$< .001$	9.45
z	13.39	1.30	$< .001$	10.25

Table G.1:
Statistical results of the validation of the vibrotactile controller
Note. For all trials (n = 4000), the SNR is displayed for each dimension (i.e., x, y, and z) according to section G.2.6. A significance level, α, of 0.05/3 = .016 is determined after Bonferroni correction.

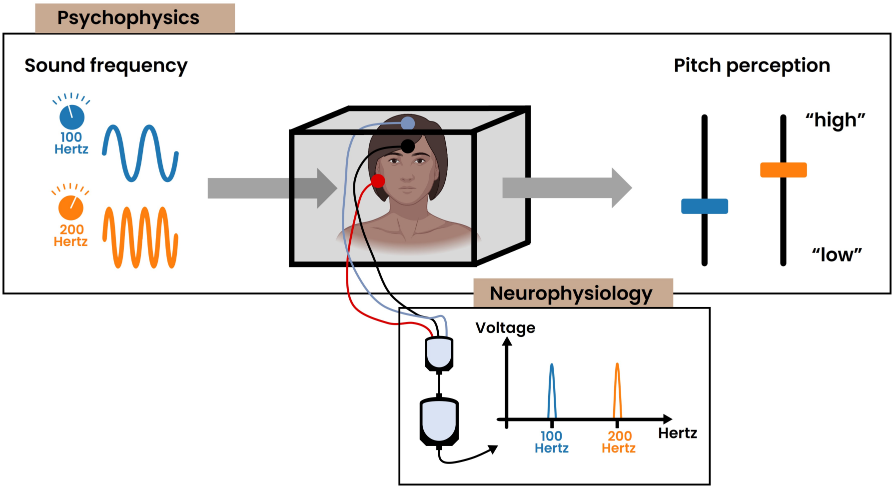


Fig. G.1: The common strategy for exploring the relationship between sensation and perception, combining the results obtained by the field of psychophysics and neurophysiology. The subject of study is displayed centrally in a box. **(Psychophysics)** Illustration depicting the positive correlation between the physical frequency of a sinusoidal stimulus and the resultant pitch perception. **(Neurophysiology)** Examining the neural mechanisms that underlie these psychophysical observations. In this example, the FFR captures the phase locking activity in the brain stem to the frequency of the auditory stimulus.

Landmark	*M*	*SD*	*p*	Cohen's *d*
V	-0.07	0.16	.031	0.42
A	0.06	0.17	.136	0.34
C	-0.21	0.71	.154	0.30
D	-0.03	0.69	.878	0.04
E	0.07	0.48	.502	0.15
F	0.07	0.75	.703	0.09
O	0.21	0.45	.016	0.48

Table G.2: item *Note.* The letters V, A, C, D, E, F and O refer to the distinct landmarks (Fig. G.3B). The peak timing difference of a landmark is displayed in ms and determined for all participants ($n = 22$). Negative values mean a later peak timing for the audio modality (audio $>$ audio-tactile). A significance level, α, of $0.05/7 = .007$ is determined after Bonferroni correction.

