NEW DISCOVERIES IN THE BENEFITS AND OUTCOMES OF COCHLEAR IMPLANTATION

EDITED BY: Fei Chen, Jing Chen and Xin Luo

PUBLISHED IN: Frontiers in Neuroscience and Frontiers in Psychology

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-83250-903-6
DOI 10.3389/978-2-83250-903-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

NEW DISCOVERIES IN THE BENEFITS AND OUTCOMES OF COCHLEAR IMPLANTATION

Topic Editors:
Fei Chen, Southern University of Science and Technology, China
Jing Chen, Peking University, China
Xin Luo, Arizona State University, United States

Citation: Chen, F., Chen, J., Luo, X., eds. (2022). New Discoveries in the Benefits and Outcomes of Cochlear Implantation. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83250-903-6

Table of Contents

TYPE Editorial
PUBLISHED 15 November 2022
DOI 10.3389/fnins.2022.1062582

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Isabelle Peretz,
Université de Montréal, Canada

*CORRESPONDENCE
Fei Chen
fchen@sustech.edu.cn

SPECIALTY SECTION
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

RECEIVED 06 October 2022
ACCEPTED 31 October 2022
PUBLISHED 15 November 2022

CITATION
Chen F, Chen J and Luo X (2022) Editorial: New discoveries in the benefits and outcomes of cochlear implantation. *Front. Neurosci.* 16:1062582. doi: 10.3389/fnins.2022.1062582

Editorial: New discoveries in the benefits and outcomes of cochlear implantation

Fei Chen[1]*, Jing Chen[2] and Xin Luo[3]

[1]Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China, [2]Key Laboratory of Machine Perception, Ministry of Education, School of Intelligence Science and Technology, College of Future Technology, Peking University, Beijing, China, [3]Speech and Hearing Science, College of Health Solutions, Arizona State University, Tempe, AZ, United States

KEYWORDS

cochlear implantation, speech perception, cochlear implant benefits, psychophysical and neurophysiological methods, music and emotion perception

Editorial on the Research Topic
New discoveries in the benefits and outcomes of cochlear implantation

Cochlear implants (CIs), hearing prostheses that bypass sensory hair cells in the cochlea to directly stimulate the auditory nerve, have been shown to restore hearing for individuals with severe to profound hearing loss. Recent research has demonstrated enormous CI benefits for speech recognition, sound localization, as well as language development in children. The effectiveness of CIs is, however, affected by many factors, including the duration of deafness/hearing loss, implantation age, and duration of CI use. Whereas recent studies have largely focused on the effects of CIs on speech perception and production, further investigation is needed to study the effects of CIs on emotion processing, music perception, prosody perception, etc. The aim of this Research Topic Collection is to bring together studies addressing recent discoveries in the benefits and outcomes of cochlear implantation. A total of 17 papers are included in this Research Topic.

Five papers utilized neuroimaging techniques, including electroencephalogram (EEG) and functional near-infrared spectroscopy (fNIRS), to investigate CI users' cortical response characteristics on a range of auditory perception tasks. (1) As P300 is closely related to cognitive processes (e.g., auditory discrimination), Tao et al. investigated the connection between auditory segregation of competing speech in Mandarin-speaking bimodal CI users and the P300 component of event-related potentials elicited by 1 vs. 2 kHz contrast. Their results showed that the P300 amplitude was significantly correlated with the speech reception threshold in the same target-masker voice gender (male) condition in the CI group, which suggests the potential of P300 amplitude as a clinically useful neural indicator of central auditory processing capabilities that are susceptible to informational masking in bimodal CI users. (2) Xie et al. explored the relationship between the ability to detect frequency changes or temporal gaps and speech perception using psychophysical and neurophysiological methods among post-lingually deafened CI users. Their multiple regression analysis showed the predictive ability of gap detection threshold (GDT) and the amplitude of acoustic change complex (ACC) response on speech perception performance. This indicates that GDT (as a psychophysical measure) may work as an easy, quick and non-linguistic tool, and ACC amplitude induced by the temporal gap (as a neurophysiological indicator) may have the potential to predict speech outcomes. (3) Cartocci et al. studied the emotion processing of children with unilateral CI *via* EEG and behavioral measures. Compared to normal-hearing (NH) children, less accurate vocal emotional state recognition was observed in children with unilateral CI, which was correlated with increased gamma activity lateralization index (relatively higher right-hemisphere activity) in response to emotional speech stimuli. The implantation side for children with unilateral CI did not affect the contralateral gamma activity, but the age at implantation influenced emotion recognition. These indicate that a deficit in engaging the left hemisphere for emotional tasks exists in unilateral CI users and early implantation may be beneficial to children's emotion recognition. (4) Raghavendra et al. investigated the sound quality of speech produced by CI users from the perspective of NH listeners' perception. They decoded and re-constructed the speech envelope from single-trial EEG recorded on the scalp of the NH listeners using a regenerative model and computed the correlation between the actual and reconstructed speech envelope waveforms. They found that the perceived sound quality rating was associated with the cortical tracking of speech envelop, and speech produced by NH speakers was more closely tracked relative to that produced by CI speakers. (5) Lu et al. focused on the cortical coding of prosody measured by fNIRS in pre-lingually deafened children with CIs. They recorded cortical responses to natural sentences with strong or weak prosodic features, and evaluated participants' speech communication ability in three tasks of picture description, video content statement and free conversation. Weaker cortical activation and characteristic deficits in perceiving strong prosodies were observed in children with CIs compared to NH children. Sensitivity to strong prosodic information was significantly correlated with the speech communication ability of all children who participated in this study. This suggests the importance of speech prosody in children's speech development.

Understanding the predictive factors of CI users' auditory performance is crucial to prognosis and clinical decision-making. Factors that have reached general consensus such as age at implantation and duration of device use cannot explain all the variance in CI performance; exploring new independent factors that have predicative power is thus needed. (1) Lu et al. retrospectively examined clinical results of children with cochlear nerve deficiency (CND) and CIs in order to identify main predictive factors and develop predictive models using machine learning. A parameter-optimized support vector machine based on the vestibulocochlear nerve (VCN) area and the number of nerve bundles (measured with high-resolution computed tomography) was constructed to predict speech perception performance. These two factors were suggested to have the ability to predict CI outcomes in children with CND (i.e., prediction accuracies of 71 and 93% in hearing rehabilitation and speech rehabilitation, respectively), providing guidance on surgical side selection and prognosis. (2) Chao et al. focused on the predictive ability of pre-implantation imaging results for electrically evoked compound action potentials (ECAPs) of auditory nerve fibers in response to electrical stimuli in children with CND. They found that the width of the bony cochlear nerve canal or the VCN diameter did not correlate with ECAP responses, while the ratio of the VCN to facial nerve (FN) diameter was significantly correlated with the slope of the ECAP input/output function and the ECAP maximum amplitude. This suggests that the VCN to FN diameter ratio may be an effective predictor of the cochlear nerve function in children with CND and CIs. (3) Zheng and Liu reviewed CI outcomes in patients with auditory neuropathy caused by Otoferlin (OTOF) gene mutations. They concluded that patients with OTOF mutations had excellent performance in both sound perception and speech recognition, and suggested the importance of genetic analysis in localizing lesions and informing clinical decision-making. Early implantation for patients with biallelic OTOF mutations was encouraged. (4) Zhang et al. investigated the predictive power of imaging results but for the long-term auditory and speech perception development of pediatric CI users with common cavity deformity (CCD). Auditory and speech behaviors [using four parent reports questionnaires: categories of Auditory Performance (CAP), Speech Intelligibility Rating (SIR), Meaningful Auditory Integration Scale/Infant-Toddler Meaningful Auditory Integration Scale (MAIS/ITMAIS), and Meaningful Use of Speech Scale (MUSS)] improved over time after implantation in children with CCD, but were poorer than those of CI users with normal inner ear structures. The volume and lumen surface range of CCD reflected inner ear development and influenced CI outcomes.

The remaining eight papers cover a broad range of new topics on the outcomes of cochlear implantation. (1) Mao et al. themed around Mandarin tone production of pre-lingually deafened children with CIs. They recorded monosyllables produced by each participant and calculated the differentiability and hit rate of different tones using acoustic analyses. In general, children with CIs exhibited significantly poorer tone productions in both measures compared with NH children. A weak correlation between age at implantation or duration of CI use and acoustic measures (i.e., differentiability and hit rate, computed based on the F0 onset and offset values or the F0 onset, midpoint, and offset values) of tone productions was noted. (2) Balkenhol et al. focused on the benefits of adding a hearing aid (HA) on the contralateral side of a CI for sentence recognition and auditory evoked potential (AEP) responses.

Their results suggest that bimodal listeners can take advantage of head shadow and binaural summation effects, but cannot make use of binaural squelch and spatial release from masking at 6 months post-CI. The perceptual benefit of bimodal hearing may be objectively evaluated by the AEP responses, supported by a significant correlation of binaural summation effects with AEP latency differences between the bimodal and CI-only conditions. (3) Matz et al. studied auditory stream segregation in CI users by changing the spectral- and amplitude-modulation rate of narrowband noise (NBN) bursts. Their results showed the deficits of CI users in segregating NBN bursts into different auditory streams when they were moderately separated in the spectral domain compared to NH listeners. Both groups were able to utilize build-up effects to segregate auditory streams when lengthening the duration of stimulus sequences. (4) Liang et al. investigated the effect of implantation side on CI outcomes through behavioral measures and brain activation measured by ACC. Their results indicated that the implantation side may affect neural plasticity patterns in the adult population, demonstrated by a unique correlation between ACC activation patterns and performance in frequency change detection in subjects with a right-ear CI. (5) Yao et al. presented a literature review on the research status and future development of cochlear reimplantation. Although techniques are relatively matured since the 1980s, several issues were identified by the authors. The need for an international consensus statement on cochlear re-implantation in relevant problems (e.g., to standardize the definition, calculation formulas of reimplantation rate, and follow-up systems) is highlighted. (6) Wang et al. studied changes in vestibular function in patients who received a minimally invasive CI surgery 1 year before the study. The functions of semicircular canals and otolith were assessed by a comprehensive battery of tests. Most of the vestibular functions could be preserved with no damage discrepancy among the otolith and three semicircular canal functions at 12 months post CI. (7) Di Nardo et al. explored the benefits of CI in tinnitus suppression. Single-channel stimulation resulted in a significant reduction of tinnitus loudness. Their results provided insights into the mechanisms of tinnitus and the development of tinnitus therapies. (8) Bissmeyer et al. examined the effect of a computer-based musical training program on musical interval identification of CI users and listeners with no known hearing loss and the correlation between low-level psychophysics and higher-level musical abilities. They observed strong correlations between pitch sensitivity and musical interval identification for the two participant groups. However, the effect of the training program in this study on musical interval identification was small among CI users. The authors discussed directions toward improving pitch access and auditory training of musical interval appreciation for CI users.

The papers presented in this Research Topic provide a snapshot of the latest discoveries in the benefits and outcomes of CI. We believe that this Research Topic provides an interdisciplinary forum for researchers working in the fields of speech and hearing science, computational neuroscience, medicine, and biomedical engineering to present the most recent ideas, methodologies, and studies for understanding and improving the benefits and outcomes of CI. Future studies on the benefits and outcomes of CI could continue in several aspects, including establishing neural biomarkers of central auditory processing, discovering new predictive factors of CI outcomes, and exploring CI benefits in speech production, emotion processing and spatial listening. Utilizing multi-disciplinary methods in future CI studies is also recommended, e.g., combining neuroimaging and psychophysical tools to assess CI outcomes and integrating traditional statistical tests and artificial intelligence algorithms to analyze clinical findings.

The editors sincerely thank all the authors, reviewers, and Frontiers editors as well as the staff involved in the preparation of this Research Topic.

Author contributions

FC, JC, and XL contributed to the drafting and revision. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 29 April 2020
doi: 10.3389/fnins.2020.00368

The Effect of Side of Implantation on the Cortical Processing of Frequency Changes in Adult Cochlear Implant Users

Chun Liang[1,2]*, *Lisa H. Wenstrup*[3], *Ravi N. Samy*[3], *Jing Xiang*[4] *and Fawen Zhang*[1]*

[1] Department of Communication Sciences and Disorders, University of Cincinnati, Cincinnati, OH, United States, [2] Child Psychiatry and Rehabilitation, Affiliated Shenzhen Maternity & Child Healthcare Hospital, Southern Medical University, Shenzhen, China, [3] Department of Otolaryngology-Head and Neck Surgery, University of Cincinnati, Cincinnati, OH, United States, [4] Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

OPEN ACCESS

Edited by:
Fei Chen,
Southern University of Science and Technology, China

Reviewed by:
Carla Gentile Matas,
University of São Paulo, Brazil
Viral Tejani,
University of Iowa Hospitals and Clinics, United States

Correspondence:
Chun Liang
chunliang13@yeah.net;
liangcn@mail.uc.edu
Fawen Zhang
Fawen.Zhang@uc.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *14 December 2019*
Accepted: *25 March 2020*
Published: *29 April 2020*

Citation:
Liang C, Wenstrup LH, Samy RN, Xiang J and Zhang F (2020) The Effect of Side of Implantation on the Cortical Processing of Frequency Changes in Adult Cochlear Implant Users. Front. Neurosci. 14:368. doi: 10.3389/fnins.2020.00368

Cochlear implants (CI) are widely used in children and adults to restore hearing function. However, CI outcomes are vary widely. The affected factors have not been well understood. It is well known that the right and left hemispheres play different roles in auditory perception in adult normal hearing listeners. It is unknown how the implantation side may affect the outcomes of CIs. In this study, the effect of the implantation side on how the brain processes frequency changes within a sound was examined in 12 right-handed adult CI users. The outcomes of CIs were assessed with behaviorally measured frequency change detection threshold (FCDT), which has been reported to significantly affect CI speech performance. The brain activation and regions were also examined using acoustic change complex (ACC, a type of cortical potential evoked by acoustic changes within a stimulus), on which the waveform analysis and the standardized low-resolution brain electromagnetic tomography (sLORETA) were performed. CI users showed activation in the temporal lobe and non-temporal areas, such as the frontal lobe. Right-ear CIs could more efficiently activate the contralateral hemisphere compared to left-ear CIs. For right-ear CIs, the increased activation in the contralateral temporal lobe together with the decreased activation in the contralateral frontal lobe was correlated with good performance of frequency change detection (lower FCDTs). Such a trend was not found in left-ear CIs. These results suggest that the implantation side may significantly affect neuroplasticity patterns in adults.

Keywords: cochlear implant, frequency change detection, acoustic change complex, standardized low-resolution brain electromagnetic tomography (sLORETA), temporal lobe, frontal lobe

INTRODUCTION

Cochlear implants (CIs) have been successful in providing auditory sensation to individuals with severe to profound hearing loss. Recently, more and more CIs have been used in both children and adults with hearing loss. However, there is large variability in CI users' speech outcomes. Previous studies have suggested that CI outcomes could be affected by many factors such as the duration of deafness, age at implantation, the duration of CI use, cognitive ability, and electrode placement

(Finley and Skinner, 2008; Lazar et al., 2010; Lee et al., 2010; Blamey et al., 2012; Lazard et al., 2012; Holden et al., 2013; Pisoni et al., 2016). However, the underlying mechanism for the large variability in CI outcomes is still not well understood. This lack of information is a barrier to customized rehabilitation and hampers further improvement of CI outcomes.

It is well known that the right and left hemispheres of the brain play distinct roles in processing auditory information. In normal hearing, the left hemisphere is dominant in processing temporal information and the right hemisphere is dominant for spectral information (Zatorre et al., 2002; Schonwiesner et al., 2005; Hyde et al., 2008; Okamoto et al., 2009). In adults wearing CIs, processing spectral information in CI users is substantially impaired due to the low number of spectral channels used to deliver sound information and the deafness-related neural deficits in the auditory system (Giraud et al., 2001; Limb and Roy, 2014). Thus, the side of cochlear implantation may influence the outcomes of the CIs, particularly for adult CI users. However, direct information for the effect of the implantation side on the CI outcome is not available.

The electroencephalography (EEG) technique allows for the recording of auditory evoked potentials with an excellent temporal resolution, enabling the examination of real-time brain processing. EEG is the most suitable tool to examine the neural substrates of sound processing in CI users (Debener et al., 2008). With EEG techniques, researchers have examined the cortically generated auditory evoked potential (CAEP), which consists of the N1 and P2 peaks occurring in a window of approximately 70–250 ms after stimulus onset. The acoustic change complex (ACC) is a special type of the CAEP elicited by an acoustic change (e.g., a change in frequency, intensity, duration, etc.) embedded in a stimulus (Ostroff et al., 1998; Abbas and Brown, 2014; Brown et al., 2015; Kim, 2015). Data from non-CI users showed that the ACC threshold (the minimum magnitude of the acoustic change required to evoke the ACC) is in agreement with behaviorally measured auditory discrimination thresholds and that the ACC amplitude is related to the salience of the perceived acoustic change (He et al., 2012; Liang et al., 2016; Fulbright et al., 2017). Moreover, multi-channel EEG data could be used in source localization analysis to identify the activated brain regions (Worrell et al., 2000; Plummer et al., 2010; Eugene et al., 2014; Talja et al., 2015).

Source localization techniques can be used to estimate the current source generators in the brain that best fit the scalp recorded EEG or MEG data. Although there is no unique solution to the neuroimaging inverse problems, the standardized Low-Resolution Brain Electromagnetic Tomography (sLORETA, available at http://www.uzh.ch/keyinst/loretaOldy.htm) can be used to calculate neural generators of EEG or MEG data with exact and zero error localization for test dipoles (Pascual-Marqui, 2002; Wagner et al., 2004; Grech et al., 2008). Moreover, sLORETA has no localization bias in the presence of measurement and biological noise (Pascual-Marqui, 2002). Comparing other techniques such as WMN and LORETA, sLORETA gives the best solution in terms of both localization error and sources (Grech et al., 2008) and this method has been validated both theoretically and experimentally (Pascual-Marqui, 2002; Plummer et al., 2010). Using sLORETA, it is possible to specifically compare the activity of regions of interest (ROIs) between the left and right hemispheres or to examine the correlations between brain activities and behavioral measures.

Previous ACC studies in CI users were restricted to the analysis of the ACC waveform without the identification of neural generators of the response (Friesen and Tremblay, 2006). Meanwhile, in the studies using source localization analysis in CI users, the analysis was limited to the response evoked by the onset of the stimulus (i.e., the onset CAEP, Wong and Gordon, 2009; Gordon et al., 2010; Song et al., 2013; Nash-Kille and Sharma, 2014) rather than the response evoked by acoustic changes embedded in the stimulus (i.e., ACC).