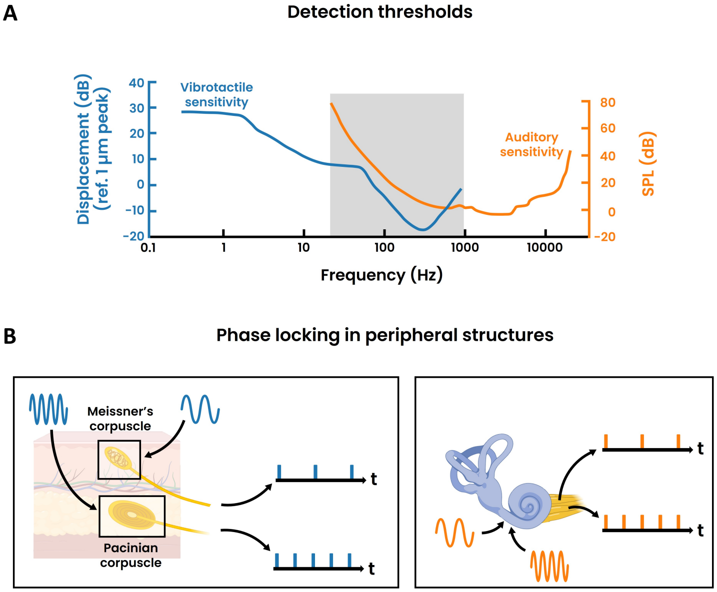


Fig. G.2: Parallels between the auditory and vibrotactile modality. **(A)** Detection thresholds for the auditory and vibrotactile modality across different frequencies. The vibrotactile sensitivity curve is derived from [12] where 700 ms sinusoidal stimuli are presented through a 0.72 cm^2 circular contactor at the index fingertip of the right hand. The auditory sensitivity curve is based on the International Standard ISO 389-7: 2003 for pure tones under free-field and diffuse-field listening conditions. The grey area indicates the frequency range where both curves overlap. **(B)** Phase locking activity in the peripheral structures of the vibrotactile modality (left) and auditory modality (right) for arbitrary stimulus frequencies respectively. Single nerve fibers are displayed for illustrative reasons. In reality, the precise synchronization should be regarded as the response of a population of nerve fibers.

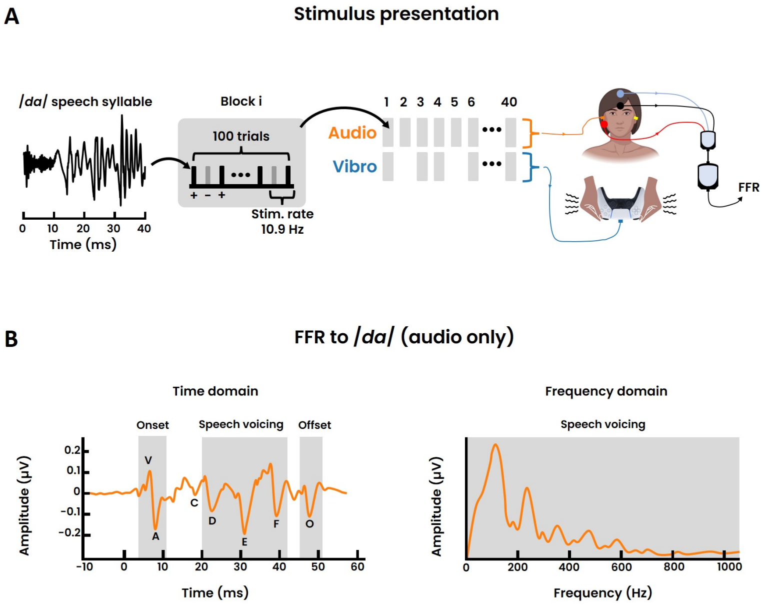


Fig. G.3: FFR recording. **(A)** Stimulus presentation of the 40-ms /*da*/ speech syllable for a single session. Each block consisted of 100 individual /*da*/ trials, presented in alternating polarity (+, -) with a stimulus rate of 10.9 Hz. In total 40 blocks were delivered with the randomized inclusion of the vibrotactile modality. Accoringly the FFR was recorded for both the audio and audio-tactile condition following a vertical electrode montage. **(B)** Normative FFR data to a 40-ms /*da*/ sound stimulus, adapted from [49]. The time domain signal on the left displays the 7 stereotypical peaks, representing: the stimulus onset response ('V', 'A'), the transition from stimulus onset to voicing onset ('C'), the voicing of the speech ('D', 'E', 'F') and the offset response ('O') respectively. The right graph displays the frequency domain of the speech voicing region of the FFR (19.5 - 44.2 ms). The largest peak in the spectrum represents the fundamental frequency.

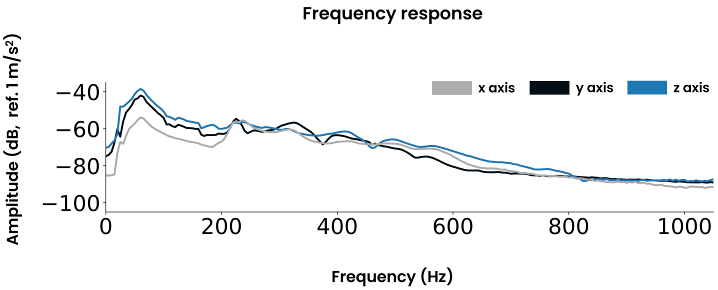


Fig. G.4: The frequency response of the vibrotactile controller for the three axes x, y and z respectively.

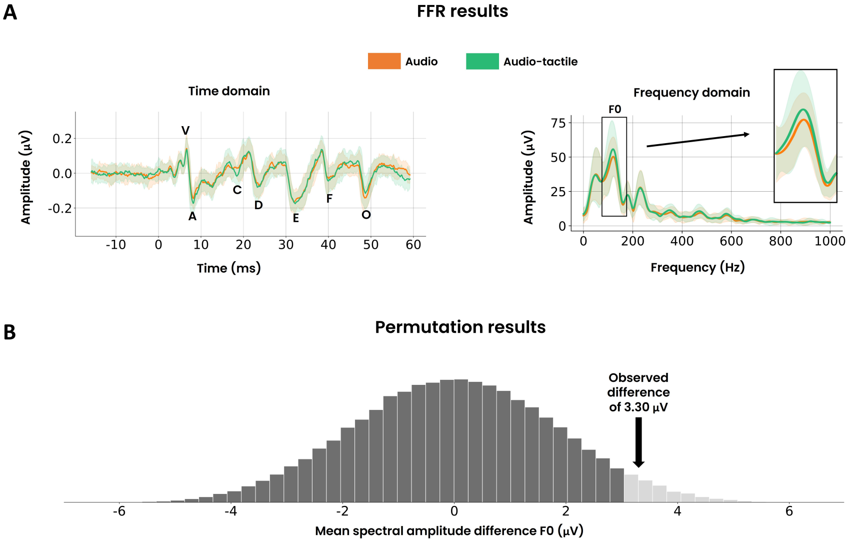


Fig. G.5: FFR results and permutation results. **(A)** Time domain representation (left) of the FFR where the added polarity was averaged across all participants. The mean value (solid curve) and standard deviation (fill areas around the mean) are displayed. The marked peaks follow the same definitions as detailed in Fig. G.3B. Spectral encoding of the time domain (right) for the window [19.5 ms - 44.2 ms], averaged across all participants. Again, the mean value (solid curve) and standard deviation (fill areas around the mean) are displayed. The fundamental frequency bin (F_0: 75 Hz - 175 Hz) is highlighted on the right. **(B)** The null distribution of the mean spectral amplitude difference for F_0, approximated by randomly permuting the condition labels. The light hued tail on the right represents the one-sided critical region for a significance level of 0.05. The black arrow demonstrates that the observed difference of 3.30 μV resides within this critical region, thereby providing statistical evidence to reject the null hypothesis.

References

[1] M. Ahissar and S. Hochstein, “The reverse hierarchy theory of visual perceptual learning,” *Trends in Cognitive Sciences*, vol. 8, no. 10, pp. 457–464, Oct. 2004.

[2] D. Brahimaj, M. Ouari, A. Kaci, F. Giraud, C. Giraud-Audine, and B. Semail, “Temporal Detection Threshold of Audio-Tactile Delays With Virtual Button,” *IEEE Transactions on Haptics*, pp. 1–6, 2023.

[3] P. Bruns, C. J. Camargo, H. Campanella, J. Esteve, H. R. Dinse, and B. Röder, “Tactile Acuity Charts: A Reliable Measure of Spatial Acuity,” *PLoS ONE*, vol. 9, p. e87384, Feb. 2014.

[4] S. Carcagno and C. J. Plack, “Subcortical Plasticity Following Perceptual Learning in a Pitch Discrimination Task,” *Journal of the Association for Research in Otolaryngology*, vol. 12, pp. 89–100, 2011.