A study from our lab reported that the CI users exhibited much poorer ACC waveforms evoked by frequency changes compared to normal hearing subjects (Liang et al., 2018). Behavioral studies from other researchers using pitch discrimination tasks or spectral modulation tasks have also reported that the capability of detecting changes in the frequency domain is crucial for speech performance in CI ears (Gifford et al., 2014; Won et al., 2014; Kenway et al., 2015). Therefore, it is important to examine brain activation patterns to frequency changes in CI users to understand the neural basis of the CI outcomes. In the current study, we performed sLORETA source analysis using ACC data collected for Liang et al.'s (2018), in which only ACC waveform results were reported, to examine brain activation patterns in response to tones containing frequency changes in the adult CI users. The focus of the current study, which is a companion paper to Liang et al. (2018), is to examine the effect of the implantation side (right- and left-ear CIs) on the cortical processing of frequency changes. Therefore, the brain activation patterns of the right- vs. left-ear CIs and additional waveform comparisons will be presented. To our knowledge, the current study is the first one to investigate the effects of the implantation side on cortical processing of frequency changes and to localize the neural substrates of the ACC N1' in CI users. This study will be valuable for guiding CI side selection and maximizing CI outcomes. We hypothesize that the brain activation patterns of the ACC N1' peak evoked by frequency changes are correlated to the behaviorally measured frequency change detection threshold, which has been reported to significantly affect CI speech performance (Zhang et al., 2019). We also hypothesize that the brain activation patterns are different in right- and left-ear CIs in patients who are right-handed.

EXPERIMENTAL SECTION

Participants

Twelve CI users (seven females, five males; 43–75 years old, with a mean age of 63 years) wearing the devices from Cochlear (Sydney, Australia) participated in this study. All participants were right-handed and native English speakers with no history of neurological or psychological disorders. They did not take medications that have been reported to affect the EEG. The CI users had severe-to-profound, bilateral, sensorineural hearing loss prior to implantation. Of the twelve CI subjects, seven subjects were bilateral CI users, four subjects were bimodal device users (one ear wore a CI and the non-implanted ear wore a hearing aid), and one subject was a unilateral CI user. Each CI ear was tested individually. In one bilateral CI user, only one CI ear was recorded, because the other CI ear was not able to detect the maximum magnitude of frequency changes in the psychoacoustic test. Therefore, both EEG and behavioral data were collected from a total of 18 ears (ten right-ear CIs and eight left-ear CIs). Individual CI subject's demographic information has been provided in Table 1 of Liang et al. (2018).

In addition, data from twelve normal hearing (NH) individuals (six females, six males; 20–30 years old, with the mean age of 23 years) were used to provide information on brain activity in individuals with a normal auditory system. The ACC waveform data from NH listeners have been reported in a previous study from our lab (Liang et al., 2016). The purpose of presenting the sLORETA data from NH listeners is not for direct comparison between the NH and CI group (since the stimuli were presented binaurally to the NH listeners and monaurally to each CI ear, Liang et al., 2016, 2018). Rather, the NH data presented here serves as a validation of the sLORETA methodology because the NH results could be compared to those in the literature. Moreover, discussion of the CI results, with the knowledge of NH results, would be beneficial for the understanding of the nature and degree of abnormalities in cortical processing of frequency changes in CI users. All NH listeners had audiometric hearing thresholds ≤20 dB HL at octave test frequencies from 250 to 8000 Hz. All participants gave informed written consent prior to their participation. This study was approved by the Institutional Review Board of the University of Cincinnati.

Stimuli

The stimuli were 160 Hz tones generated using Audacity[1] at a sample rate of 44.1 kHz. The duration of the tones was 1 s, including linear ramps of 10 ms at the onset and offset. The 160 Hz tone was selected because this frequency is in the range of the fundamental frequency (F_0) of the human voice (between the F_0 of female and male voiced speech, Gelfer and Bennett, 2013). The 160 Hz tone contained upward frequency changes of different magnitudes at 500 ms after the tone onset. The frequency change occurred for an integer number of cycles of the base frequency at 0 phase (i.e., zero crossing). Therefore, the onset cue of the frequency change was removed, and it did not produce audible transients (Dimitrijevic et al., 2008; Pratt et al., 2009). The stimuli were initially presented at 85 dB (peSPL) through a loudspeaker placed at ear level, 50 cm in front of the participant. CI users were tested using their typical everyday speech processor settings, but were allowed to adjust the volume so that the loudness level of the stimuli corresponded to the loudness level 7 (the most comfortable level) on a 0-10-point (inaudible to uncomfortably loud) numerical scale (Hoppe et al., 2001). The most comfortable level has been widely used in EEG studies involving CI users to minimize the loudness differences across CI users (Ponton et al., 1996; Friesen and Tremblay, 2006)." According to Pfingst et al. (1994), the mean frequency discrimination performance for CI users is improved when the loudness level increases up to level 4 and then does not change significantly up to loudness level 10.

[1] http://audacity.sourceforge.net

Psychoacoustic Test of the Frequency Change Detection Threshold

Participants were seated in a sound-treated booth for the testing. An adaptive, 2-alternative forced-choice procedure with an up-down stepping rule using APEX software (Francart et al., 2008) was employed to measure the frequency change detection threshold (FCDT). In each trial, a standard stimulus (the 160 Hz tone) and a target stimulus (the 160 Hz tone containing a frequency change with a magnitude of up to 65%; the step size was 5% from 10–65% range, 0.5% from 0.5–10% range, and 0.05% from 0.05–0.5% range) were included. The order of standard and target stimulus was randomized and the interval between the stimuli in a trial was 0.5 s. The participant was instructed to choose the target signal by pressing a button on a computer screen and was given a visual feedback regarding the response. Each run generated a total of five reversals. The asymptotic amount of frequency change (the average of the last three trials) was used as the FCDT. Each CI ear was tested separately. When one CI ear was tested for bilateral users, the opposite CI was turned off. Bimodal users were only tested on the CI side with the hearing aid on the contralateral side turned off and blocked with an earplug.

EEG Recording

Participants were seated on a comfortable chair in a sound-treated booth for the experiments. A 40-channel Neuroscan multi-channel EEG system (NuAmps, Compumedics Neuroscan, Inc., Charlotte, NC) was used to record the EEG. Electro-ocular activity (EOG) was monitored so that eye movement artifacts could be identified and rejected during the offline analysis. The continuous EEG data were recorded with a band-pass filter setting from 0.1 to 100 Hz and a sampling rate of 1,000 Hz. The average electrode impedance was lower than 10 kΩ. EEG signals from a total of 1–3 electrodes over the CI coil were not available.

During testing, participants were instructed to avoid excessive eye and body movements. Participants read self-selected magazines to keep alert and were asked to ignore the acoustic stimuli. A total of 400 trials of each of the three types of stimuli (three frequency changes: 0%, 5%, and 50%) were presented. The stimulus conditions were randomized to prevent order effects. The inter-stimulus interval was 800 ms. Each CI ear was tested separately.

Data Processing and Statistical Analysis

Waveform Analysis

The detailed procedures for waveform analysis were described in Liang et al. (2018). Briefly, continuous EEG data collected from each participant were digitally filtered (0.1–30 Hz) and then divided into segments over windows of -100 ms

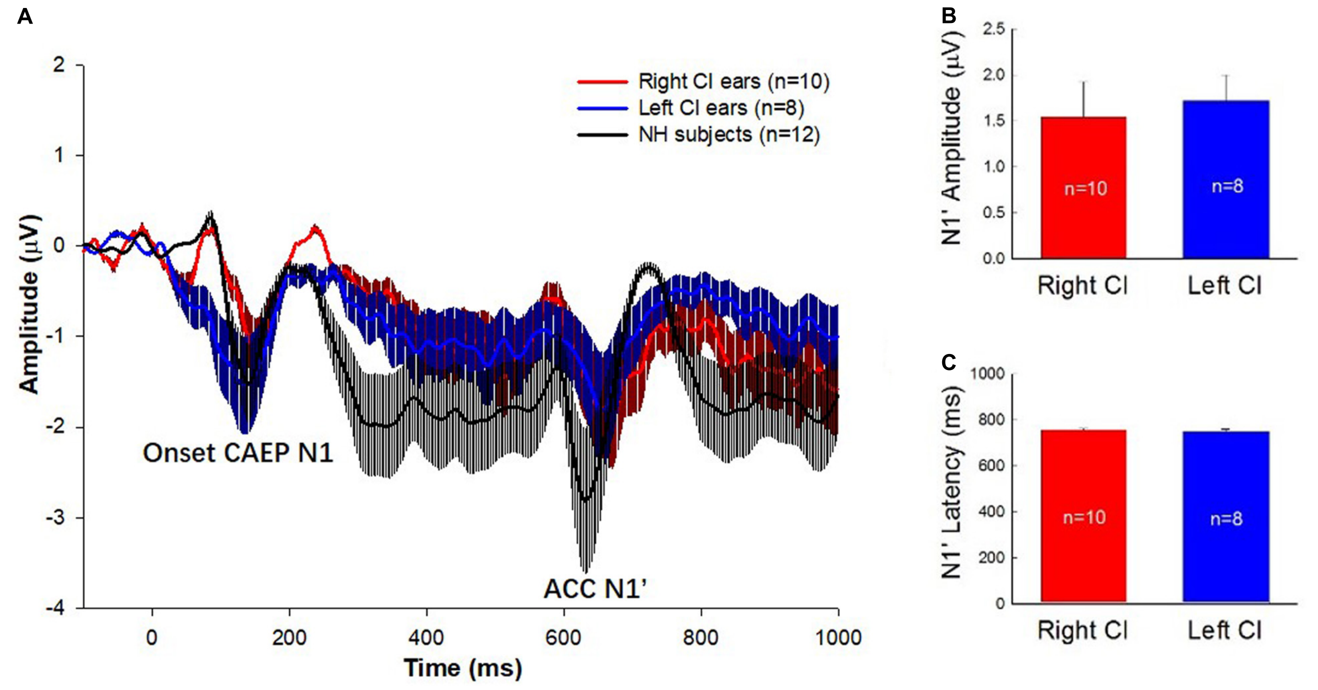

FIGURE 1 | (A) ERP waveforms with error bars for the right-ear CIs (red traces), left-ear CIs (blue traces), and the NH listeners (black traces) are displayed. The stimulus was a 1-s pure tone at 160 Hz that contained an upward frequency change of 50% in the middle of the stimulus. The N1 peak of the onset CAEP evoked by the tone onset and the N1' of the ACC evoked by the 50% frequency change are marked. Note that the N1' is the focus of this study. **(B,C)** The means and standard errors of the N1' amplitude and latency for left- and right-ear CIs. There is no significant difference in the amplitude and latency of N1' waves between right- and left-ear CIs.

to 1000 ms relative to the tone onset. Each data segment was visually inspected, and epochs contaminated by non-stereotyped artifacts were rejected and excluded from the analysis (Delorme and Makeig, 2004). There were at least 200 epochs left for each participant. Further data processing was performed by the use of the EEGLAB toolbox (Delorme and Makeig, 2004) running under MATLAB (MathWorks, United States). The data were baseline-corrected and then re-referenced using a common average reference. Independent component analysis (ICA) was then applied to identify and remove non-neurological activities (Gilley et al., 2006; Debener et al., 2008; Zhang et al., 2009, 2011). EEG from the electrodes close to the CI coil were replaced by linearly interpolated values computed from neighboring EEG signals. Then, the averaged waveform was derived for each type of stimuli (0%, 5%, and 50% frequency change) separately. The N1'peak of the ACC was identified in the range of 610–710 ms after stimulus onset or 110–210 ms after the occurrence of the frequency change.

Only EEG data elicited by the 50% frequency change were used for source density reconstruction because the ACC was present in all CI ears for the 50% frequency change, but was missing in some CI ears for the 5% frequency change and absent for the 0% change (see Liang et al., 2018). The presence of the ACC was determined based on criteria: (i) an expected wave morphology within the expected time window (approximately 610–710 ms after the tone onset) based on mutual agreement between two researchers (Lister et al., 2011; He et al., 2015); and (ii) a visual difference in the waveforms between the frequency change conditions vs. no change condition.

Sloreta Analysis

The waveform data were further imported into the sLORETA software package for source localization for the negative peak evoked by the frequency change (ACC N1'). The data from the right-ear CIs and left-ear CIs were analyzed separately. The following processes were conducted for sLORETA analysis: (1) The 3-dimensional sLORETA maps were generated to show the current source density (CSD) distribution patterns of ACC N1' (a timeframe ranging from 610 ms to 710 ms after stimulus onset); and (2) Regions of Interests (ROIs) were defined based on the activated brain regions observed in individual CIs. The correlations between the activities in the ROIs and the FCDT were also examined. The following ROIs of four brain regions were examined: the left and right temporal lobe (including Brodmann areas 21, 22, 38, 39, 41, and 42), and the left and right frontal lobe (including Brodmann areas 6, 9, 10, 11, 44, 45, 46, and 47) (Campbell and Sharma, 2013). These ROIs were selected because the greatest activities are in the temporal lobes and frontal lobes in the mean current source density distribution patterns of the ACC N1'. The ROI file for the 4 seed points for the center voxel was constructed. Each of the ROI values consisted of the mean current source density from

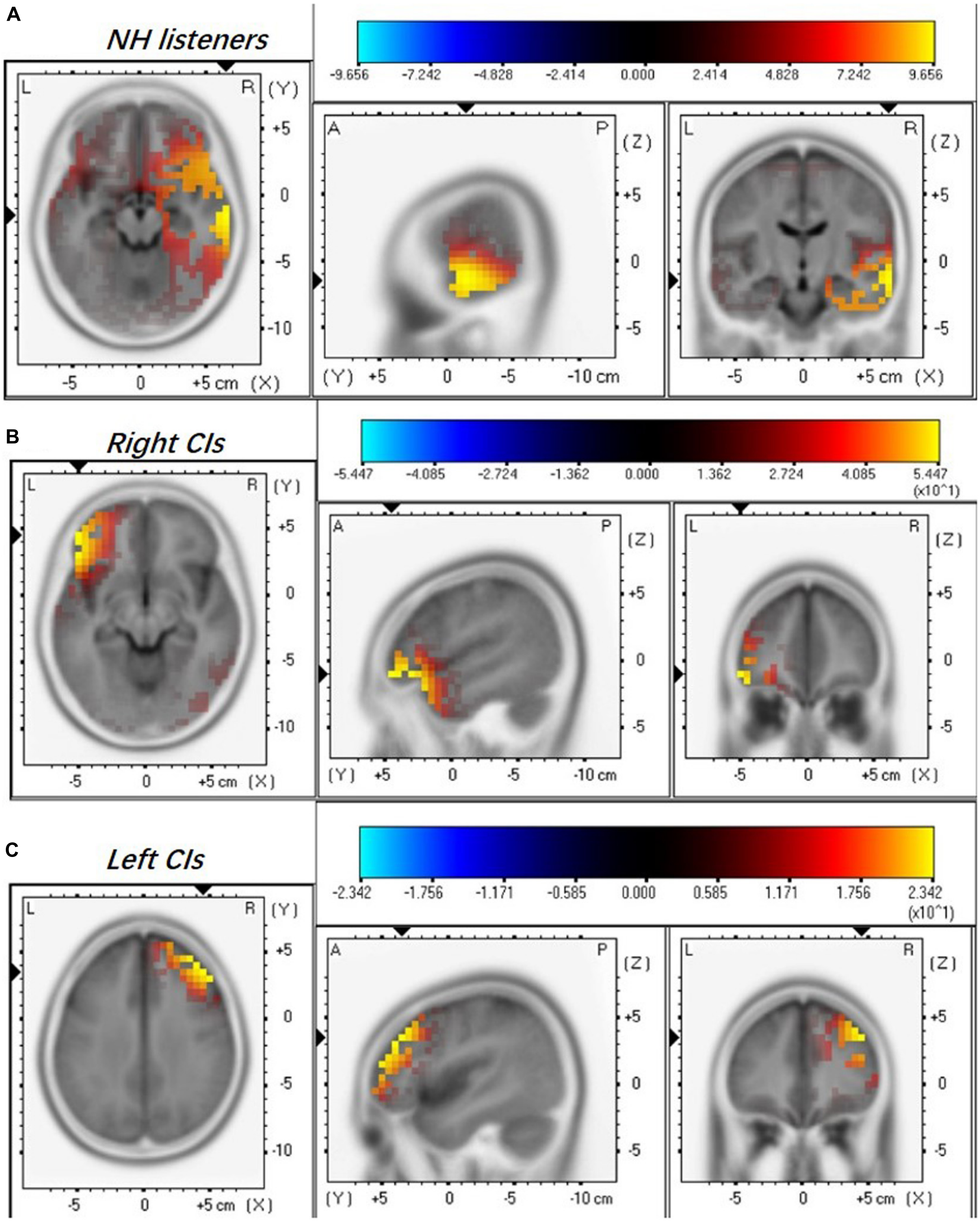

FIGURE 2 | Mean regional normalized sLORETA solutions modeling the distributed sources for ACC N1' of NH listeners **(A)**, right-ear CIs **(B)**, and left-ear CIs **(C)**. Yellow and blue colors represent increased and decreased current source density, respectively. Note that the NH listeners were stimulated binaurally and each CI ear was stimulated monaurally. There are differences in the activated brain regions among NH, right-, and left-ear CIs.

each ROI seed, including all cortical gray matter voxels within a 15-mm distance from the center. The resulting file produced the log transformed average CSD for all participants for each seed (Cannon and Baldwin, 2012). The CSD data for each brain region were imported into the Sigmaplot software program for statistical analyses.

Statistical Analysis

Data were presented as mean ± SEM and plotted with SigmaPlot v10 (SPSS Inc., Chicago, IL, United States). The statistical analyses were performed by Excel (Microsoft Office 365) or SPSS (SPSS Inc.) using *t* test or AVOVA with a Bonferroni correction for the *post hoc* tests. The level of statistical significance was set

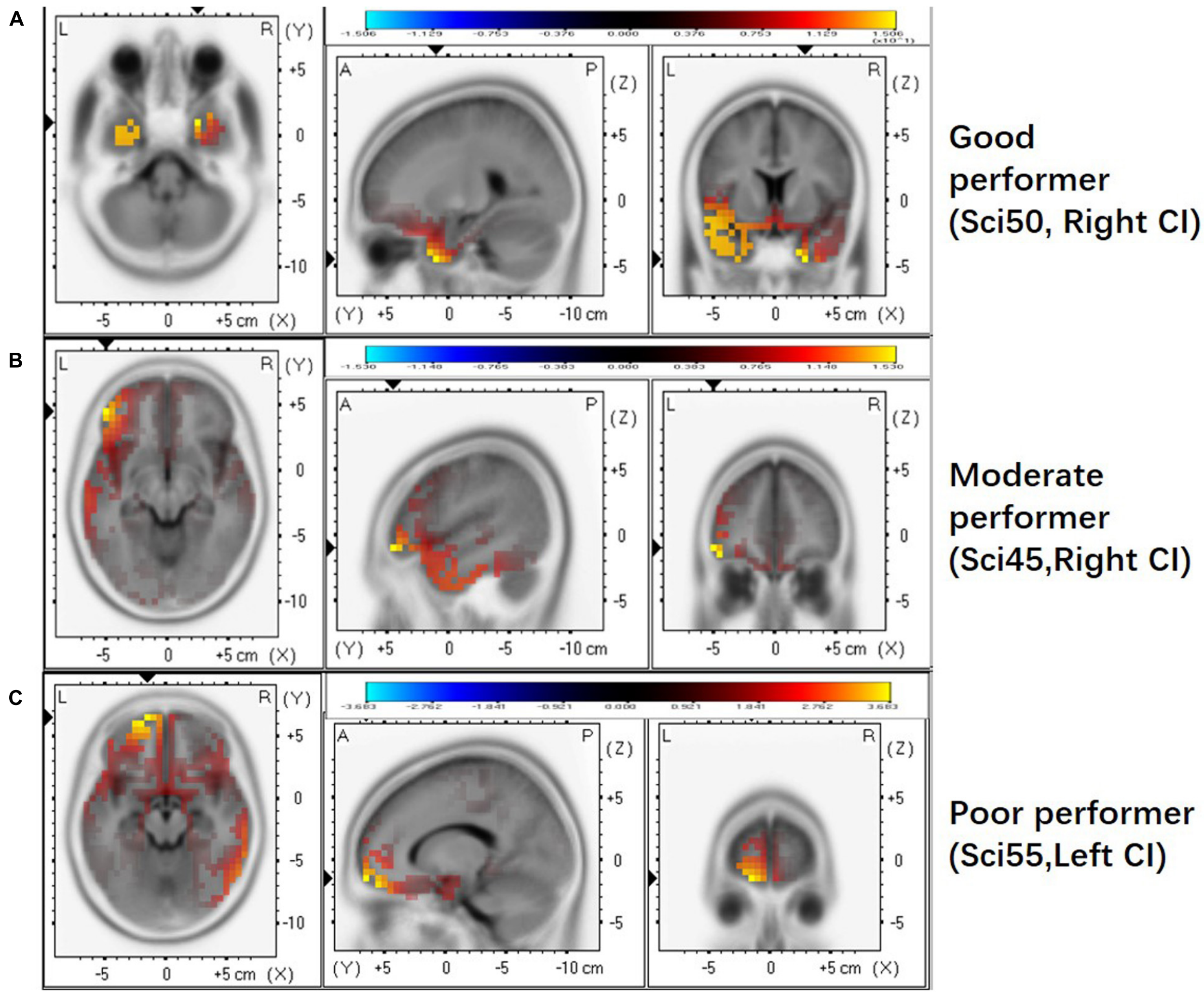

FIGURE 3 | The activated brain regions displayed for individual CI users with different FCDTs. **(A)** good performer (SCI 50 R, 1.83%). **(B)** poorer performer (SCI45 R, 5.17%). **(C)** poorest performer (SCI55 L, 9.33%). The strong activation in the temporal lobe appears in the good performer with a low FCDT.

at $p < 0.05$. All sample sizes and P values were reported in the figures or the figure legends.