[5] E. B. J. Coffey, T. Nicol, T. White-Schwoch, B. Chandrasekaran, J. Krizman, E. Skoe, R. J. Zatorre, and N. Kraus, “Evolving perspectives on the sources of the frequency-following response,” *Nature Communications*, vol. 10, no. 1, p. 5036, 2019.

[6] L. E. Crommett, D. Madala, and J. M. Yau, “Multisensory perceptual interactions between higher-order temporal frequency signals.” *Journal of Experimental Psychology: General*, vol. 148, no. 7, pp. 1124–1137, Jul. 2019.

[7] L. E. Crommett, A. Pérez-Bellido, and J. M. Yau, “Auditory adaptation improves tactile frequency perception,” *Journal of Neurophysiology*, vol. 117, no. 3, pp. 1352–1362, Mar. 2017.

[8] A. Farina, “Simultaneous measurement of impulse response and distortion with a swept-sine technique,” p. 24, 2000.

[9] ——, “Advancements in impulse response measurements by sine sweeps,” in *Audio Engineering Society Convention 122*, May 2007.

[10] 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, Jul. 2023.

[11] R. Fulford, J. Ginsborg, and J. Goldbart, “Learning not to listen: The experiences of musicians with hearing impairments,” *Music Education*

Research, vol. 13, no. 4, pp. 447–464, 2011.

[12] G. A. Gescheider, S. J. Bolanowski, J. V. Pope, and R. T. Verrillo, "A four-channel analysis of the tactile sensitivity of the fingertip: Frequency selectivity, spatial summation, and temporal summation," *Somatosensory & Motor Research*, vol. 19, no. 2, pp. 114–124, Jan. 2002.

[13] H. Gillmeister and M. Eimer, "Tactile enhancement of auditory detection and perceived loudness," *Brain Research*, vol. 1160, pp. 58–68, Jul. 2007.

[14] C. A. Holt and S. P. Sullivan, "Permutation tests for experimental data," *Experimental Economics*, vol. 26, no. 4, pp. 775–812, Sep. 2023.

[15] R. Jain and S. Shore, "External inferior colliculus integrates trigeminal and acoustic information: Unit responses to trigeminal nucleus and acoustic stimulation in the guinea pig," *Neuroscience Letters*, vol. 395, no. 1, pp. 71–75, Feb. 2006.

[16] D. H. Klatt, "Software for a cascade/parallel formant synthesizer," *The Journal of the Acoustical Society of America*, vol. 67, no. 3, pp. 971–995, 1980.

[17] N. Kraus, *Of Sound Mind: How Our Brain Constructs a Meaningful Sonic World.* MIT Press, 2021.

[18] N. Kraus, S. Anderson, and T. White-Schwoch, "The Frequency-Following Response: A Window into Human Communication," in *The Frequency-Following Response: A Window into Human Communication*, N. Kraus, S. Anderson, T. White-Schwoch, R. R. Fay, and A. N. Popper, Eds. Cham: Springer International Publishing, 2017, pp. 1–15.

[19] A. Krishnan, J. T. Gandour, and G. M. Bidelman, "The effects of tone language experience on pitch processing in the brainstem," *Journal of Neurolinguistics*, vol. 23, no. 1, pp. 81–95, 2010.

[20] A. Krishnan, Y. Xu, J. Gandour, and P. Cariani, "Encoding of pitch in the human brainstem is sensitive to language experience," *Cognitive Brain Research*, vol. 25, no. 1, pp. 161–168, 2005.

[21] J. Krizman, S. Bonacina, and N. Kraus, "Sex differences in subcortical auditory processing emerge across development," *Hearing Research*, vol. 380, pp. 166–174, 2019.

[22] J. Krizman and N. Kraus, "Analyzing the FFR: A tutorial for decoding the richness of auditory function," *Hearing Research*, vol. 382, p. 107779, 2019.

[23] P. Lakatos, C.-M. Chen, M. N. O'Connell, A. Mills, and C. E. Schroeder, "Neuronal Oscillations and Multisensory Interaction in Primary Auditory

Cortex," *Neuron*, vol. 53, no. 2, pp. 279–292, Jan. 2007.

[24] P. Lakatos, M. N. O'Connell, A. Barczak, A. Mills, D. C. Javitt, and C. E. Schroeder, "The Leading Sense: Supramodal Control of Neurophysiological Context by Attention," *Neuron*, vol. 64, no. 3, pp. 419–430, Nov. 2009.

[25] M. Lohse, J. C. Dahmen, V. M. Bajo, and A. J. King, "Subcortical circuits mediate communication between primary sensory cortical areas in mice," *Nature Communications*, vol. 12, no. 1, p. 3916, 2021.

[26] M. Lohse, P. Zimmer-Harwood, J. C. Dahmen, and A. J. King, "Integration of somatosensory and motor-related information in the auditory system," *Frontiers in Neuroscience*, vol. 16, p. 1010211, 2022.

[27] Maddieson, *Patterns of Sounds.* Cambridge University Press, 1984.

[28] L. R. Manfredi, H. P. Saal, K. J. Brown, M. C. Zielinski, J. F. Dammann, V. S. Polashock, and S. J. Bensmaia, "Natural scenes in tactile texture," *Journal of Neurophysiology*, vol. 111, no. 9, pp. 1792–1802, May 2014.

[29] S. Merchel and M. E. Altinsoy, "Auditory-Tactile Experience of Music," in *Musical Haptics*, S. Papetti and C. Saitis, Eds. Cham: Springer International Publishing, 2018, pp. 123–148.

[30] ——, "Psychophysical comparison of the auditory and tactile perception: A survey," *Journal on Multimodal User Interfaces*, vol. 14, pp. 271–283, 2020.

[31] M. A. Meredith and B. E. Stein, "Interactions Among Converging Sensory Inputs in the Superior Colliculus," *Science*, vol. 221, no. 4608, pp. 389–391, 1983.

[32] V. B. Mountcastle, R. H. LaMotte, and G. Carli, "Detection thresholds for stimuli in humans and monkeys: Comparison with threshold events in mechanoreceptive afferent nerve fibers innervating the monkey hand." *Journal of Neurophysiology*, vol. 35, no. 1, pp. 122–136, Jan. 1972.

[33] V. B. Mountcastle, W. H. Talbot, H. Sakata, and J. Hyvärinen, "Cortical neuronal mechanisms in flutter-vibration studied in unanesthetized monkeys. Neuronal periodicity and frequency discrimination." *Journal of Neurophysiology*, vol. 32, no. 3, pp. 452–484, May 1969.

[34] G. Musacchia, M. Sams, E. Skoe, and N. Kraus, "Musicians have enhanced subcortical auditory and audiovisual processing of speech and music," *Proceedings of the National Academy of Sciences*, vol. 104, no. 40, pp. 15 894–15 898, 2007.

[35] P. F. Nordmark, J. A. Pruszynski, and R. S. Johansson, “BOLD responses to tactile stimuli in visual and auditory cortex depend on the frequency content of stimulation,” *Journal of Cognitive Neuroscience*, vol. 24, no. 10, pp. 2120–2134, Oct. 2012.

[36] A. Pérez-Bellido, K. Anne Barnes, L. E. Crommett, and J. M. Yau, “Auditory Frequency Representations in Human Somatosensory Cortex,” *Cerebral Cortex*, vol. 28, no. 11, pp. 3908–3921, Nov. 2018.

[37] B. Petry, J. Huber, and S. Nanayakkara, “Scaffolding the Music Listening and Music Making Experience for the Deaf,” in *Assistive Augmentation*, J. Huber, R. Shilkrot, P. Maes, and S. Nanayakkara, Eds. Singapore: Springer Singapore, 2018, pp. 23–48.

[38] C. J. Plack, D. Barker, and D. A. Hall, “Pitch coding and pitch processing in the human brain,” *Hearing Research*, vol. 307, pp. 53–64, Jan. 2014.

[39] M. Prsa, D. Kilicel, A. Nourizonoz, K.-S. Lee, and D. Huber, “A common computational principle for vibrotactile pitch perception in mouse and human,” *Nature Communications*, vol. 12, no. 1, p. 5336, Sep. 2021.