RESULTS

Event-Related Potentials and ACC in Right-ear CIs, Left-ear CIs, and NH Subjects

Figure 1A shows the mean event-related potentials (ERPs, solid lines) and the standard errors at Cz evoked by the 160 Hz tones containing a 50% change for the right-ear CIs (red trace), left-ear CIs (blue trace), and NH listeners (black trace). As shown by the figure, there were two types of responses visible in the waveforms. One is an onset CAEP with the N1-P2 complex occurring approximately 110–210 ms after the onset of stimuli. The other one is the ACC with the N1'-P2' complex occurring approximately 610–710 ms after stimulus onset or 110–210 ms after the occurrence of the frequency change. There is a prominent difference in the amplitude and latency of ACC N1' between NH and CI subjects. **Figures 1B,C** show the means of N1' amplitude and latency for left- and right-ear CIs. The t-tests showed that there were no significant differences in N1' amplitude and latency between right- and left-ear CIs (t = 0.16, df = 16, p = 0.87 for latency; t = -0.35, df = 16, p = 0.73 for amplitude).

Difference in Activated Brain Regions in Right-ear CIs, Left-ear CIs, and NH Subjects

Current source density (CSD) maps derived using the sLORETA reflect the intensity of activation in the brain regions evoked by the presented stimulus. **Figure 2** illustrates the CSD patterns of the ACC N1' peak in the grand mean waveform for NH

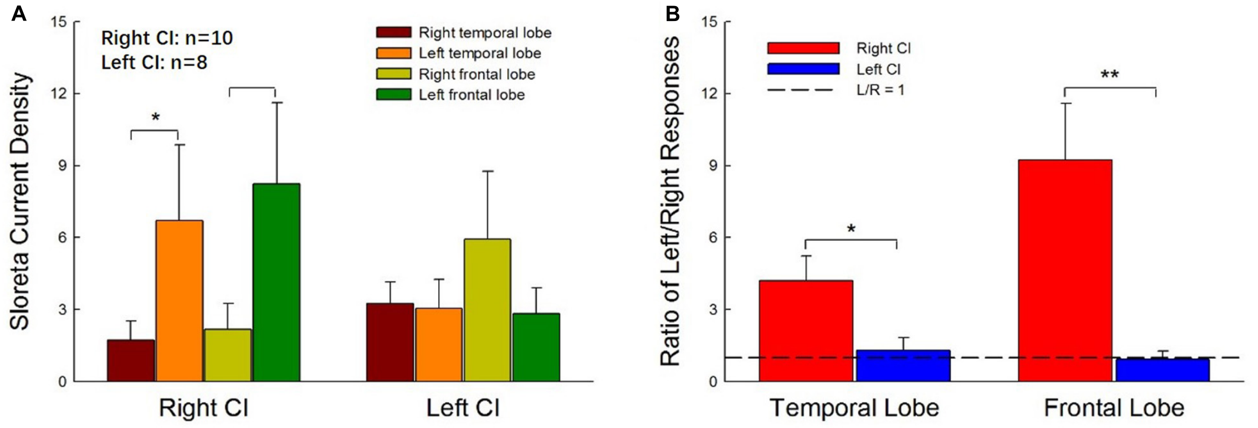

FIGURE 4 | Hemispheric differences in the current source density values for right- and left-ear CIs. **(A)** The current source density values in four regions of interest (ROIs, right, and left temporal lobes and frontal lobes) for right- and left-ear CIs. For the right-ear CIs, the current source densities in the contralateral hemisphere (left temporal lobe and left frontal lobe) are larger than those in the ipsilateral (right) hemisphere. The contralateral-dominant pattern is not observed in the left-ear CIs. **(B)** The left-dominant pattern is obvious in the right-ear CIs but not visible in the left-ear CIs. The left/right (L/R) ratios of responses in both the temporal lobe and frontal lobe for the right-ear CIs are greater than 1, whereas the L/R ratios of responses in both the temporal lobe and the frontal lobe for the left-ear CIs are close to 1. There are significant differences in the ratio of L/R responses between right-ear CIs and left-ear CIs ($^*p < 0.05$, $^{**}p < 0.01$, *t*-test).

listeners, right-ear CIs, and left-ear CIs. In NH listeners, the brain activation is mainly visible in the right temporal lobe (**Figure 2A**). However, the CI ears show an apparent difference in the activated brain regions (**Figures 2B,C**). For the right-ear CIs, the increase in CSD is visible in both the left temporal lobe and left frontal lobe (**Figure 2B**), while for the left-ear CIs the CSD increase is only visible in the right frontal lobe (**Figure 2C**).

Good Performers in Frequency Change Detection Show More Activation in the Temporal Lobe

We further examined whether the activated brain regions are associated with performance on the frequency change detection task. **Figure 3** shows brain activation patterns for ACC N1' peak in a good performer (Sci50R, FCDT: 1.83%), a moderate performer (Sci45R, FCDT: 5.17%), and a poor performer (Sci55L, FCDT: 9.33%) in the frequency change detection task. The good performer with low FCDTs showed more activation in the temporal lobe compared to the frontal lobe (**Figure 3A**), whereas the moderate and poor performers with high FCDTs had more activation in the non-temporal lobe including the frontal lobe (**Figures 3B,C**). The good or poor performance on the frequency change detection task was defined by comparing the FCDTs across all subjects tested. Our earlier publication using the FCDT task to examine the ability to detect frequency changes in CI users at base frequencies of 250, 1000, and 4000 Hz reported that poor performers (CNC, AzBio quiet and noise, and triple digit test) showed FCDTs of approximately 10% and above, while good performers showed a FCDT of approximately 1%. Using other methods (pitch discrimination/ranking), other studies also examined frequency discrimination in CI users. For example, Goldsworthy (2015) reported that CI users' pitch discrimination thresholds for pure tones (0.5, 1, and 2 kHz) ranged between 1.5 and 9.9%.

Difference in the Hemisphere Dominance Patterns of Current Source Density in Right- and Left-ear CIs

Figure 4 shows the mean CSD values in the four ROIs, the right and left temporal and frontal lobes, for right- and left-ear CIs. In general, stimulation in the right-ear CIs resulted in larger CSD values in the contralateral (left) hemisphere (**Figure 4A**) than in the ipsilateral (right) hemisphere. For the left-ear CIs, the activation in the contralateral (right) frontal lobe appeared to be stronger than that in the ipsilateral (left) side and the hemispheric difference did not exist in the temporal lobes. The t-tests showed that the activation is significantly stronger in the temporal and frontal lobes of the left side than that of the right side for the right-ear CIs (**Figure 4A**, $t = -2.23$, $df = 18$, $p = 0.04$ and $t = -2.31$, $df = 18$, $p = 0.03$, respectively, for the temporal and frontal lobe comparisons). One-way ANOVA did not show significant difference among the four ROIs for the left-ear CIs [$F_{(1,3)} = 1.54$, $p = 0.31$]. The hemispheric difference was further examined with the response ratio of the left/right (L/R ratio) hemispheres and the results were shown in **Figure 4B**. For the right-ear CIs, the L/R ratios in the temporal lobe and the frontal lobe were 4.20 ± 1.30 and 9.24 ± 2.35, respectively. For the left-ear CIs, the L/R ratios of responses in the temporal lobe and the frontal lobe were 1.30 ± 0.519 and 0.93 ± 0.34, respectively (**Figure 4B**), indicating that there was no apparent difference in the responses evoked in the left and right brain regions for the left-ear CIs. When comparing the right- and left-ear CIs, the L/R ratios of responses in both temporal lobe and frontal lobe for the right-ear CIs were significantly higher than these for the left-ear CIs (**Figure 4B**, $t = 2.50$, $df = 16$, $p = 0.026$ and $t = 3.13$,

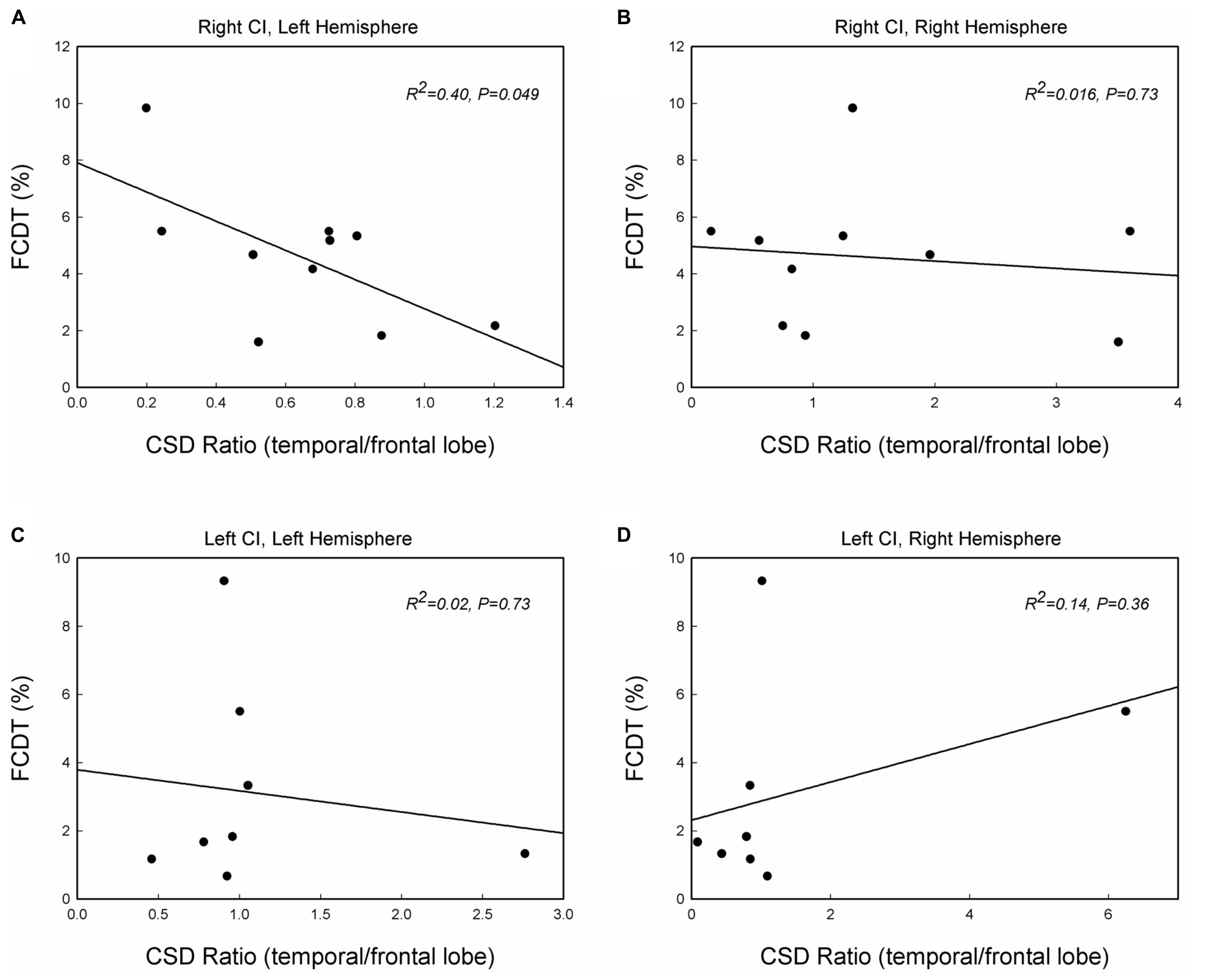

FIGURE 5 | The correlations between the FCDT and the ratio of current source density (CSD) values in the temporal lobe and frontal lobe in the left and right hemispheres for the right-ear CIs **(A,B)** and left-ear CIs **(C,D)**. Straight lines represent linear regression lines. Note that only the right-ear CIs in panel **(A)** shows a significant correlation, with good performers (low FCDTs) showing larger CSD ratios in the left hemisphere ($p < 0.05$).

$df = 16$, $p = 0.006$ for temporal and frontal lobes, respectively). In summary, the contralateral dominance of brain activation was observed for the right-ear CIs, but not for the left-ear CIs.

Correlation Between Activities in Temporal and Frontal Lobes and FCDT

Previous studies suggested that the auditory change detection depends on the interaction between the temporal and the frontal cortex (Doeller et al., 2003). We further examined the correlation between FCDT and the interactive activations in the temporal lobe and frontal lobe. **Figure 5** shows the scatter-plot of the FCDT vs. the CSD ratio (temporal lobe/frontal lobe) in the right and left hemispheres for right-ear and left-ear CIs. The CSD ratio in the left hemisphere for the right-ear CIs was negatively correlated to the FCDT ($R^2 = 0.40$, $p = 0.049$, **Figure 5A**). However, there was no significant correlation between FCDT and CSD ratio in the right hemisphere for the right-ear CIs. In addition, no statistically significant correlation was found for the FCDTs and the CSD-ratio for the left-ear CIs ($p > 0.05$. **Figures 5C,D**).

DISCUSSION

In this study, we examined the brain responses of right-ear CIs and left-ear CIs in right-handed adults. The primary findings were: (1) cortical activation patterns for the ACC-N1' are different in right-ear CIs and left-ear CIs (**Figures 1**, **2**); (2) Right-ear CIs could evoke stronger activities in the contralateral temporal and frontal lobes, but this contralateral hemisphere dominance was not observed for left-ear CIs (**Figures 3**, **4**), and (3) For the right-ear CIs, the increased activation in the left temporal lobe, along with the reduced activation in the frontal lobe (increased temporal/frontal CSD ratio), is correlated with good performance in the detection of frequency changes (**Figure 5**). Such a correlation was not found for the left-ear CIs. These results suggest that in adults, the implantation side can significantly influence the brain activation patterns evoked by the within-stimulus frequency changes.

Previous studies on brain activation patterns in CI users have focused on the onset CAEP, i.e., the cortical auditory evoked response to stimulus onset (Debener et al., 2008; Sandmann et al., 2009). Using dipole source analysis, Debener et al. (2008) reported the onset CAEP in one patient who had successfully used a CI in the right ear for four years. The results showed that the contralateral response is larger than the ipsilateral response. Sandmann et al. (2009) examined the onset CAEP to dyadic tones with pitch intervals that sound like music in CI users and NH listeners. While NH listeners showed a contralateral dominance effect for the left ear stimulation (Hine and Debener, 2007), CI users showed a contralateral dominance effect in the auditory regional source activity specifically for the right-ear stimulation. This suggested that the hemispheric asymmetry in CI users differs from that in NH listeners.

Brain activation patterns for the ACC N1' peak has not been studied in CI users. Unlike the onset-CAEP that indicates the cortical processing of stimulus onset, the ACC indicates the cortical processing of acoustic change embedded in a stimulus. In NH listeners, the processing of frequency changes is specialized in the right hemisphere, with either a binaural or monaural stimulation (Liegeois-Chauvel et al., 2001; Zatorre and Belin, 2001; Zatorre et al., 2002; Molholm et al., 2005; Schonwiesner et al., 2005; Dimitrijevic et al., 2008; Hyde et al., 2008; Itoh et al., 2012). When the acoustic change is small, the frontal lobe may also be activated (Opitz et al., 2002; Schönwiesner et al., 2007). Lesion studies also suggest that damage in the right hemisphere results in an impaired capability to discriminate frequencies, supporting that the right hemisphere has a specialized function of processing spectral changes (Sidtis and Volpe, 1988; Zatorre, 1988; Robin et al., 1990; Johnsrude et al., 2000). The activation in the right temporal lobe for the NH listeners' ACC N1' in the current study (**Figure 2**) is consistent with what has been reported in previous studies.

This study provides evidence that the temporal lobe and frontal lobe are involved in frequency change detection in CI users. This study also shows that CI users demonstrated brain activation patterns for the ACC N1' that are different from those in NH listeners. While the right-ear CIs result in activation in the contralateral temporal and frontal lobes, the left-ear CIs generate less activation in the contralateral temporal lobe (**Figures 3**, **4**). The differences in the brain activation patterns in the CI users relative to the NH listeners may arise from CI technology limitations and/or brain reorganization in CI users. Specifically, the CI delivers the sound information through a limited number of frequency channels, resulting in a dramatic decrease in CI users' frequency resolution. Such a compromised frequency resolution is exacerbated by the neural deficits related to the long-term deafness prior to the implantation. In this study, 160 Hz was used because this frequency is in the frequency range of the fundamental frequency (F0). As this frequency is located at the filter slope of the first band at the CI speech processing stage, the detection of a 50% frequency change (from 160 Hz to 240 Hz) may rely on the temporal rate cues as the result of different signal intensities on the filter response curve in the CI output, and/or the place cues at a result of activating different electrodes. Therefore, we speculate that frequency change detection relies on both temporal and spectral cues in CI users. The ACC results from CI users in the current study suggested that the auditory brain can automatically encode the 50% frequency change from the 160 Hz, although the activation pattern may differ from that in NH listeners.

We found that the CI-activated brain regions included the temporal lobe as well as non-temporal regions, such as the frontal lobe (**Figures 2–4**). Moreover, the brain activation and activated patterns were related to the behaviorally measured FCDT (**Figures 3**, **5**). Specifically, the right-ear CIs with good performance had similar brain activation patterns as the NH listeners, mainly activating the temporal lobe (**Figures 3**, **4**). The CI ears with poor performance were likely to have activation in non-temporal areas (**Figure 3**). The ears with better performance had a greater temporal lobe/frontal lobe CSD ratio (stronger activation in the temporal lobe and weaker activation in the frontal lobe) on the contralateral side for the right-ear CIs (**Figure 5A**). Green et al. (2005) reported that the activation strength in the temporal lobe is positively correlated with the speech perception performance (Green et al., 2005). In future studies, it is necessary to evaluate the correlations between the ACC brain activation and speech perception performance in CI users.