[40] D. Purves, “The somatosensory system: Touch and proprioception,” in *Neuroscience*, D. Purves, G. Augustine, D. Fitzpatrick, W. Hall, A.-S. LaMantia, R. Mooney, M. Platt, and L. White, Eds. Oxford University Press, 2018, pp. 193–212.

[41] M. S. Rahman, K. A. Barnes, L. E. Crommett, M. Tommerdahl, and J. M. Yau, “Auditory and tactile frequency representations are co-embedded in modality-defined cortical sensory systems,” *NeuroImage*, vol. 215, p. 116837, Jul. 2020.

[42] 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, no. 19, p. 6575, Sep. 2021.

[43] F. A. Russo, “Music Beyond Sound: Weighing the Contributions of Touch, Sight, and Balance,” *Acoustics Today*, vol. 16, no. 1, p. 37, 2020.

[44] F. A. Russo, P. Ammirante, and D. I. Fels, “Vibrotactile discrimination of musical timbre.” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 38, no. 4, pp. 822–826, 2012.

[45] H. P. Saal, X. Wang, and S. J. Bensmaia, “Importance of spike timing in touch: An analogy with hearing?” *Current Opinion in Neurobiology*, vol. 40, pp. 142–149, Oct. 2016.

[46] M. Schürmann, G. Caetano, Y. Hlushchuk, V. Jousmäki, and R. Hari, "Touch activates human auditory cortex," *Neuroimage*, vol. 30, no. 4, pp. 1325–1331, 2006.

[47] 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, no. 2, pp. 830–832, 2004.

[48] A. Sharp, M. S. Houde, M. Maheu, I. Ibrahim, and F. Champoux, "Improved tactile frequency discrimination in musicians," *Experimental Brain Research*, vol. 237, no. 6, pp. 1575–1580, 2019.

[49] E. Skoe, J. Krizman, S. Anderson, and N. Kraus, "Stability and Plasticity of Auditory Brainstem Function Across the Lifespan," *Cerebral Cortex*, vol. 25, no. 6, pp. 1415–1426, 2015.

[50] E. Skoe and N. Kraus, "Auditory Brain Stem Response to Complex Sounds: A Tutorial," *Ear & Hearing*, vol. 31, no. 3, pp. 302–324, 2010.

[51] H. F. Sperdin, "Early, low-level auditory-somatosensory multisensory interactions impact reaction time speed," *Frontiers in Integrative Neuroscience*, vol. 3, 2009.

[52] B. E. Stein, T. R. Stanford, and B. A. Rowland, "Development of multisensory integration from the perspective of the individual neuron," *Nature Reviews Neuroscience*, vol. 15, no. 8, pp. 520–535, Aug. 2014.

[53] E. Verschooten, S. Shamma, A. J. Oxenham, B. C. Moore, P. X. Joris, M. G. Heinz, and C. J. Plack, "The upper frequency limit for the use of phase locking to code temporal fine structure in humans: A compilation of viewpoints," *Hearing Research*, vol. 377, pp. 109–121, Jun. 2019.

[54] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental algorithms for scientific computing in python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[55] P. Vogt and L. F. Kasper, "The missing fundamental tone in everyday life and in experiments," *The Physics Teacher*, vol. 61, no. 3, pp. 228–229, Mar.

2023.

[56] G. Von Békésy, “Similarities between hearing and skin sensations.” *Psychological Review*, vol. 66, no. 1, pp. 1–22, 1959.

[57] E. C. Wilson, L. D. Braida, and C. M. Reed, “Perceptual interactions in the loudness of combined auditory and vibrotactile stimuli,” *The Journal of the Acoustical Society of America*, vol. 127, no. 5, pp. 3038–3043, 2010.

[58] E. C. Wilson, C. M. Reed, and L. D. Braida, “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.

[59] J. M. Yau, J. B. Olenczak, J. F. Dammann, and S. J. Bensmaia, “Temporal Frequency Channels Are Linked across Audition and Touch,” *Current Biology*, vol. 19, no. 7, pp. 561–566, Apr. 2009.