The major finding in this study is that in comparison with left-ear CIs, right-ear CIs can more efficiently activate the temporal lobe in the contralateral hemisphere (**Figures 3**, **4**). Previous studies examining the onset cortical responses have reported that the contralateral dominance is greater in CI users using a right CI compared to a left CI (Debener et al., 2008; Sandmann et al., 2009). Studies using neuroimaging techniques, such as positron emission tomography (PET), also showed a higher activation in the cortex contralateral to the CI than that on the ipsilateral side (Herzog et al., 1991). In this study, we not only used sLORETA to locate activated brain regions but also examined the behavioral measure of frequency change detection. We found that, in comparison with left-ear CIs, the right-ear CIs are more efficient in activating the contralateral brain regions (**Figures 3–5**).

It is unknown why the right-ear CIs in this study demonstrated more activity in the contralateral hemisphere than in the ipsilateral hemisphere, while the left-ear CI s did not show this contralateral dominance. One possibility is that such a phenomenon is related to the fact that all tested subjects in this study were right-handed. Brain activations evoked by acoustic stimuli show different patterns for right-handed and left-handed NH individuals (Provins and Jeeves, 1975; Maehara et al., 1999). It is possible that the pre-existing hemispheric differences for right-handed individuals before implantation may be differentially affected by right- vs. left-ear implantation. A review study (Kraaijenga et al., 2018) based on 20 articles that were eligible for critical evaluation of the effect of the implantation side on CI outcomes reported that the majority of studies reveal evidence for a right-ear advantage in post-lingually deafened adults. The right-ear advantage was also reported in prelingually deafened children wearing CIs (Henkin et al., 2014). Although it is too premature to advise implanting in the right ear, our findings in the current study do support the idea that cortical processing of frequency changes show different brain activation patterns between the right- and left-ear CIs, at least in right-handed adults.

CONCLUSION AND FUTURE STUDIES

This study reveals that the side of implantation can significantly influence the brain activation patterns evoked by frequency changes in adults with right-handedness. Right-ear CIs result in stronger brain activities in the contralateral hemisphere than in the ipsilateral hemisphere. Such a contralateral dominance was not observed for left-ear CIs. For right-ear CIs, good performance in frequency change detection is correlated with larger temporal activation, along with weaker frontal activation. The findings of this study provide valuable evidence that the right-ear implantation appears to support the contralateral dominance for right-handed patients. The data also demonstrated that the sLORETA is a promising source location approach and can be used to longitudinally examine brain plasticity after cochlear implantation, to examine the development of auditory cortex reorganization, to actively guide rehabilitation strategies, and to monitor the progress of an individual during rehabilitation.

There are some possible directions for future studies. First, we will further examine the brain activation patterns to frequency changes in both left-handed and right-handed CI patients with a larger sample size. With a larger sample size, the correlation observed in **Figure 5** of this study might be stronger; second, it would be valuable to add NH listeners who are age-, gender-matched, and test to their individual ears corresponding to those in their CI counterparts. Finally, the CI outcomes in terms of speech perception will be obtained to establish the brain-behavioral relationships.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding authors.

ETHICS STATEMENT

This study involving human participants was reviewed and approved by Fawen Zhang's affiliated institution, University of Cincinnati. The patients/participants provided their written informed consents to participate in this study.

AUTHOR CONTRIBUTIONS

CL and FZ conceptualized, formally collected and analyzed the data. CL, FZ, RS, JX, and LW contributed to patient recruitment and methodology. CL and FZ prepared and wrote the manuscript. All authors reviewed and edited the manuscript.

FUNDING

This research was partially supported by the University Research Council (URC) grant at the University of Cincinnati, the Center for Clinical and Translational Science and Training (CCTST) grant, and the National Institute of Health R15 grant (NIH 1 R15 DC016463-01) to FZ. This project is also a part of the dissertation work of CL at the University of Cincinnati. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health and other funding agencies.

ACKNOWLEDGMENTS

We thank all participants for their participation in this research.

REFERENCES

Abbas, P. J., and Brown, C. J. (2014). Assessment of responses to cochlear implant stimulation at different levels of the auditory pathway. *Hear. Res.* 322, 67–76. doi: 10.1016/j.heares.2014.10.011

Blamey, P., Artieres, F., Başkent, D., Bergeron, F., Beynon, A., Burke, E., et al. (2012). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: an update with 2251 patients. *Audiol. Neurotol.* 18, 36–47. doi: 10.1159/000343189

Brown, C. J., Jeon, E. K., Chiou, L.-K., Kirby, B., Karsten, S. A., Turner, C. W., et al. (2015). Cortical auditory evoked potentials recorded from nucleus hybrid cochlear implant users. *Ear Hear.* 36, 723–732. doi: 10.1097/AUD.0000000000000206

Campbell, J., and Sharma, A. (2013). Compensatory changes in cortical resource allocation in adults with hearing loss. *Front. Syst. Neurosci.* 7:71. doi: 10.3389/fnsys.2013.00071

Cannon, R. L., and Baldwin, D. R. (2012). EEG current source density and the phenomenology of the default network. *Clin. EEG Neurosci.* 43, 257–267. doi: 10.1177/1550059412449780

Debener, S., Hine, J., Bleeck, S., and Eyles, J. (2008). Source localization of auditory evoked potentials after cochlear implantation. *Psychophysiology* 45, 20–24. doi: 10.1111/j.1469-8986.2007.00610.x

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Dimitrijevic, A., Michalewski, H. J., Zeng, F. G., Pratt, H., and Starr, A. (2008). Frequency changes in a continuous tone: auditory cortical potentials. *Clin. Neurophysiol.* 119, 2111–2124. doi: 10.1016/j.clinph.2008.06.002

Doeller, C. F., Opitz, B., Mecklinger, A., Krick, C., Reith, W., and Schröger, E. (2003). Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. *NeuroImage* 20, 1270–1282. doi: 10.1016/S1053-8119(03)00389-6

Eugene, A. R., Masiak, J., and Kapica, J. (2014). Electrophysiological neuroimaging using sLORETA comparing 12 anorexia nervosa patients to 12 controls. *Brain (Bacau)* 5, 26–33. doi: 10.1037/emo0000122.Do

Finley, C. C., and Skinner, M. W. (2008). Role of electrode placement as a contributor to variability in cochlear implant outcomes Charles. *Otol Neurotol.* 29, 920–928. doi: 10.1097/MAO.0b013e318184f492.Role

Francart, T., van Wieringen, A., and Wouters, J. (2008). APEX 3: a multi-purpose test platform for auditory psychophysical experiments. *J. Neurosci. Methods* 172, 283–293. doi: 10.1016/j.jneumeth.2008.04.020

Friesen, L. M., and Tremblay, K. L. (2006). Acoustic change complexes recorded in adult cochlear implant listeners. *Ear Hear.* 27, 678–685. doi: 10.1097/01.aud.0000240620.63453.c3

Fulbright, A. N. C., Le Prell, C. G., Griffiths, S. K., and Lobarinas, E. (2017). Effects of recreational noise on threshold and suprathreshold measures of auditory function. *Semin. Hear.* 38, 298–318. doi: 10.1055/s-0037-1606325

Gelfer, M. P., and Bennett, Q. E. (2013). Speaking fundamental frequency and vowel formant frequencies: effects on perception of gender. *J. Voice* 27, 556–566. doi: 10.1016/j.jvoice.2012.11.008

Gifford, R. H., Hedley-Williams, A., and Spahr, A. J. (2014). Clinical assessment of spectral modulation detection for adult cochlear implant recipients: a non-language based measure of performance outcomes. *Int. J. Audiol.* 53, 159–164. doi: 10.3109/14992027.2013.851800

Gilley, P. M., Sharma, A., Dorman, M., Finley, C. C., Panch, A. S., and Martin, K. (2006). Minimization of cochlear implant stimulus artifact in cortical auditory evoked potentials. *Clin. Neurophysiol.* 117, 1772–1782. doi: 10.1016/j.clinph.2006.04.018

Giraud, A. L., Price, C. J., Graham, J. M., and Frackowiak, R. S. (2001). Functional plasticity of language-related brain areas after cochlear implantation. *Brain* 124(Pt 7), 1307–1316. doi: 10.1093/brain/124.7.1307

Goldsworthy, R. L. (2015). Correlations between pitch and phoneme perception in cochlear implant users and their normal hearing peers. *J. Assoc. Res. Otolaryngol.* 16, 797–809. doi: 10.1007/s10162-015-0541-9

Gordon, K. A., Wong, D. D. E., and Papsin, B. C. (2010). Cortical function in children receiving bilateral cochlear implants simultaneously or after a period of interimplant delay. *Otol. Neurotol.* 31, 1293–1299. doi: 10.1097/MAO.0b013e3181e8f965

Grech, R., Cassar, T., Muscat, J., Camilleri, K. P., Fabri, S. G., Zervakis, M., et al. (2008). Review on solving the inverse problem in EEG source analysis. *J. Neuroeng. Rehabil.* 5:25. doi: 10.1186/1743-0003-5-25

Green, K. M. J., Julyan, P. J., Hastings, D. L., and Ramsden, R. T. (2005). Auditory cortical activation and speech perception in cochlear implant users: effects of implant experience and duration of deafness. *Hear. Res.* 205, 184–192. doi: 10.1016/j.heares.2005.03.016

He, S., Grose, J. H., and Buchman, C. A. (2012). Auditory discrimination: the relationship between psychophysical and electrophysiological measures. *Int. J. Audiol.* 51, 771–782. doi: 10.3109/14992027.2012.699198

He, S., Grose, J. H., Teagle, H. F. B., Woodard, J., Park, L. R., Hatch, D. R., et al. (2015). Acoustically evoked auditory change complex in children with auditory neuropathy spectrum disorder: a potential objective tool for identifying cochlear implant candidates. *Ear Hear.* 36, 289–301. doi: 10.1097/AUD.0000000000000119

Henkin, Y., Swead, R. T., Roth, D. A.-E., Kishon-Rabin, L., Shapira, Y., Migirov, L., et al. (2014). Evidence for a right cochlear implant advantage in simultaneous bilateral cochlear implantation. *Laryngoscope* 124, 1937–1941. doi: 10.1002/lary.24635

Herzog, H., Lamprecht, A., Kuhn, A., Roden, W., Vosteen, K. H., and Feinendegen, L. E. (1991). Cortical activation in profoundly deaf patients during cochlear implant stimulation demonstrated by H2(15)O PET. *J. Comput. Assist Tomogr.* 15, 369–375. doi: 10.1097/00004728-199105000-00005

Hine, J., and Debener, S. (2007). Late auditory evoked potentials asymmetry revisited. *Clin. Neurophysiol.* 118, 1274–1285. doi: 10.1016/j.clinph.2007.03.012

Holden, L. K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., Potts, L. G., et al. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear Hear.* 34, 342–360. doi: 10.1097/AUD.0b013e3182741aa7

Hoppe, U., Rosanowski, F., Iro, H., and Eysholdt, U. (2001). Loudness perception and late auditory evoked potentials in adult cochlear implant users. *Scand. Audiol.* 30, 119–125. doi: 10.1080/010503901300112239

Hyde, K. L., Peretz, I., and Zatorre, R. J. (2008). Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia* 46, 632–639. doi: 10.1016/j.neuropsychologia.2007.09.004

Itoh, K., Okumiya-Kanke, Y., Nakayama, Y., Kwee, I. L., and Nakada, T. (2012). Effects of musical training on the early auditory cortical representation of pitch transitions as indexed by change-N1. *Eur. J. Neurosci.* 36, 3580–3592. doi: 10.1111/j.1460-9568.2012.08278.x

Johnsrude, I. S., Penhune, V. B., and Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain* 123, 155–163. doi: 10.1093/brain/123.1.155

Kenway, B., Tam, Y. C., Vanat, Z., Harris, F., Gray, R., Birchall, J., et al. (2015). Pitch discrimination: an independent factor in cochlear implant performance outcomes. *Otol. Neurotol.* 36, 1472–1479. doi: 10.1097/MAO.0000000000000845

Kim, J.-R. (2015). Acoustic change complex: clinical implications. *Korean J. Audiol.* 19, 120–124. doi: 10.7874/jao.2015.19.3.120

Kraaijenga, V. J. C., Derksen, T. C., Stegeman, I., and Smit, A. L. (2018). The effect of side of implantation on unilateral cochlear implant performance in patients with prelingual and postlingual sensorineural hearing loss: a systematic review. *Clin. Otolaryngol.* 43, 440–449. doi: 10.1111/coa.12988

Lazar, D. S., Bordure, P., Lina-Granade, G., Magnan, J., Meller, R., Meyer, B., et al. (2010). Speech perception performance for 100 post-lingually deaf adults fitted with Neurelec cochlear implants: comparison between Digisonic® Convex and Digisonic® SP devices after a 1-year follow-up. *Acta Oto-Laryngol.* 130, 1267–1273. doi: 10.3109/00016481003769972

Lazard, D. S., Giraud, A. L., Gnansia, D., Meyer, B., and Sterkers, O. (2012). Understanding the deafened brain: implications for cochlear implant rehabilitation. *Eur. Ann. Otorhinolaryngol. Head Neck Dis.* 129, 98–103. doi: 10.1016/j.anorl.2011.06.001

Lee, J., Nadol, J. B., and Eddington, D. K. (2010). Depth of electrode insertion and postoperative performance in humans with cochlear implants: a histopathologic study. *Audiol. Neurotol.* 15, 323–331. doi: 10.1159/000289571

Liang, C., Earl, B., Thompson, I., Whitaker, K., Cahn, S., Xiang, J., et al. (2016). Musicians are better than non-musicians in frequency change detection: behavioral and electrophysiological evidence. *Front. Neurosci.* 10:464. doi: 10.3389/fnins.2016.00464

Liang, C., Houston, L. M., Samy, R. N., Abedelrehim, L. M. I., and Zhang, F. (2018). Cortical processing of frequency changes reflected by the acoustic change complex in adult cochlear implant users. *Audiol. Neurotol.* 23, 152–164. doi: 10.1159/000492170

Liegeois-Chauvel, C., Giraud, K., Badier, J.-M., Marquis, P., and Chauvel, P. (2001). "Intracerebral evoked potentials in pitch perception reveal a function of asymmetry of human auditory cortex," in *The Cognitive Neuroscience of Music*, eds U. Montreal, et al. (London: Oxford University Press), 452.

Limb, C. J., and Roy, A. T. (2014). Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hear. Res.* 308, 13–26. doi: 10.1016/j.heares.2013.04.009

Lister, J. J., Maxfield, N. D., Pitt, G. J., and Gonzalez, V. B. (2011). Auditory evoked response to gaps in noise: older adults. *Int. J. Audiol.* 50, 211–225. doi: 10.3109/14992027.2010.526967

Maehara, T., Ozaki, H., Wakabayashi, C., Iizuka, Y., and Hosokawa, A. (1999). Functional magnetic resonance imaging of auditory cortex: with special reference to the side of aural stimulation. *Radiat. Med.* 17, 145–149.

Molholm, S., Martinez, A., Ritter, W., Javitt, D. C., and Foxe, J. J. (2005). The neural circuitry of pre-attentive auditory change-detection: An fMRI study of pitch and duration mismatch negativity generators. *Cereb. Cortex* 15, 545–551. doi: 10.1093/cercor/bhh155

Nash-Kille, A., and Sharma, A. (2014). Inter-trial coherence as a marker of cortical phase synchrony in children with sensorineural hearing loss and auditory neuropathy spectrum disorder fitted with hearing aids and cochlear implants. *Clin. Neurophysiol.* 125, 1459–1470. doi: 10.1016/j.clinph.2013.11.017

Okamoto, H., Stracke, H., Draganova, R., and Pantev, C. (2009). Hemispheric asymmetry of auditory evoked fields elicited by spectral versus temporal stimulus change. *Cereb. Cortex* 19, 2290–2297. doi: 10.1093/cercor/bhn245

Opitz, B., Rinne, T., Mecklinger, A., von Cramon, D. Y. Y., Schröger, E., Doeller, C. F., et al. (2002). Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *NeuroImage* 15, 167–174. doi: 10.1006/nimg.2001.0970

Ostroff, J. M., Martin, B. A., and Boothroyd, A. (1998). Cortical evoked response to acoustic change within a syllable. *Ear Hear.* 19, 290–297. doi: 10.1097/00003446-199808000-00004

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Findings Exp. Clin. Pharmacol.* 24D, 5–12.

Pfingst, B. E., Holloway, L. A., Poopat, N., Subramanya, A. R., Warren, M. F., and Zwolan, T. A. (1994). Effects of stimulus level on nonspectral frequency discrimination by human subjects. *Hear. Res.* 78, 197–209. doi: 10.1016/0378-5955(94)90026-4

Pisoni, D. B., Kronenberger, W. G., Chandramouli, S. H., and Conway, C. M. (2016). Learning and memory processes following cochlear implantation: the missing piece of the puzzle. *Front. Psychol.* 7:493. doi: 10.3389/fpsyg.2016.00493

Plummer, C., Wagner, M., Fuchs, M., Vogrin, S., Litewka, L., Farish, S., et al. (2010). Clinical utility of distributed source modelling of interictal scalp EEG in focal epilepsy. *Clin. Neurophysiol.* 121, 1726–1739. doi: 10.1016/j.clinph.2010.04.002

Ponton, C. W., Don, M., Eggermont, J. J., Waring, M. D., and Masuda, A. (1996). Maturation of human cortical auditory function: Differences between normal-hearing children and children with cochlear implants. *Ear Hear.* 17, 430–437. doi: 10.1097/00003446-199610000-00009

Pratt, H., Starr, A., Michalewski, H. J., Dimitrijevic, A., Bleich, N., and Mittelman, N. (2009). Auditory-evoked potentials to frequency increase and decrease of high- and low-frequency tones. *Clin. Neurophysiol.* 120, 360–373. doi: 10.1016/j.clinph.2008.10.158

Provins, K. A., and Jeeves, M. A. (1975). Hemisphere differences in response time to simple auditory stimuli. *Neuropsychologia* 13, 207–211. doi: 10.1016/0028-3932(75)90029-9

Robin, D. A., Tranel, D., and Damasio, H. (1990). Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain Lang.* 39, 539–555. doi: 10.1016/0093-934X(90)90161-9

Sandmann, P., Eichele, T., Buechler, M., Debener, S., Jäncke, L., Dillier, N., et al. (2009). Evaluation of evoked potentials to dyadic tones after cochlear implantation. *Brain* 132, 1967–1979. doi: 10.1093/brain/awp034

Schönwiesner, M., Novitski, N., Pakarinen, S., Carlson, S., Tervaniemi, M., and Näätänen, R. (2007). Heschl's gyrus, posterior superior temporal gyrus, and mid-ventrolateral prefrontal cortex have different roles in the detection of acoustic changes. *J. Neurophysiol.* 97, 2075–2082. doi: 10.1152/jn.01083.2006

Schonwiesner, M., Rübsamen, R., and Von Cramon, D. Y. (2005). Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *Eur. J. Neurosci.* 22, 1521–1528. doi: 10.1111/j.1460-9568.2005.04315.x

Sidtis, J. J., and Volpe, B. T. (1988). Selective loss of complex-pitch or speech discrimination after unilateral lesion. *Brain Lang.* 34, 235–245. doi: 10.1016/0093-934X(88)90135-6

Song, J. J., Punte, A. K., De Ridder, D., Vanneste, S., and Van de Heyning, P. (2013). Neural substrates predicting improvement of tinnitus after cochlear implantation in patients with single-sided deafness. *Hear. Res.* 299, 1–9. doi: 10.1016/j.heares.2013.02.001

Talja, S., Alho, K., and Rinne, T. (2015). Source analysis of event-related potentials during pitch discrimination and pitch memory tasks. *Brain Topogr.* 28, 445–458. doi: 10.1007/s10548-013-0307-9

Wagner, M., Fuchs, M., and Kastner, J. (2004). Evaluation of sLORETA in the presence of noise and multiple sources. *Brain Topogr.* 16, 277–280. doi: 10.1023/B:BRAT.0000032865.58382.62

Won, J. H., Humphrey, E. L., Yeager, K. R., Martinez, A. A., Robinson, C. H., Mills, K. E., et al. (2014). Relationship among the physiologic channel interactions, spectral-ripple discrimination, and vowel identification in cochlear implant users. *J. Acoust. Soc. Am.* 136, 2714. doi: 10.1121/1.4895702

Wong, D. D. E., and Gordon, K. A. (2009). Beamformer suppression of cochlear implant artifacts in an electroencephalography dataset. *IEEE Trans. Biomed. Eng.* 56, 2851–2857. doi: 10.1109/TBME.2009.2029239

Worrell, G. A., Lagerlund, T. D., Sharbrough, F. W., Brinkmann, B. H., Busacker, N. E., Cicora, K. M., et al. (2000). Localization of the epileptic focus by low-resolution electromagnetic tomography in patients with a lesion demonstrated by MRI. *Brain Topogr.* 12, 273–282. doi: 10.1023/A:1023407521772

Zatorre, R. J. (1988). Pitch perception of complex tones and human temporal-lobe function. *J. Acoust. Soc. Am.* 84, 566–572. doi: 10.1121/1.396834

Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953. doi: 10.1093/cercor/11.10.946

Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cognit. Sci.* 6, 37–46. doi: 10.1016/S1364-6613(00)01816-7

Zhang, F., Deshpande, A., Benson, C., Smith, M., Eliassen, J., and Fu, Q. J. (2011). The adaptive pattern of the auditory N1 peak revealed by standardized low-resolution brain electromagnetic tomography. *Brain Res.* 1400, 42–52. doi: 10.1016/j.brainres.2011.05.036

Zhang, F., Samy, R. N., Anderson, J. M., and Houston, L. (2009). Recovery function of the late auditory evoked potential in cochlear implant users and normal-hearing listeners. *J. Am. Acad. Audiol.* 20, 397–408. doi: 10.3766/jaaa.20.7.2

Zhang, F., Underwood, G., McGuire, K., Liang, C., Moore, D. R., and Fu, Q.-J. (2019). Frequency change detection and speech perception in cochlear implant users. *Hear. Res.* 379, 12–20. doi: 10.1016/j.heares.2019.04.007

Conflict of Interest: RS and LW received research and travel funding from Cochlear Ltd. The funders had no role in the design of the study, the collection, analyses, or interpretation of data, the writing of the manuscript, or the decision to publish the results.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

MINI REVIEW
published: 21 May 2020
doi: 10.3389/fnins.2020.00447

Cochlear Implantation Outcomes in Patients With *OTOF* Mutations

Dandan Zheng and Xiao Liu*

Department of Otorhinolaryngology, 2nd Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China

Auditory neuropathy is a special type of hearing loss caused by dysfunction of the synapse of the inner hair cells, the auditory nerve, and/or the auditory nerve itself. For patients with auditory neuropathy who have severe to profound hearing loss or failed auditory skills development with hearing-aids, cochlear implantation (CI) serves as the only possible effective treatment. It is accepted that the exact sites of lesion causing auditory neuropathy determine the CI performance. Mutations in the *OTOF* gene were the first identified and the most common cause of congenital auditory neuropathy. The site of lesion in patients with auditory neuropathy caused by biallelic *OTOF* mutations (*OTOF*-related auditory neuropathy) is presumed to be presynaptic, leaving auditory nerve function intact. Thus, *OTOF*-related auditory neuropathy is expected to have good CI performances. In this review, we describe the CI outcomes in patients with *OTOF* mutations. We will focus on whether biallelic *OTOF* mutations are ideal indications for CI in patients with auditory neuropathy. Also, the factors that may still influence the CI outcomes in patients with *OTOF* mutations are discussed.

Keywords: auditory neuropathy, cochlear implantation, *OTOF*, rehabilitation, outcomes

OPEN ACCESS

Edited by:
Virginia Penhune,
Concordia University, Canada

Reviewed by:
Anu Sharma,
University of Colorado Boulder, United States
Shujiro Minami,
Tokyo Medical Center (NHO), Japan
Ignacio del Castillo,
Hospital Universitario Ramón y Cajal, Spain

****Correspondence:***
Xiao Liu
11618337@zju.edu.cn

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *06 December 2019*
Accepted: *09 April 2020*
Published: *21 May 2020*

Citation:
Zheng D and Liu X (2020) Cochlear Implantation Outcomes in Patients With OTOF Mutations. Front. Neurosci. 14:447. doi: 10.3389/fnins.2020.00447

INTRODUCTION

Auditory neuropathy is a type of hearing loss caused by dysfunction of the synapse of the inner hair cells, the auditory nerve, and/or the auditory nerve itself (Hayes and Sininger, 2008). Individuals with auditory neuropathy typically show normal or near-normal otoacoustic emission (OAE) or cochlear microphonics (CM), but absent or abnormal auditory brainstem response (ABR) and/or middle ear muscle reflexes, usually accompanied by poor speech discrimination scores and poor understanding (Starr et al., 1996). For hearing rehabilitation in patients with auditory neuropathy, both cochlear implantation (CI) and wearing hearing-aids (HA) are options. However, in patients who have failed auditory skills development with HA or who are with severe to profound hearing loss, CI is considered the only possible effective treatment (Yawn et al., 2019).

As the transmission of the signal from electrical stimulation of the spiral ganglion provided by the cochlear implants could be affected, it had been thought that the CI outcomes were relatively poor in patients with auditory neuropathy (Starr et al., 1996). Recently, however, studies showed that CI could help to develop auditory skill in some of the patients with auditory neuropathy, yet the benefits were uncertain due to a wide range of etiologies (Rance et al., 1999; Berlin et al., 2010; Humphriss et al., 2013). The cause of auditory neuropathy includes loss of inner hair cells (IHCs) or IHC ribbon synapses, impaired synaptic transmission to spiral ganglion neurons (SGNs), and disrupted propagation of auditory information along the auditory nerve (Moser and Starr, 2016). It is obvious that the exact sites of lesion causing auditory neuropathy determine the CI performance. That is, lesions located in the membranous labyrinth (presynaptic) are associated with good CI performance, while the lesions in the auditory nerve itself (postsynaptic) are not (Eppsteiner et al., 2012).

In the last two decades, genetic defects have been proved that can cause auditory neuropathy [reviewed in Moser and Starr (2016)]. Among these genetic defects, mutations in *OTOF* gene (MIM# 603681) were the first identified and the most common cause of congenital auditory neuropathy (Rodriguez-Ballesteros et al., 2003; Varga et al., 2003; Rodriguez-Ballesteros et al., 2008; Zhang Q. J. et al., 2016). Otoferlin, encoded by the *OTOF* gene, plays an essential role in vesicle releasing and replenishing at the auditory ribbon synapses between IHCs and SGNs (Roux et al., 2006; Pangrsic et al., 2010). Mutations of *OTOF* lead to a reduction of synaptic vesicle exocytosis at ribbon synapse (Roux et al., 2006; Pangrsic et al., 2010; Michalski et al., 2017). Therefore, the site of lesion in auditory neuropathy patients with biallelic *OTOF* mutations (*OTOF*-related auditory neuropathy) is presumed to be presynaptic, and auditory nerve function is assumed to be intact. Theoretically, *OTOF* gene mutations are associated with good CI performances. Indeed, several studies did report "excellent" CI outcomes in patients with *OTOF* mutations. Nevertheless, the evidence that biallelic *OTOF* mutations are associated with good CI outcomes is not yet sufficient due to the small number ($n \leq 10$) of subjects of these studies.

In this mini review, we will focus on the CI outcomes in patients with *OTOF* mutations. The main goal is to confirm whether biallelic *OTOF* mutations are ideal indications for CI in patients with auditory neuropathy. Also, the factors that may influence the CI outcomes are discussed.

METHODS OF LITERATURE SEARCH

Following Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (Liberati et al., 2009; Moher et al., 2009), the databases PubMed, Web of Science, and the Cochrane Library were searched for relevant articles published between April 1999 and March 2020. The following search strategy was used to identify eligible studies: ('*OTOF*' OR "otoferlin") AND ('cochlear implant' OR 'cochlear implantation' OR 'CI'). All publications were searched and screened by two individuals independently. Additional articles were identified by manually searching known articles. Only full-text, peer-reviewed articles written in English were considered for inclusion. The exclusion of irrelevant studies, animal experiments, and book sections or conferences was made by screening titles and abstracts of the articles. The inclusion criteria and selection were performed through the reading of the full text. The included studies are required to report original CI outcomes in patients with *OTOF* mutations. The flow diagram is shown in **Figure 1**.

"EXCELLENT" CI OUTCOMES IN PATIENTS WITH *OTOF* MUTATIONS

CI for *OTOF*-related auditory neuropathy was first reported in 2003 by Rodriguez-Ballesteros et al. (2003). Ten subjects who met the diagnostic criteria of auditory neuropathy (TEOAEs were present, while ABRs were absent) underwent CI in the study (Rodriguez-Ballesteros et al., 2003). Despite no quantitative indicators, the results of CI in all the 10 subjects were considered to be "successful" (Rodriguez-Ballesteros et al., 2003). Subsequently, CI outcomes were assessed quantitatively in several other case reports and series (Loundon et al., 2005; Rouillon et al., 2006; Runge et al., 2013; Zhang et al., 2013; Santarelli et al., 2015; Miyagawa et al., 2016; Zhang Q. et al., 2016; Park et al., 2017; Chen et al., 2018; Hosoya et al., 2018). All of these studies showed improvement of sound perception and speech recognition with cochlear implants. Most recently, two studies of larger sample sizes ($n = 10$) reviewed the CI outcomes in patients with *OTOF*-related auditory neuropathy, and the results were compared with typical sensorineural hearing loss (SNHL) (Kim et al., 2018; Wu et al., 2018). To the best of our knowledge, CI outcomes have only been reported in approximately 60 patients with *OTOF* mutations (detailed in **Table 1**, some cases may be shared among studies). Remarkably, almost all of the patients with *OTOF* mutations developed great skills in sound perception and/or speech recognition after CI.

Sound Perception After CI

Perceiving sound is the initial step and a prerequisite for hearing rehabilitation with CI. The ability to perceive sound after CI was evaluated by audiometry. Loundon et al. (2005) and Rouillon et al. (2006) found that mean pure tone thresholds of 250, 500, 1000 and 2000 Hz in the patients with *OTOF*-related auditory neuropathy improved from 75 dB with HA to 37 and 45 dB with the cochlear implants after 1–1.5 years of rehabilitation. Similarly, Zhang et al. (2013), Zhang Q. et al. (2016), and Chen et al. (2018) found that the mean pure tone thresholds of 500, 1000, 2000 and 4000 Hz received 25 to 37.5 dB with cochlear implants after more than 2-year of rehabilitation. Although it is still unknown whether the audiometric thresholds would continue to improve with the extension of the rehabilitation time, the available data have shown that the sound perception in patients with *OTOF* mutations can be significantly improved by CI.

Speech Recognition After CI

Being able to understand speech is one of the main purposes of CI rehabilitation. Thus, speech recognition is a direct indicator of CI outcomes evaluation. Objective indicators, such as speech perception testing and speech recognition thresholds, showed that all the patients with *OTOF* mutations benefited from CI. In terms of speech perception, most of the patients with *OTOF* mutations got a $\geq$90% score in closed-set or open-set perception (detailed in **Table 1**; Loundon et al., 2005; Rouillon et al., 2006; Santarelli et al., 2015; Zhang Q. et al., 2016; Chen et al., 2018). More recently, Wu et al. (2018) reviewed 10 cases of *OTOF*-related auditory neuropathy and found that the speech

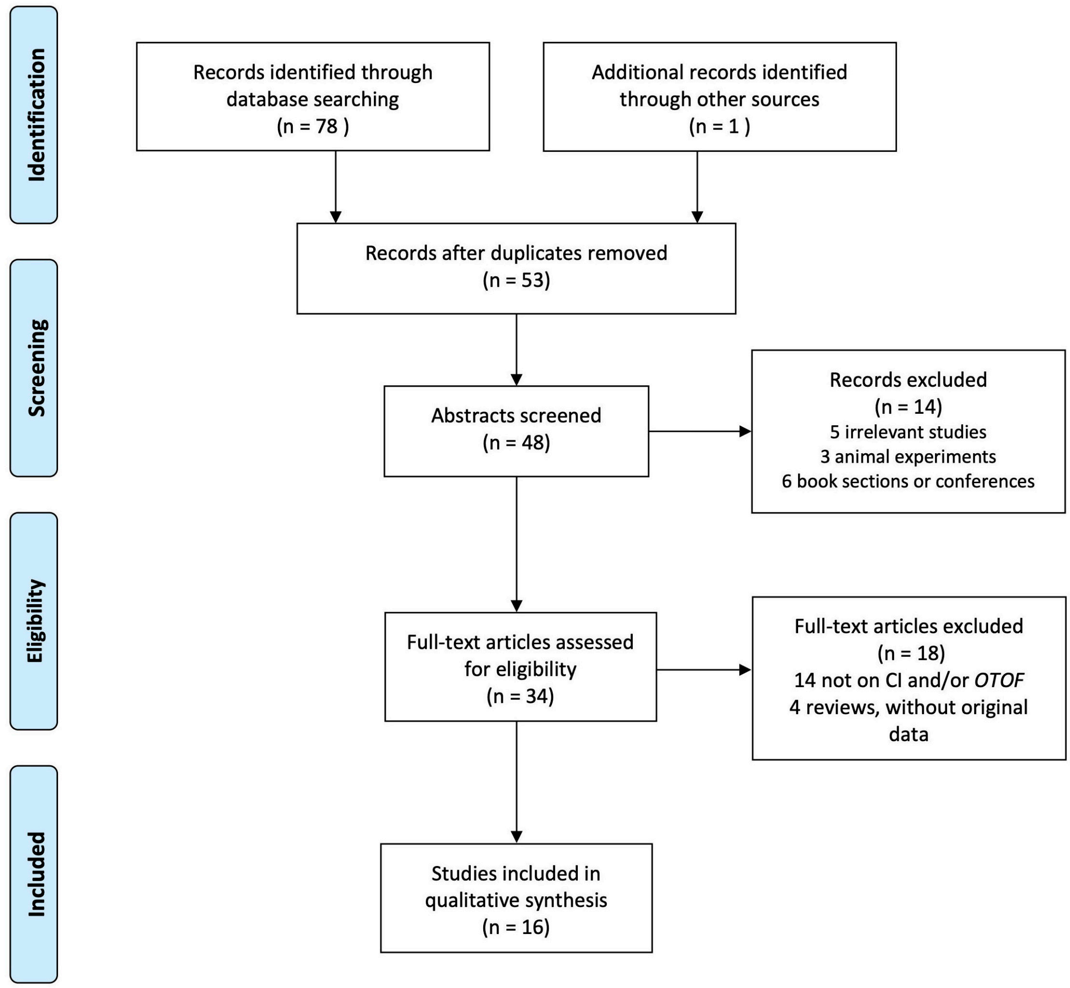

FIGURE 1 | PRISMA flow diagram.

discrimination score received 77.5 ± 37.1% at 3 years. In speech recognition thresholds, Runge et al. (2013) reported 2 cases (siblings) of *OTOF*-related auditory neuropathy and found that the thresholds were 44 and 65 dB in quiet, 52 and 70 dB in noise, respectively.

In addition, scales, such as meaningful auditory integration scale/infant-toddler meaningful auditory integration scale (MAIS/IT-MAIS), categories of auditory performance (CAP) and speech intelligibility rating (SIR) are widely used to evaluate the CI performances. In the case reports by Loundon et al. (2005) and Rouillon et al. (2006), the MAIS/IT-MAIS scores increased from 4/40 and 4/40 with HA to 40/40 and 31/40 with cochlear implants. In terms of CAP and SIR, patients with *OTOF* mutations showed rapid improvement of scores after CI (Wu et al., 2018). The CAP scores reached 6/6–7/7 during the 2–3 year follow up (Wu et al., 2011; Park et al., 2017; Kim et al., 2018). Moreover, studies by Hosoya et al. (2018) and Wu et al. (2018) showed that there was no significant difference in CAP or SIR scores among patients with *OTOF*, *GJB2, SLC26A4* mutations or cytomegalovirus infections.

According to the literature, *OTOF*-related auditory neuropathy is associated with excellent CI outcomes. Patients with this type of auditory neuropathy can not only "hear" the sound, but also "understand" the speech well with the help of cochlear implants. Unlike other types of auditory neuropathy, the CI performances of the patients with biallelic *OTOF* mutations are predictable and comparable to those of "typical" SNHL. A detection of *OTOF* mutations can help in accurately localizing the site of lesion and informing therapy-related clinical decision making in patients with auditory neuropathy.

FACTORS THAT MAY INFLUENCE CI OUTCOMES IN *OTOF*-RELATED AUDITORY NEUROPATHY

Although all studies were in coherence with that the auditory neuropathy caused by *OTOF* mutations tend to have good CI outcomes, individual variations still exist among cases. For example, Runge et al. (2013) reported a sibling pair diagnosed

TABLE 1 | Detailed cochlear implantation performances in patients with *OTOF* mutations.