[60] J. M. Yau, A. I. Weber, and S. J. Bensmaia, “Separate Mechanisms for Audio-Tactile Pitch and Loudness Interactions,” *Frontiers in Psychology*, vol. 1, 2010.

[61] W. A. Yost, “Pitch perception,” *Attention, Perception & Psychophysics*, vol. 71, no. 8, pp. 1701–1715, Nov. 2009.

Paper H

Multisensory Integration Design in Music for Cochlear Implant Users

D. Cavdir, F. Ganis, R. Paisa, P. Williams, S. Serafin

The article has been submitted to
The 19th Sound and Music Computing Conference Conference Proceedings, 2022.

© 2022, The Authors, under CC 4.0
The layout has been revised.

Original Publication

D. Cavdir, F. Ganis, R. Paisa, P. Williams, and S. Serafin, "Multisensory Integration Design in Music for Cochlear Implant Users," in *Proceedings of the 19th Sound and Music Computing Conference*, Saint-Étienne, France, Zenodo, June 2022, pp. 334–341. doi: 10.5281/ZENODO.6573315.

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.

H.1 Introduction

While CIs have achieved a high level of complexity in terms of hardware and ergonomics, training and rehabilitation programs for CI 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 [6, 7]. 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 CI users due to "age, cognitive processing residual hearing, hearing aid (HA) 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.

H.2 Related Work

H.2.1 Accessible and Inclusive Design in Music

When designing accessible digital musical instruments (ADMI) or accessible musical technologies (AMT), 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 [5]. 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 HAs. Similar participatory practices and their implications for rehabilitation are explored by accessibility researchers [27]; however their application to musical experience design, specifically for D/HOH 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 CI use and music.

H.2.2 CI Use and Music

CIs 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 [4, 8]. Furthermore, there is a general weak representation of the fundamental frequency (F_0) 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.

H.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 [9, 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 [1].

Another chair installation was designed by Nanayakkara et al. with the help of the hearing impaired community [22]. 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* [23]. 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 [23].

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

---

[1] https://subpac.com/

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.

H.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].

H.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 H.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*.

H.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

---

[2] Facebook CI Group

CI, experiencing a positive transition from HA to CIs. 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 CI in his left ear, and a HA 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 CI in the right ear, and near deaf in the left ear, with hearing threshold at +95dB. He has been using HAs 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*.

H.4 Design and Implementations

H.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.

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.

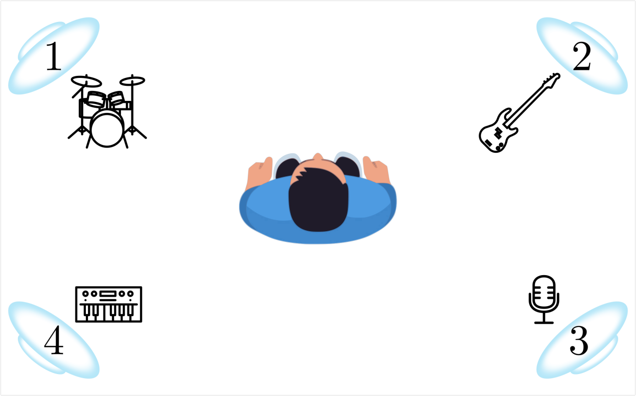


Fig. H.1: Scheme of installation 1.

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.

H.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.

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.