Author	Year	No. of patients	Age at HL diagnosed	Age at first CI mean(rang)	Follow up duration	CI performances
Rodriguez-Ballesteros et al., 2003	2003	10	–	–	–	The results of CI were successful in terms of sound detection and communication skills.
Loundon et al., 2005	2005	1	10 m	∼4 y	12 m	**PTA (mean of 250–2000 Hz):** 45dB (vs. 75dB with HA) **Speech perception:** 100% (vs. 0% with HA); **IT-MAIS:** 31/40 (vs. 4/40 with HA); **NRT:** good responses on the tested electrodes.
Rouillon et al., 2006	2006	2*	10 m 22 m	4 y 25 m	18 m 36 m	**PTA (mean of 250–2000 Hz):** 37dB and 45dB (vs. 75dB and 75dB with HA); **Closed-set sentences:** 100%; **Open-set words and sentences:** 45–100%; **MAIS:** 40/40 and 31/40 (vs. 4/40 and 4/40 with HA); **Nottingham scale:** grade 4 and 2; **NRT:** all the electrodes with positive responses.
Chiu et al., 2010	2010	3	6 m 6 m 1 y	–	>1 y	A preliminary evaluation of the speech perception performance revealed excellent outcomes in all three patients, comparable to cochlear implantees with *OTOF* mutations Rodriguez-Ballesteros et al., 2003.
Zadro et al., 2010	2010	3	13 m 2 y 30 m	46 m 32 m 36 m	24 m 24 m 12 m	The patients showed awareness to speech sounds, and hearing perceptive abilities achieved the identification level.
Wu et al., 2011	2011	1	–	6 y	3 y	**CAP:** 7 at 3 years.
Runge et al., 2013	2013	2 (siblings)	0.5 9 m	18 m 16 m	3 y	**Speech recognition thresholds:** 44dB in quiet, 52dB in noise, and 65dB in quiet, 70dB in noise; **Lexical Neighborhood Test Easy:** 80% and 44% correct; **ECAP recovery:** one patient had a higher recovery exponent than the average of the pediatric and adult subjects; one had a recovery exponent within the average range.
Zhang et al., 2013	2013	1	12 m	20 m	∼3 y	**PTA (mean of 500–4000 Hz):** 25dB; **NRI:** waveform testing was within normal limits.
Santarelli et al., 2015	2015	6	4 m–2 y	2.1 (1–4)y	1–1.5 y	**Open-set disyllable recognition test:** 90–100%; **ECAP:** increasing stimulation levels resulted in a higher amplitude and a slight decrease in the latency.
Miyagawa et al., 2016	2016	1	–	–	12 m	**LittlEARS auditory questionnaire:** 0, 9, 24, and 30 at 0, 3, 6, and 12 months after CI
Zhang Q. et al., 2016	2016	1	30 m	4 y	2 y	**PTA (mean of 500–4000 Hz):** 37.5 dB; **Open-set words recognition:** 93%; **Open-set sentences recognition:** 98%.
Park et al., 2017	2017	5	–	–	36 m	**CAP:** 6–7 at 24 m in early implantees (age < 24 m), and 3–4 at 24 m in late implantees (age > 24 m).
Chen et al., 2018	2018	1	18 m	4.5 y	24 m	**PTA (mean of 500–4000 Hz):** 31.25dB at 18 m, 30dB at 24 m; **Open-set words:** 90% at 18 m, 95% at 24 m.
Hosoya et al., 2018	2018	4	–	27.8 (21–40) m	–	**CAP:** no significant difference among the patients with *OTOF, GJB2, SLC26A4* mutations and CMV infection; **EABR:** longer wave V, wave III, and Wave III–Wave V latencies.
Kim et al., 2018	2018	10†	–	19.2 (13–26) m	36 m	**CAP:** 4–5 at 12 m, 4–7 at 25 m, 7 at 36 m. More rapid improvement in early implantees (age ≤ 18 m) than late implantees (age > 18 m).
Wu et al., 2018	2018	10	–	2.9 (1–5.6)y	3 m–5 y	**Speech discrimination score:** 77.5 ± 37.1% at 3 years; **CAP and SIR:** no significant difference among the patients with *OTOF, GJB2* and *SLC26A4* mutations. **NRT and NRI:** all 10 patients revealed robust ECAPs.

**One case may be shared with Loundon et al. (2005); †some cases may be shared with Park et al. (2017); m: month(s); y: year(s); PTA: pure tone audiometry; CAP: categories of auditory performance; CMV: cytomegalovirus; CI: cochlear implant(s); EABR: electrically evoked auditory brainstem responses; ECAP: electrically evoked compound action potential; HA: hearing aid(s); HL: hearing loss; IT-MAIS: infant-toddler meaningful auditory integration scale; MAIS: meaningful auditory integration scale; NRI: neural response imaging; NRT: neural response telemetry; SIR: speech intelligibility rating.*

with *OTOF*-related auditory neuropathy. The genotypes of these siblings were the same, but the speech perception performance differed between the siblings. In another study, Park et al. (2017) followed up four subjects with *OTOF*-related auditory neuropathy who underwent CI and found that the CAP scales ranged from 3 to 7 at 24 months post-CI. Due to the limited number of cases, it is impossible to ascertain what is the exact factors that may influence the CI outcomes in *OTOF*-related auditory neuropathy. However, clues may be provided by these cases.

Age at Implantation

Earlier implantation is associated with better CI outcomes in patients with *OTOF* mutations. It has been widely accepted that early implantation is good for CI outcomes in patients with typical SNHL (Niparko et al., 2010; Black et al., 2011; Panda et al., 2019). 0 to 3.5 years of age is considered a critical period for first language acquisition, and implantation after that period tends to have poorer outcomes (Sharma et al., 2002; Kral and Sharma, 2012). For patients with auditory neuropathy, the critical period seems to be narrower than those with typical severe to profound SNHL. That is, patients with auditory neuropathy who undergo cochlear implantation before the age of 2 years may have better auditory outcomes than those after the age of 2 years (Cardon and Sharma, 2013; Liu et al., 2014). This could be explained by that the disordered pattern of neural input to the cortex, as a result of auditory nerve dys-synchrony in patients with auditory neuropathy have negative effects on central auditory maturation (Cardon and Sharma, 2013; Sharma and Cardon, 2015).

As mentioned before, the site of lesion in patients with *OTOF* mutations is assumed to be presynaptic (Roux et al., 2006; Pangrsic et al., 2010). Thus, the pathological mechanism of *OTOF*-related auditory neuropathy is thought to be more like a typical SNHL but not an auditory neuropathy. However, Park et al. (2017) found that early ($<$24 months) implantees experienced notably better outcomes than late ($>$24 months) implantees. Similarly, Kim et al. (2018) found that the early ($\leq$18 months) implantees had better outcomes than the late ($>$18 months) implantees at 6 months after CI. Furthermore, listening skills improved more rapidly in early implantees than late implantees (Kim et al., 2018). All the above suggest that patients with auditory neuropathy caused by mutations of *OTOF* seemed to be more affected by delayed implantation than those with typical SNHL that caused by mutations of *SLC26A4* and *GJB2* (Park et al., 2017).

The above results still require further confirmation as the sample size of existing comparative studies were small ($n \leq 10$). Besides, late implantation (6 years of age) reported by Wu et al. (2011) also showed with a "good" CI outcome, which implied that late implantation was still beneficial for some of the patients. However, from a clinicians' point of view, most of the patients with *OTOF* mutations did not experience spontaneous recovery of auditory performance, it was more likely that early implantation ($<$24 years) could achieve optimal CI performances (Wu et al., 2018).

The Integrity of the Auditory Nerve Function

The integrity of auditory nerve function is a key determinant of CI performances, especially for patients with auditory neuropathy. The auditory nerve function in *OTOF*-related auditory neuropathy is usually presumed intact. This assertion was supported by testing the neural responses of SGNs. Neural responses eluted by CI are objective indicators for evaluating the ability of the auditory pathway to receive, transmit and process complex electrical signals. By testing electrically evoked compound action potentials (ECAPs) or electrically evoked auditory brainstem responses (EABRs), the auditory nerve has been proved to have "good response" to cochlear implant stimuli (Loundon et al., 2005; Rouillon et al., 2006; Runge et al., 2013; Zhang et al., 2013; Santarelli et al., 2015; Wu et al., 2018).

However, some other studies might challenge the above view. Runge et al. (2013) quantitatively analyzed ECAP recovery rates in the sibling pair with *OTOF*-related auditory neuropathy. Though with the same genotype, one sibling had a recovery exponent within the average range of SNHL, while the other one had a more than one standard deviation (SD) higher recovery exponent than the average range of SNHL. Hosoya et al. (2018) tested the EABRs in patients with congenital hearing loss and found that the Wave III, Wave V, and Wave III–Wave V latencies were significantly longer in patients with *OTOF* mutations than those in SNHL. These two studies implicated that *OTOF* mutations might affect the more central auditory pathway beyond the synapse between the IHCs and SGNs. In addition, neurological and/or central pathologies should not necessarily be ruled out even when a patient was diagnosed with *OTOF*-related auditory neuropathy.

Genotypes of the OTOF Gene

As mutations of the *OTOF* gene are the cause of the disease, it is reasonable to speculate whether CI outcomes are associated with distinct genotypes. To date, more than 130 variants in the *OTOF* gene have been implicated pathogenic or likely pathogenic[1]. Although patients who underwent CI showed a high frequency of p.Gln829* in European (Rodriguez-Ballesteros et al., 2003; Loundon et al., 2005; Rouillon et al., 2006), p.Glu1700Gln in Chinese (Chiu et al., 2010; Wu et al., 2011; Wu et al., 2018) and p.Arg1939Glu in Korean (Park et al., 2017; Kim et al., 2018) population, more than 30 different *OTOF* mutations have been detected in the CI recipients (listed in **Supplementary Table 1**). Due to the different methods of CI performance evaluation and the variety of genotypes, it is impracticable to compare CI outcomes among patients with different genotypes directly.

Nevertheless, it is still viable to investigate whether there is a correlation between mutation types and CI outcomes. Otoferlin has six C2 domains, one Fer-like structure and one transmembrane domain (TMD). Nonsense mutations (like p.Gln829*) usually lead to the loss of C2 domain(s) and the TMD, and result in a complete loss of otoferlin function (Migliosi et al., 2002), while missense mutations (like p.Glu1700Gln and

[1] http://deafnessvariationdatabase.org/gene_page/OTOF

p.Arg1939Glu) might affect only one C2 domain, the TMD, or neither of them (Varga et al., 2003; Chiu et al., 2010; Zhang Q. J. et al., 2016). Based on the available data, all patients with homozygous p.Gln829*, p.Glu1700Gln or p.Arg1939Glu revealed excellent outcomes in sound perception and/or speech recognition (Rodriguez-Ballesteros et al., 2003; Chiu et al., 2010; Wu et al., 2011; Kim et al., 2018). Therefore, there is insufficient evidence that CI outcomes are correlated with distinct *OTOF* genotypes.

Recommendations Regarding the CI in *OTOF*-Related Auditory Neuropathy

As most of the patients with *OTOF*-related auditory neuropathy presented a phenotype of stable, severe to profound non-syndromic hearing loss (Zhang Q. J. et al., 2016), it seems that CI is the optimal and the only defective treatment option. However, CI may not be suitable for all patients with *OTOF* mutations. Firstly, the confirmation of *OTOF*-related auditory neuropathy could be challenging. There is no well-accepted hotspot mutation in *OTOF* except p.Gln829* in Spanish (Migliosi et al., 2002), and the number of novel *OTOF* mutations is growing, but most of the mutations lack functional studies. In a patient with *OTOF* mutations, it is difficult to confirm that the hearing loss is caused by otoferlin (*OTOF*) deficiency and rule out other causes. Secondly, some of the patients with *OTOF* mutations manifested as temperature sensitive auditory neuropathy (TS-AN), i.e. the hearing thresholds fluctuate with a variation of core body temperature and may improve with age (Zhang Q. et al., 2016). In these cases, HA could be an effective treatment. Thus, HA trials are still recommended in the patients with mild to moderate fluctuating hearing loss, but once a patient was identified with severe to profound hearing loss or fail to develop age-appropriate language skills, CI would be considered. Due to a possible narrower critical period for CI, earlier (<24 months) implantation is recommended (Cardon and Sharma, 2013). Before the operation, clinical manifestations, molecular test results, and the auditory nerve functions should be comprehensively assessed to exclude TS-AN and auditory neuropathy caused by other reasons.

CONCLUSION

The existing literature consistently revealed that patients with *OTOF* mutations are associated with excellent CI performance in both sound perception and speech recognition. Genetic analysis of *OTOF* can provide great help in localizing the site of lesion and informing therapy-related clinical decision making. Auditory neuropathy with biallelic *OTOF* mutations is an ideal surgical indication for CI. Notably, compared with typical SNHL, a narrower critical period for CI was implied in patients with *OTOF* mutations. Thus, once diagnosed as an *OTOF*-related auditory neuropathy with severe to profound hearing loss, early implantation is recommended. In addition, although the auditory nerve function is normal in most of the patients with *OTOF* mutations, neurological and/or central pathologies should not be ruled out in these cases. There is no evidence that CI outcomes are correlated with distinct *OTOF* genotypes. Compared with typical SNHL, the sample size of the studies on CI outcomes in patients with *OTOF* mutations is small. Future studies with larger sample sizes are required to confirm this conclusion.

AUTHOR CONTRIBUTIONS

XL conceived the review. DZ participated in drafting the manuscript. Both authors wrote the manuscript.

FUNDING

This work was supported by the Ministry of Science and Technology of China under grant no. 2014CB541700; and Department of Science and Technology of Zhejiang Province under grant no. 2015C03035.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2020.00447/full#supplementary-material

REFERENCES

Berlin, C. I., Hood, L. J., Morlet, T., Wilensky, D., Li, L., Mattingly, K. R., et al. (2010). Multi-site diagnosis and management of 260 patients with auditory neuropathy/dys-synchrony (auditory neuropathy spectrum disorder). *Int. J. Audiol.* 49, 30–43. doi: 10.3109/14992020903160892

Black, J., Hickson, L., Black, B., and Perry, C. (2011). Prognostic indicators in paediatric cochlear implant surgery: a systematic literature review. *Cochlear. Implants Int.* 12, 67–93. doi: 10.1179/146701010X486417

Cardon, G., and Sharma, A. (2013). Central auditory maturation and behavioral outcome in children with auditory neuropathy spectrum disorder who use cochlear implants. *Int. J. Audiol.* 52, 577–586. doi: 10.3109/14992027.2013.799786

Chen, K., Liu, M., Wu, X., Zong, L., and Jiang, H. (2018). Targeted next generation sequencing reveals OTOF mutations in auditory neuropathy spectrum disorder. *Int. J. Pediatr. Otorhinolaryngol.* 115, 19–23. doi: 10.1016/j.ijporl.2018.09.008

Chiu, Y. H., Wu, C. C., Lu, Y. C., Chen, P. J., Lee, W. Y., Liu, A. Y., et al. (2010). Mutations in the OTOF gene in Taiwanese patients with auditory neuropathy. *Audiol. Neurootol.* 15, 364–374. doi: 10.1159/000293992

Eppsteiner, R. W., Shearer, A. E., Hildebrand, M. S., DeLuca, A. P., Ji, H. H., Dunn, C. C., et al. (2012). Prediction of cochlear implant performance by genetic mutation: the spiral ganglion hypothesis. *Hear. Res.* 292, 51–58. doi: 10.1016/j.heares.2012.08.007

Hayes, D., and Sininger, Y. S. (2008). "Guidelines: identification and management of infants and young children with auditory neuropathy spectrum disorder," in *Proceedings of the Guidelines Development Conference*, Como. doi: 10.3109/14992027.2012.684404

Hosoya, M., Minami, S. B., Enomoto, C., Matsunaga, T., and Kaga, K. (2018). Elongated EABR wave latencies observed in patients with auditory neuropathy caused by OTOF mutation. *Laryngosc. Investig. Otolaryngol.* 3, 388–393. doi: 10.1002/lio2.210

Humphriss, R., Hall, A., Maddocks, J., Macleod, J., Sawaya, K., and Midgley, E. (2013). Does cochlear implantation improve speech recognition in children with auditory neuropathy spectrum disorder? A systematic review. *Int. J. Audiol.* 52, 442–454. doi: 10.3109/14992027.2013.786190

Kim, B. J., Jang, J. H., Han, J. H., Park, H.-R., Oh, D. Y., Lee, S., et al. (2018). Mutational and phenotypic spectrum of OTOF-related auditory neuropathy in Koreans: eliciting reciprocal interaction between bench and clinics. *J. Transl. Med.* 16:330. doi: 10.1186/s12967-018-1708-z

Kral, A., and Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends Neurosci.* 35, 111–122. doi: 10.1016/j.tins.2011.09.004

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J. P., et al. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med.* 6:e1000100. doi: 10.1371/journal.pmed.1000100

Liu, Y., Dong, R., Li, Y., Xu, T., Li, Y., Chen, X., et al. (2014). Effect of age at cochlear implantation on auditory and speech development of children with auditory neuropathy spectrum disorder. *Auris Nasus Larynx* 41, 502–506. doi: 10.1016/j.anl.2014.06.001

Loundon, N., Marcolla, A., Roux, I., Rouillon, I., Denoyelle, F., Feldmann, D., et al. (2005). Auditory neuropathy or endocochlear hearing loss? *Otol. Neurotol.* 26, 748–754. doi: 10.1097/01.mao.0000169044.63970.4a

Michalski, N., Goutman, J. D., Auclair, S. M., Boutet de Monvel, J., Tertrais, M., Emptoz, A., et al. (2017). Otoferlin acts as a Ca(2+) sensor for vesicle fusion and vesicle pool replenishment at auditory hair cell ribbon synapses. *eLife* 6:e31013. doi: 10.7554/eLife.31013

Migliosi, V., Modamio-Hoybjor, S., Moreno-Pelayo, M. A., Rodriguez-Ballesteros, M., Villamar, M., Telleria, D., et al. (2002). Q829X, a novel mutation in the gene encoding otoferlin (OTOF), is frequently found in Spanish patients with prelingual non-syndromic hearing loss. *J. Med. Genet.* 39, 502–506. doi: 10.1136/jmg.39.7.502

Miyagawa, M., Nishio, S. Y., and Usami, S. (2016). A comprehensive study on the etiology of patients receiving cochlear implantation with special emphasis on genetic epidemiology. *Otol. Neurotol.* 37, e126–e134. doi: 10.1097/MAO.0000000000000936

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 6:e1000097. doi: 10.1371/journal.pmed.1000097

Moser, T., and Starr, A. (2016). Auditory neuropathy–neural and synaptic mechanisms. *Nat. Rev. Neurol.* 12, 135–149. doi: 10.1038/nrneurol.2016.10

Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N. Y., Quittner, A. L., et al. (2010). Spoken language development in children following cochlear implantation. *JAMA* 303, 1498–1506. doi: 10.1001/jama.2010.451

Panda, S., Sikka, K., Singh, V., Agarwal, S., Kumar, R., Thakar, A., et al. (2019). Comprehensive analysis of factors leading to poor performance in prelingual cochlear implant recipients. *Otol. Neurotol.* 40, 754–760. doi: 10.1097/mao.0000000000002237

Pangrsic, T., Lasarow, L., Reuter, K., Takago, H., Schwander, M., Riedel, D., et al. (2010). Hearing requires otoferlin-dependent efficient replenishment of synaptic vesicles in hair cells. *Nat. Neurosci.* 13, 869–876. doi: 10.1038/nn.2578

Park, J. H., Kim, A. R., Han, J. H., Kim, S. D., Kim, S. H., Koo, J.-W., et al. (2017). Outcome of cochlear implantation in prelingually deafened children according to molecular genetic etiology. *Ear Hear.* 38, E316–E324. doi: 10.1097/AUD.0000000000000437

Rance, G., Beer, D. E., Cone-Wesson, B., Shepherd, R. K., Dowell, R. C., King, A. M., et al. (1999). Clinical findings for a group of infants and young children with auditory neuropathy. *Ear Hear.* 20, 238–252. doi: 10.1097/00003446-199906000-00006

Rodriguez-Ballesteros, M., del Castillo, F. J., Martin, Y., Moreno-Pelayo, M. A., Morera, C., Prieto, F., et al. (2003). Auditory neuropathy in patients carrying mutations in the otoferlin gene (OTOF). *Hum. Mutat.* 22, 451–456. doi: 10.1002/humu.10274

Rodriguez-Ballesteros, M., Reynoso, R., Olarte, M., Villamar, M., Morera, C., Santarelli, R., et al. (2008). A multicenter study on the prevalence and spectrum of mutations in the otoferlin gene (OTOF) in subjects with nonsyndromic hearing impairment and auditory neuropathy. *Hum. Mutat.* 29, 823–831. doi: 10.1002/humu.20708

Rouillon, I., Marcolla, A., Roux, I., Marlin, S., Feldmann, D., Couderc, R., et al. (2006). Results of cochlear implantation in two children with mutations in the OTOF gene. *Int. J. Pediatr. Otorhinolaryngol.* 70, 689–696. doi: 10.1016/j.ijporl.2005.09.006

Roux, I., Safieddine, S., Nouvian, R., Grati, M., Simmler, M. C., Bahloul, A., et al. (2006). Otoferlin, defective in a human deafness form, is essential for exocytosis at the auditory ribbon synapse. *Cell* 127, 277–289. doi: 10.1016/j.cell.2006.08.040

Runge, C. L., Erbe, C. B., McNally, M. T., Van Dusen, C., Friedland, D. R., Kwitek, A. E., et al. (2013). A novel otoferlin splice-site mutation in siblings with auditory neuropathy spectrum disorder. *Audiol. Neurootol.* 18, 374–382. doi: 10.1159/000354978

Santarelli, R., del Castillo, I., Cama, E., Scimemi, P., and Starr, A. (2015). Audibility, speech perception and processing of temporal cues in ribbon synaptic disorders due to OTOF mutations. *Hear. Res.* 330, 200–212. doi: 10.1016/j.heares.2015.07.007

Sharma, A., and Cardon, G. (2015). Cortical development and neuroplasticity in auditory neuropathy spectrum disorder. *Hear. Res.* 330(Pt B), 221–232. doi: 10.1016/j.heares.2015.06.001

Sharma, A., Dorman, M. F., and Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear Hear.* 23, 532–539. doi: 10.1097/00003446-200212000-00004

Starr, A., Picton, T. W., Sininger, Y., Hood, L. J., and Berlin, C. I. (1996). Auditory neuropathy. *Brain* 119, 741–753. doi: 10.1093/brain/119.3.741

Varga, R., Kelley, P. M., Keats, B. J., Starr, A., Leal, S. M., Cohn, E., et al. (2003). Non-syndromic recessive auditory neuropathy is the result of mutations in the otoferlin (OTOF) gene. *J. Med. Genet.* 40, 45–50. doi: 10.1136/jmg.40.1.45

Wu, C. C., Hsu, C. J., Huang, F. L., Lin, Y. H., Lin, Y. H., Liu, T. C., et al. (2018). Timing of cochlear implantation in auditory neuropathy patients with OTOF mutations: our experience with 10 patients. *Clin. Otolaryngol.* 43, 352–357. doi: 10.1111/coa.12949

Wu, C. C., Liu, T. C., Wang, S. H., Hsu, C. J., and Wu, C. M. (2011). Genetic characteristics in children with cochlear implants and the corresponding auditory performance. *Laryngoscope* 121, 1287–1293. doi: 10.1002/lary.21751

Yawn, R. J., Nassiri, A. M., and Rivas, A. (2019). Auditory neuropathy: bridging the gap between hearing aids and cochlear implants. *Otolaryngol. Clin. North Am.* 52, 349–355. doi: 10.1016/j.otc.2018.11.016

Zadro, C., Ciorba, A., Fabris, A., Morgutti, M., Trevisi, P., Gasparini, P., et al. (2010). Five new OTOF gene mutations and auditory neuropathy. *Int. J. Pediatr. Otorhinolaryngol.* 74, 494–498. doi: 10.1016/j.ijpor1.2010.02.004

Zhang, L. P., Chai, Y. C., Yang, T., and Wu, H. (2013). Identification of novel OTOF compound heterozygous mutations by targeted next-generation sequencing in a Chinese patient with auditory neuropathy spectrum disorder. *Int. J. Pediatr. Otorhinolaryngol.* 77, 1749–1752. doi: 10.1016/j.ijporl.2013.08.007

Zhang, Q., Lan, L., Shi, W., Yu, L., Xie, L.-Y., Xiong, F., et al. (2016). Temperature sensitive auditory neuropathy. *Hear. Res.* 335, 53–63. doi: 10.1016/j.heares.2016.01.008

Zhang, Q. J., Han, B., Lan, L., Zong, L., Shi, W., Wang, H. Y., et al. (2016). High frequency of OTOF mutations in Chinese infants with congenital auditory neuropathy spectrum disorder. *Clin. Genet.* 90, 238–246. doi: 10.1111/cge.12744 doi: 10.1111/cge.12744

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 04 November 2020
doi: 10.3389/fnins.2020.592954

Check for updates

Acoustic Assessment of Tone Production of Prelingually-Deafened Mandarin-Speaking Children With Cochlear Implants

*Yitao Mao[1], Hongsheng Chen[2], Shumin Xie[2] and Li Xu[3]**

[1] Department of Radiology, Xiangya Hospital, Central South University, Changsha, China, [2] Department of Otolaryngology-Head and Neck Surgery, Xiangya Hospital, Central South University, Changsha, China, [3] Communication Sciences and Disorders, Ohio University, Athens, OH, United States

OPEN ACCESS

Edited by:
Fei Chen,
Southern University of Science and Technology, China

Reviewed by:
Duo-duo Tao,
Soochow University, China
Richard Charles Dowell,
The University of Melbourne, Australia

****Correspondence:***
Li Xu
xul@ohio.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *08 August 2020*
Accepted: *12 October 2020*
Published: *04 November 2020*

Citation:
Mao Y, Chen H, Xie S and Xu L (2020) Acoustic Assessment of Tone Production of Prelingually-Deafened Mandarin-Speaking Children With Cochlear Implants. Front. Neurosci. 14:592954. doi: 10.3389/fnins.2020.592954

Objective: The purpose of the present study was to investigate Mandarin tone production performance of prelingually deafened children with cochlear implants (CIs) using modified acoustic analyses and to evaluate the relationship between demographic factors of those CI children and their tone production ability.

Methods: Two hundred seventy-eight prelingually deafened children with CIs and 173 age-matched normal-hearing (NH) children participated in the study. Thirty-six monosyllabic Mandarin Chinese words were recorded from each subject. The fundamental frequencies (F0) were extracted from the tone tokens. Two acoustic measures (i.e., differentiability and hit rate) were computed based on the F0 onset and offset values (i.e., the tone ellipses of the two-dimensional [2D] method) or the F0 onset, midpoint, and offset values (i.e., the tone ellipsoids of the 3D method). The correlations between the acoustic measures as well as between the methods were performed. The relationship between demographic factors and acoustic measures were also explored.

Results: The children with CIs showed significantly poorer performance in tone differentiability and hit rate than the NH children. For both CI and NH groups, performance on the two acoustic measures was highly correlated with each other (r values: 0.895–0.961). The performance between the two methods (i.e., 2D and 3D methods) was also highly correlated (r values: 0.774–0.914). Age at implantation and duration of CI use showed a weak correlation with the scores of acoustic measures under both methods. These two factors jointly accounted for 15.4–18.9% of the total variance of tone production performance.

Conclusion: There were significant deficits in tone production ability in most prelingually deafened children with CIs, even after prolonged use of the devices. The strong correlation between the two methods suggested that the simpler, 2D method seemed to be efficient in acoustic assessment for lexical tones in hearing-impaired children. Age at implantation and especially the duration of CI use were significant, although weak, predictors for tone development in pediatric CI users. Although a large part of tone production ability could not be attributed to these two factors, the results still encourage early implantation and continual CI use for better lexical tone development in Mandarin-speaking pediatric CI users.

Keywords: cochlear implant, tone production, Mandarin Chinese, lexical tone, acoustic analysis, pediatric

INTRODUCTION

The modern cochlear implant (CI) is currently the most successful neural prosthesis in wide clinical application. It can restore the sense of hearing for hearing-impaired individuals by bypassing the damaged sensory cells in the inner ear and stimulating the auditory nerve directly (Wilson, 2019). Previous evidence showed that severely to profoundly deafened children obtained enormous benefits for their speech and language development after cochlear implantation (Niparko et al., 2010). Speech production ability and language development of CI users, especially of those children with prelingual deafness, have been the major focus of the postoperative rehabilitation process. Detailed acoustic analyses of the production of vowels and fricatives in CI children have shown significant progress in phoneme development, but there are still significant remaining deficits in speech production in those children (Uchanski and Geers, 2003; Yang et al., 2015, 2017a; Yang and Xu, 2017). For CI children who speak tonal languages, such as Mandarin Chinese, speech production is compounded by the involvement of lexical tones. In tonal languages, the word meaning depends not only on the phonemes (such as the combination of consonants and vowels), but also on the pattern of tones (i.e., the fundamental frequency, F0) of the syllables. In other words, for a specific syllable, changing the F0 contour will bring about a change in the meaning of the syllable. There are four tones in Mandarin, namely tones 1, 2, 3, and 4. The F0 contours are (1) high and flat, (2) low at the beginning and then rising, (3) falling at the beginning and then rising with a dip in the middle, and (4) high-falling, respectively. Such tonal information, primarily carried by the F0, is not adequately coded in current CI devices (Han et al., 2009; Xu and Zhou, 2011; Limb and Roy, 2014; Deroche et al., 2019). Previous studies that focused on tone perception have revealed significant deficits of CI children in tone recognition tasks, with tremendous variability observed across CI subjects (Lee et al., 2002; Peng et al., 2004; Han et al., 2009; Zhou et al., 2013; Mao and Xu, 2017; Holt et al., 2018). As a result of tone recognition deficits, the tone production ability of prelingually deafened children with CIs might also be compromised.

The specific mechanism of the influence of auditory feedback on oral speech is not entirely clear. It was suggested that auditory feedback has a significant and immediate effect on oral speech (Davidson, 1959; Lane and Tranel, 1971). For example, when exposed to noise, the talker's vocal intensity would increase involuntarily, a phenomenon known as the "Lombard effect" (Lane and Tranel, 1971). If the auditory feedback is deliberately delayed, it will cause the speech speed to slow down (Davidson, 1959). The frequency information is also affected by the auditory feedback frequency (Elman, 1981). For phoneme pronunciation, Houde and Jordan (1998) found that if the first three formant frequencies of vowels in auditory feedback are deliberately changed, the produced vowels would be unconsciously replaced by other vowels to compensate for these formant changes. All these findings supported the hypothesis that there is a closely coupled loop between auditory perception and vocal production (Deroche et al., 2019), and auditory feedback can regulate oral speech instantaneously (Natke and Kalveram, 2001; Amir et al., 2003; Mora et al., 2012; Liu et al., 2020).

For postlingually deafened adults, the connection between perception and production will decline gradually due to the loss of auditory feedback, whereas for prelingually deafened children, this connection has not been established. Earlier studies have shown that for postlingually deafened adults, the loss of hearing does not have a significant impact on their speech intelligibility but only gradually changes some acoustic parameters of their oral speech with a very slow rate (Waldstein, 1990; Leder and Spitzer, 1993; Plant, 1993). In the absence of auditory feedback, they seem to use their knowledge and experiences to regulate their vocal organs to make the desired sound (Matthies et al., 1994). However, for children with prelingual deafness, the connection between hearing and vocal production has not been well established in their speech acquisition stage, which harms their speech intelligibility. Because of the absence of effective auditory feedback, prelingually deafened children would likely rely on visual or somatosensory inputs to establish a feedback connection with their vocal production (Tobey et al., 1991; Osberger et al., 1993; Nava et al., 2014; Selleck and Sataloff, 2014). With their auditory function partially restored with CIs, prelingually deafened children still face challenges in their speech production (Uchanski and Geers, 2003; Yang et al., 2015, 2017a; Yang and Xu, 2017).

Pitch information is not adequately coded in the contemporary envelope-based speech processing strategies in which fixed-rate electrical stimulations delivered to a small number of CI electrodes result in poor pitch perception in CI users (Wilson and Dorman, 2008; Xu and Zhou, 2011). At present, numerous studies have reported that there are considerable deficits in Mandarin tone recognition for CI children (see Tan et al., 2016; Chen and Wong, 2017; Liu et al., 2017 for reviews). For example, Zhou et al. (2013) and Mao and Xu (2017) reported that CI children achieved Mandarin tone recognition of 67.3 to 82.3% correct, whereas their normal-hearing (NH) counterpart obtained > 95% correct. The inadequate tonal information since childhood is likely to make tone production problematic in the speech development of those CI children who use Mandarin Chinese as their mother tongue. Several previous studies with relatively small sample sizes have found that the tone production ability of Mandarin-speaking pediatric CI users was significantly poorer compared with NH children at a similar age range (Wei et al., 2000; Peng et al., 2004; Xu et al., 2004, 2011; Han et al., 2007; Zhou et al., 2013; Tang et al., 2019). Peng et al. (2004) reported tone production accuracy in 30 CI children aged between 6.0 and 12.5 years old based on the subjective judgment of NH adults. The average tone production accuracy was only 53.1% correct. Zhou et al. (2013) also reported that tone intelligibility was only 46.8% correct for their 76 CI children with an age range of 2.4–16.2 years old. A common finding by these studies was that these CI children had tremendous individual variability in tone production ability and that their tone production was distributed from the chance level to near-perfect performance.

With the increasing number of prelingually deaf children who have received cochlear implantation in China in the past decades, it is of great importance to explore their vocal tone ability using a more objective way of evaluation. In our previous studies (Xu et al., 2007; Zhou et al., 2013; Mao et al., 2017), we used an artificial neural network to evaluate the tone production ability of children with CIs. The artificial neural network yielded an objective and efficient way to assess tone production ability; however, it could not reveal what the deficits in tone production were in CI children. Acoustic analyses might be of great value in pinpointing such deficits. Barry and Blamey (2004) proposed a method of acoustic analysis to assess Cantonese tone production. Zhou and Xu (2008) modified this method and applied it in Mandarin tone production evaluation of CI children. This acoustic method was based on the F0 contours of the produced tone tokens. In particular, the onset and offset frequencies of the F0 contours were extracted, and the tonal ellipses were generated over the scatter plots of the F0 onset versus F0 offset values. The spread and degree of overlap among tonal ellipses were quantified by a series of acoustic indices to reflect various aspects of the tone production ability. In a recent study, Tang et al. (2019) examined the F0 contours of tone tokens produced by prelingually deafened children with CIs. The authors quantified tone production accuracy based on the curvature of the F0 contours. In the 72 pediatric CI users, those who received CIs between 1 and 2 years of age demonstrated near-normal tone contours, whereas all other CI children's tone patterns tended to be flattened (Tang et al., 2019).

In the present study, we recruited a large cohort of prelingually deafened children with CIs (N = 278) and age-matched NH children (N = 173). A modified acoustic analysis method was developed and used to evaluate the tone production skills of the children. The purpose is to verify the effectiveness of this modified acoustic analysis method in the evaluation of tone production of pediatric CI users and to explore the different tone-production characteristics of the hearing-impaired group from those of the NH group. Correlational analyses were implemented between several demographic factors of the CI group and the acoustic indices obtained by our modified method in the present study. In addition, a generalized linear model (GLM; Song et al., 2013) was also used to explore further the effects of demographic factors on tone production performance.

MATERIALS AND METHODS

Subjects

A total of 278 prelingually deafened, Mandarin-speaking children were recruited to participate in the present study. The inclusion criteria were as follows: (1) prelingual sensorineural hearing loss, (2) bilateral severe to profound hearing loss (≥85 dB HL) and implanted unilaterally, (3) limited or no hearing aid use experiences before CI implantation, (4) chronological age was >2 years old, (5) the age at implantation was <12 years old, (6) using Mandarin as the mother tongue or the rehabilitation language, and (7) hearing impairment was the only health problem. In this CI group, there were 152 boys and 126 girls, ranging in chronological age from 2.13 to 19.04 (mean ± SD: 6.64 ± 3.46) years old, the age at implantation was from 0.50 to 11.02 (3.38 ± 2.25) years old, and the CI use duration was from 0.14 to 11.20 (3.26 ± 2.64) years.

As the control group, 173 Mandarin-speaking NH children from kindergartens and primary schools with ages between 2.28 and 12.51 (6.83 ± 2.85) years old were recruited in the present study. The parents reported the NH status. In the NH group, there were 94 boys and 79 girls. The mean chronological ages of these two groups were not statistically different (t-test, t = 0.479, $p > 0.05$). The use of human subjects was reviewed and approved by the Institutional Review Board of Ohio University.

Test Materials

Eighteen monosyllables (i.e., bei, bi, chi, chuang, deng, hu, jian, mao, mi, qiang, san, shu, tang, tu, wa, wu, ye, and yu) in Mandarin Chinese were selected as the targets. Each monosyllable was assigned two tones to make up a tone contrast. Therefore, the test materials consisted of 36 Chinese words (a complete list of the 36 words can be found in Han et al., 2009). All the 36 words were at the vocabulary level of young children and were used in previous studies (Han et al., 2009; Zhou et al., 2013). Each of the tone contrasts (i.e., tone 1 vs. 2, tone 1 vs. 3, tone 1 vs. 4, tone 2 vs. 3, tone 2 vs. 4, and tone 3 vs. 4) had three pairs of monosyllabic words, and each tone type (i.e., tone 1, tone 2, tone 3, and tone 4) had nine monosyllabic words, thus making it balanced among the number of monosyllabic words for tone contrasts or tone types.

Test Procedure

The test was conducted in a sound-treated room. The 36 test words were presented to the subjects in the form of cards used to elicit vocal production. Each card displayed a picture illustrating the meaning of the target word, the Chinese character, and the corresponding Pinyin (i.e., an alphabetic form indicating the pronunciation of the Chinese character). The experimenter first explained the test requirements to the subjects to make sure they understood the tasks. A recorder microphone was then placed in front of the subjects with a distance of approximately 10 cm from the subject's lips. With the help of test cards, the experimenter guided the subjects to speak out the target words, which were recorded at a sampling rate of 44.1 kHz and an amplitude resolution of 16 bits.

Acoustic Analysis

An autocorrelation algorithm was used to extract the F0s of each produced tone token (Xu et al., 2006, 2007; Zhou et al., 2008). The F0 contours were then drawn on a narrowband spectrogram for accuracy comparison. Occasionally, there were some errors in the extracted F0 contours, which, for a large part, were doubling and halving errors. Those errors were corrected manually on the spectrogram.

To eliminate the impact of individual vocal pitch on the differentiability of the four tones when the data were pooled together across all subjects, the F0 data were normalized subject by subject. The normalization algorithm was as follows: (1) we took the mean F0 of all tokens in tone 1 of one subject and called it *M,* (2) all F0 data of this subject was converted to semitones based on the equation below, and (3) the normalization was then applied for all subjects in both groups.

$$\text{Semitone} = 12 \times \log_2\left(\frac{F0}{M}\right)$$

In Zhou and Xu (2008), the F0 onset and offset of the F0 contours were extracted, and four tonal ellipses based on the four scatter plots of F0 onset versus F0 offset data of the four tones were defined. The center of the ellipses was the center of the scatter distribution, and the major and minor axes of the ellipses were of two standard deviations (SDs) of the distribution in length. Three acoustic indices were calculated based on the tonal ellipses. Index 1 was defined as the ratio of the area of quadrangle formed by joining the centers of the four tonal ellipses relative to the averaged area of the four ellipses. Index 2 was defined as the ratio of the averaged distance of the centers of the four tonal ellipses from each other relative to the averaged lengths of the two axes for four Mandarin tonal ellipses. Index 3 was the averaged proportion of the number of points of a specific tone inside that specific tonal ellipse. The three indices were found highly correlated with each other, with all correlation coefficients >0.94. In the present study, we modified these indices into two features: tone differentiability and tone hit rate. In addition, besides the two endpoints on the F0 contours, we also incorporated the middle point of the F0 contours in our computation to capture potentially distinctive characteristics of tone contour, which could be especially meaningful for tone 3 (Tupper et al., 2020). This latter method was referred to as the three-dimensional (3D) method to differentiate it from the 2D method that used the F0 onset and offset values only in the present study.