---

[3] https://thebuttkicker.com/

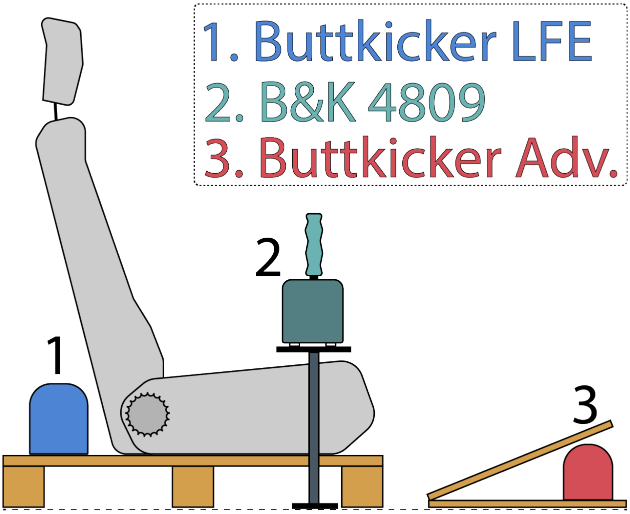


Fig. H.2: One configuration experienced by all participants

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 H.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 H.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 H.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 individually 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 [2]. 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.

Participants selected extra audio material for their preferred setup. All

---

[4] http://tactilelabs.com/

three preferred the setup in Figure H.2. P1 listened to *Fleetwood Mac - Dreams*, P2 *Vienna Philharmonic - An der schönen, blauen Donau* (excerpts), and P3 *Dizzy Mizz Lizzy - Silverflame.*

H.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 [3]. 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.

H.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.

H.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 H.4.2.

---

[5] https://tinyurl.com/2p8axhwp

[6] https://tinyurl.com/yck63zbz

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 H.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.

H.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 HA.

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.

H.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 H.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.

H.6 Discussion and Future Directions

H.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, HA 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 CI use and music can support their participation in design and research, enhancing their musical experiences.

H.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.

H.7 Conclusions

In this paper, we study 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 CIs 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 CI 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] 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.

[2] 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, no. 5, 2014.

[3] 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.

[4] 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, no. 2, 2015.

[5] 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.

[6] 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.

[7] 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.

[8] 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, no. 1, 2020.

[9] F. Fontana, I. Camponogara, P. Cesari, M. Vallicella, and M. Ruzzenente, “An exploration on whole-body and foot-based vibrotactile sensitivity to melodic consonance,” in *SMC 2016 - 13th Sound and Music Computing Conference, Proceedings*, ser. SMC 2016 - 13th Sound and Music Computing Conference, Proceedings, R. Grossmann and G. Hajdu, Eds. Zentrum fur Mikrotonale Musik und Multimediale Komposition (ZM4), Hochschule fur Musik und Theater, 2016, pp. 143–150.

[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, Aug. 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*, vol. 13, p. 1229, Nov. 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, D. A. Mcpherson, and D. T. Stockman, “Designing Tactile Musical Devices with and for Deaf Users: A Case Study,” p. 7, 2015.

[16] L. A. Jones and N. B. Sarter, “Tactile Displays: Guidance for Their Design and Application,” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 50, no. 1, pp. 90–111, Feb. 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, no. 3, pp. 160–169, Jul. 2009.

[18] C. Kayser, C. Petkov, M. Augath, and N. Logothetis, “Integration of touch and sound in auditory cortex,” *Neuron*, vol. 48, pp. 373–84, Nov. 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, 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,” ser. Conference on Human Factors in Computing Systems - Proceedings, Apr. 2009, pp. 337–346.

[23] S. C. Nanayakkara, L. Wyse, S. H. Ong, and E. A. Taylor, “Enhancing

Musical Experience for the Hearing-Impaired Using Visual and Haptic Displays," p. 47, 2013.

[24] R. Okazaki, H. Kajimoto, and V. Hayward, "Vibrotactile stimulation can affect auditory loudness: A pilot study," Jun. 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, no. 17, 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. E. Stein and M. A. Meredith, *The Merging of the Senses.* MIT press, 1993.

[32] A. Tilley, H. Dreyfuss, and H. D. Associates, *The Measure of Man and Woman: Human Factors in Design.* Whitney Library of Design, 1993.

[33] ALVAR. Wilska, "On the vibrational sensitivity in different regions of the body surface," *Acta Physiologica Scandinavica*, vol. 31, no. 2-3, 1954.

[34] B. S. Wilson and M. F. Dorman, "Cochlear implants: A remarkable past and a brilliant future," *Hearing Research*, vol. 242, no. 1-2, pp. 3–21, Aug. 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, Dec. 2016.