Tone Differentiability

We modified the algorithm of the Index 1 and Index 2 from Zhou and Xu (2008) for tone differentiability and decomposed it into the differentiable degree of each tone contrast (i.e., tone 1 vs. 2, tone 1 vs. 3, tone 1 vs. 4, tone 2 vs. 3, tone 2 vs. 4, and tone 3 vs. 4). Taking different tones 1 vs. 2 as an example, our algorithm was as follows: assuming A_i represented the intersected area of tonal ellipse 1 and tonal ellipse 2, A_1 and A_2 represented the area of tonal ellipse 1 and tonal ellipse 2, respectively. Then, the differentiability between tone 1 and tone 2 was calculated using the following equation:

$$\text{Differentiability (tone 1 vs. 2)} = \frac{\left(1 - \frac{A_i}{A_1}\right) + \left(1 - \frac{A_i}{A_2}\right)}{2}$$

The differentiability in the 3D method was calculated similarly except that the area of ellipses was changed to the volume of ellipsoids. The center of a tone ellipsoid was placed at the means of the distributions of F0 onset, middle, and offset values of a particular tone and the principal semiaxes were equal to two SDs of the distributions. Tone differentiability became percentage data so that it was more intuitive than the previous index values.

Tone Hit Rate

The algorithm of tone hit rate was similar to that of Index 3 in the Zhou and Xu (2008) study. For example, the hit rate of tone 1 was defined as the number of points with tone 1 as the target inside tonal ellipse 1 (or ellipsoid 1) divided by the number of all points (i.e., all tones) inside tonal ellipse 1 (or ellipsoid 1), which was technically the proportion of the points of tone 1 inside tonal ellipse 1 (or ellipsoid 1). In the present study, the proportions of the points of tones 2, 3, and 4 inside tonal ellipse 1(or ellipsoid 1) were also separately calculated to display the hit-rate data in the form of a confusion matrix. The compilation of a hit-rate confusion matrix, which was not conducted in the previous study (Zhou and Xu, 2008), might further provide insight into tone production deficits in prelingually deafened children with CIs.

Statistical Analyses

The calculated indices in the present study for tone differentiability and hit rate were percentage data and were arcsine transformed before statistical analyses, as the percentage data were not recommended to analyze directly due to the heterogeneous variance. Arcsine transformation was a way to homogenize the variance, making the data more suitable for ANOVA or other statistical analyses (Studebaker, 1985). A two-way ANOVA was conducted to explore the effects of hearing status (i.e., NH or CI) and methods used (i.e., 2D method or 3D method) on the averaged tone differentiability, as well as the averaged tone hit rate. The possible interactions between the main factors were also examined in each two-way ANOVA. A one-way repeated-measures ANOVA was adopted to assess the possible effects of tone types (i.e., tone 1, tone 2, tone 3, or tone 4) on tone hit rate, as well as the effect of tone contrasts (i.e., tone 1 vs. 2, tone 1 vs. 3, tone 1 vs. 4, tone 2 vs. 3, tone 2 vs. 4, and tone 3 vs. 4) on the tone differentiability. In addition, Pearson correlational analyses were conducted for the potential relationship between averaged tone differentiability and hit rate and between the 2D and 3D methods. Pearson correlational analyses were also implemented to examine whether these acoustic indices (tone differentiability and hit rate) were correlated with any of the demographic factors, including age at implantation, chronological age, and CI use duration. In addition, GLM analyses were implemented to examine further the combined contributions of these demographic factors. As chronological age was actually a linear combination of the other two factors (i.e., the sum of age at implantation and duration of

CI use), this factor was thus excluded in the GLM. Therefore, the GLM analyses explored the effects of age at implantation, CI use duration, and the interaction of these two main factors on tone production performance.

RESULTS

Tone Production Performance in Normal-Hearing and Cochlear Implant Groups

Figure 1 illustrated the tonal ellipses of the two groups based on the 2D method (upper panels) and four representative subjects from either group (lower panels). The representative subjects were randomly selected, one from each quartile of the differentiability score, in respective groups. The boundaries of the four tonal ellipses in the NH group were relatively clearly separated. For tone 1, both F0 onset and offset were relatively high. Thus, the points were mainly located in the upper left quadrant of the scatter plot. For tone 2, the F0 onset was low, and the offset was high, and the data points were mainly located in the upper right quadrant. For tone 3, the heights of both F0 onset and offset were the lowest, making the data points located in the lower-left quadrant. For tone 4, the F0 onset was high, and the offset was low; thus, the corresponding points were located in the lower-right quadrant. Hence, the four ellipses of the NH group were differentiable from each other. However, for the CI group, the scattered F0 data points of the four tones were overlapped with each other to a greater extent. At the individual levels, it was difficult to distinguish some of the tone categories from each other except for the very best performers in the CI group (**Figure 1**).

Figure 2 shows the tonal ellipsoids based on the 3D method of the two groups and four representative subjects from either group (lower panels). The representative subjects were randomly selected, one from each quartile of the differentiability score, in respective groups. Like the 2D method, the boundaries of the four tonal ellipsoids in the NH group were relatively clearly separated. The four ellipsoids representing the four tones had their own unique positions in a 3D space and were spatially differentiable from each other. However, the four ellipsoids in the CI group were overlapped with each other to a great degree and were not separable spatially as a whole. At the individual levels, some of the better performers in the CI group showed well-differentiated tonal ellipsoids, and their differentiability scores surpassed those of the poorer performers in the NH group.

Differentiability of Tone Production in Normal-Hearing and Cochlear Implant Groups

The tone differentiability score was computed for each subject to quantify the differentiability of tone contrast in the production. The upper panel of **Figure 3** shows the differentiability of each tone contrast based on the 2D method. For the NH group, the differentiability between tone 2 and tone 3 (i.e., contrast tone 2 vs. 3) was the lowest, followed by tone 1 vs. 2 and tone 1 vs. 3. As for the CI group, the differentiability between tone 2 and tone 3 was also the lowest, followed by tone 1 vs. 3 and tone 1 vs. 2. With the 3D method (**Figure 3**, lower panel), the lowest differentiability was found in tone 2 vs. 3 again for both groups, but there were only minor differences among the other five tone contrasts (i.e., tone 1 vs. 2, tone 1 vs. 3, tone 1 vs. 4, tone 2 vs. 4, and tone 3 vs. 4).

A two-way ANOVA showed significant main effects of both subject group ($F = 490.41$, $p < 0.001$) and method used ($F = 180.68$, $p < 0.001$). Additionally, the interaction between these two factors was not significant ($F = 0.28$, $p = 0.595$). A one-way repeated-measures ANOVA was adopted under each method condition to evaluate further the effects of different tone contrasts on differentiability. For both groups, the one-way repeated-measures ANOVA showed significant differences among the six tone contrasts under both methods (all $p < 0.001$). For the 2D method, *post hoc* comparisons indicated that tone 2 vs. 3 yielded the lowest differentiability score, whereas tone 1 vs. 4 and tone 2 vs. 4 produced higher differentiability scores. For the 3D method, the differentiability score for tone 2 vs. 3 was significantly lower than those of all other tone contrasts.

Hit Rate of Tone Production in Normal-Hearing and Cochlear Implant Groups

Figure 4 shows the confusion matrices of the calculated tone hit rates. A two-way ANOVA revealed significant main effects of both subject group ($F = 499.26$, $p < 0.001$) and method used ($F = 145.08$, $p < 0.001$). The interaction between these two factors was not significant ($F = 2.98$, $p = 0.084$). A one-way repeated-measures ANOVA was performed under each method to evaluate further the effects of different tone types on hit rate. For both groups, the one-way repeated-measures ANOVA revealed significant differences among the four tone types under both methods (all $p < 0.001$). For both NH and CI groups, *post hoc* comparisons indicated that different hit rates among tone types were mainly due to the significantly higher hit rates of tone 1 and tone 4.

Correlational and Generalized Linear Model Analyses

Pearson correlations were performed between the averaged tone differentiability and hit rate and between 2D and 3D methods. The averaged tone differentiability used here was the average tone differentiability across all six tone contrasts, and the averaged hit rate here was the average value along the diagonal line in the confusion matrix. **Figure 5** shows these correlational analyses for the NH and CI groups. The differentiability scores and the hit rates in both NH and CI groups were highly correlated for both 2D and 3D methods (**Figure 5**, upper panels). In addition, the differentiability scores derived from the 2D and 3D methods were highly correlated with each other (**Figure 5**, lower left panels). Likewise, the hit rates derived from the 2D and 3D methods were also highly correlated with each other (**Figure 5**, lower right panels). Note that all these correlations depicted in **Figure 5** were statistically highly significant (all $p < 0.0001$).

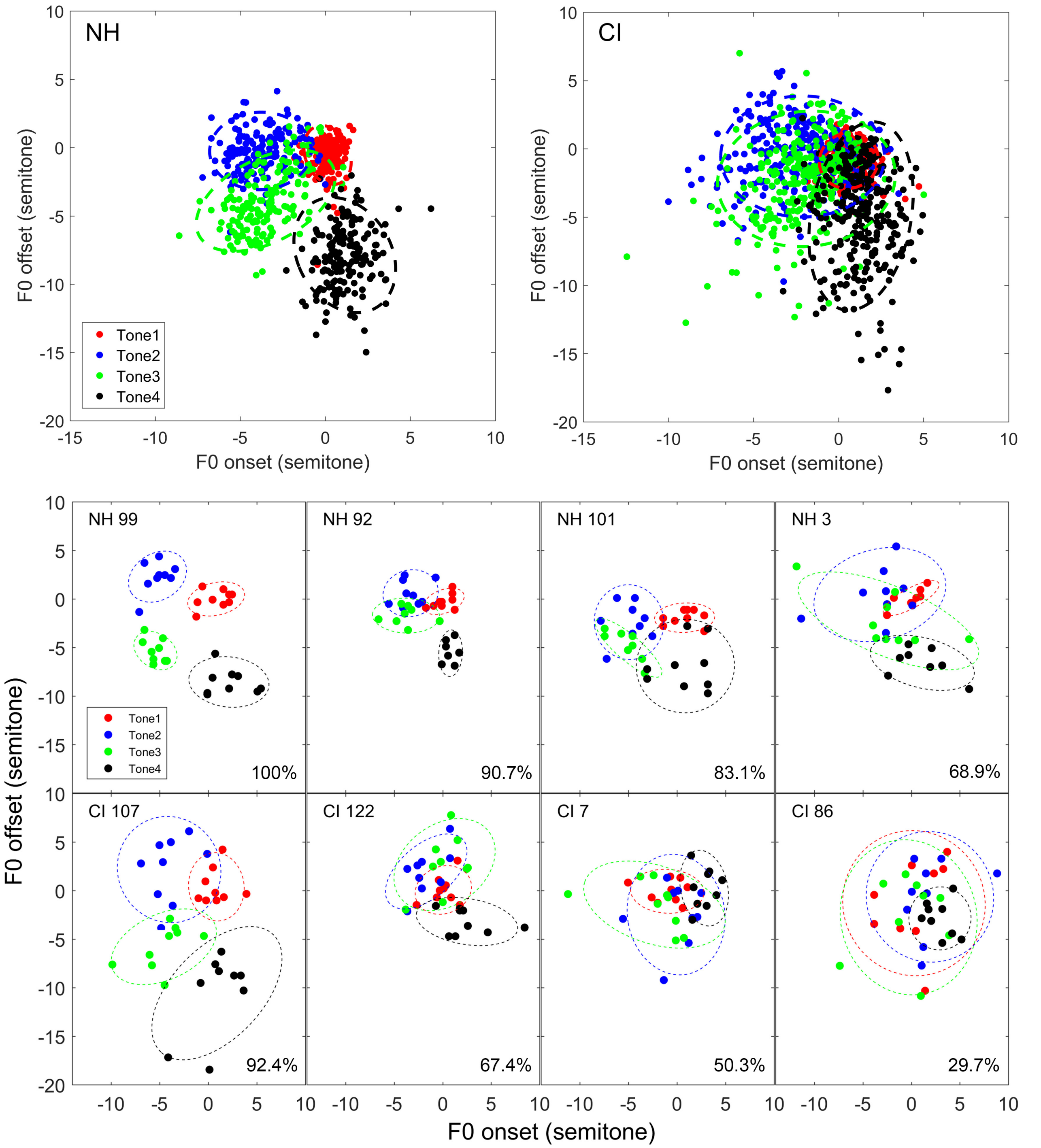

FIGURE 1 | Tone ellipses based on F0 onset and offset values (2D method). Upper panels: Tone ellipses of the NH **(left)** and CI **(right)** groups. Each symbol represents the mean data of one subject. Lower panels: Tone ellipses of individual representative subjects. For either NH or CI group, four subjects were randomly selected based on their mean tone differentiability scores in the interval of 75th–100th, 50th–75th, 25th–50th, and 0–25th quantiles in the respective groups. Mean differentiability score is displayed in the lower right corner. Each data point represents a pair of F0 onset–offset values of a produced token. Four different colors represent four different tone types, as indicated in the legend.

Pearson correlations were also performed between demographic factors of the CI group (i.e., age at implantation, chronological age, and CI use duration) and acoustic indices (i.e., tone differentiability and hit rate) under both 2D and 3D method conditions. Under both method conditions, age at implantation was significantly negatively correlated with

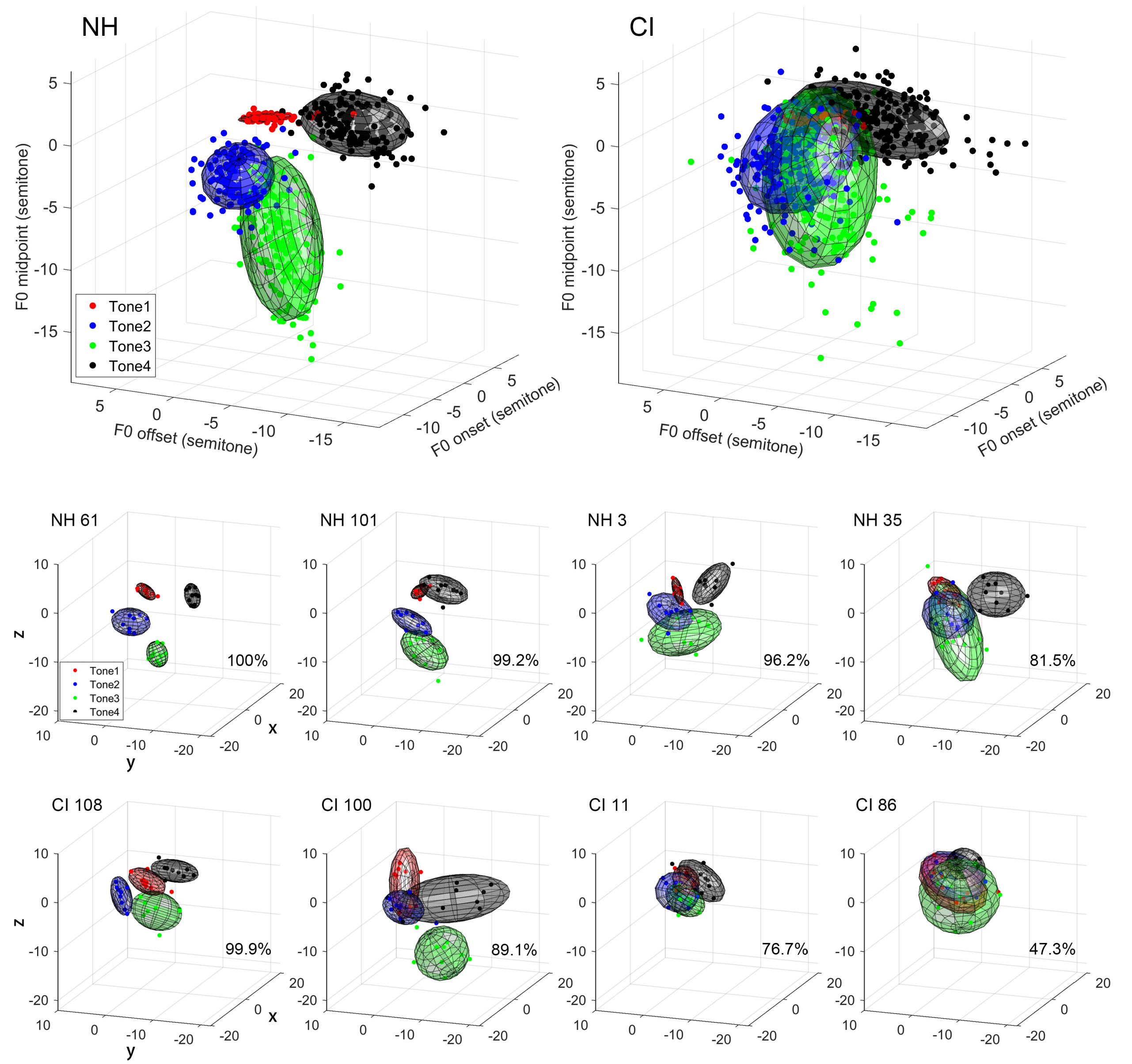

FIGURE 2 | Tone ellipsoids based on F0 onset, midpoint, and offset values (3D method). Upper panels: Tone ellipsoids of the NH **(left)** and CI **(right)** groups. Each symbol represents the mean data of one subject. Lower panels: Tone ellipsoids of individual representative subjects. For either NH or CI group, four subjects were randomly selected based on their mean tone differentiability scores in the interval of 75th–100th, 50th–75th, 25th–50th, and 0–25th quantiles in the respective groups. Mean differentiability score is displayed on the lower right corner. Each data point represents F0 onset–midpoint–offset values of a produced token. Four different colors represent four different tone types as indicated in the legend.

tone differentiability but not correlated with the hit rate. Chronological age was significantly positively correlated with hit rate but not correlated with tone differentiability, whereas CI use duration was significantly positively correlated with both tone differentiability and hit rate. The corresponding correlation coefficients *r* and *p* values are summarized in **Table 1**. Note that these three demographic factors were not independent of each other. Chronological age was equal to the sum of age at implantation and duration of CI use. Thus, we should interpret these correlations with caution. Although these correlational analyses reveal some significant correlations, the absolute values of the correlation coefficients were small, ranging from 0.165 to 0.343. Note also that the *p*-values in **Table 1** are not corrected for multiple comparisons. If we had performed multiple comparisons with Bonferroni correction, then only the correlations related to the duration of CI use would be significant at $p < 0.05$.

Age at implantation and duration of CI use was further subject to GLM analyses (**Table 2**). In these analyses, the duration of CI use but not age at implantation showed significant effects

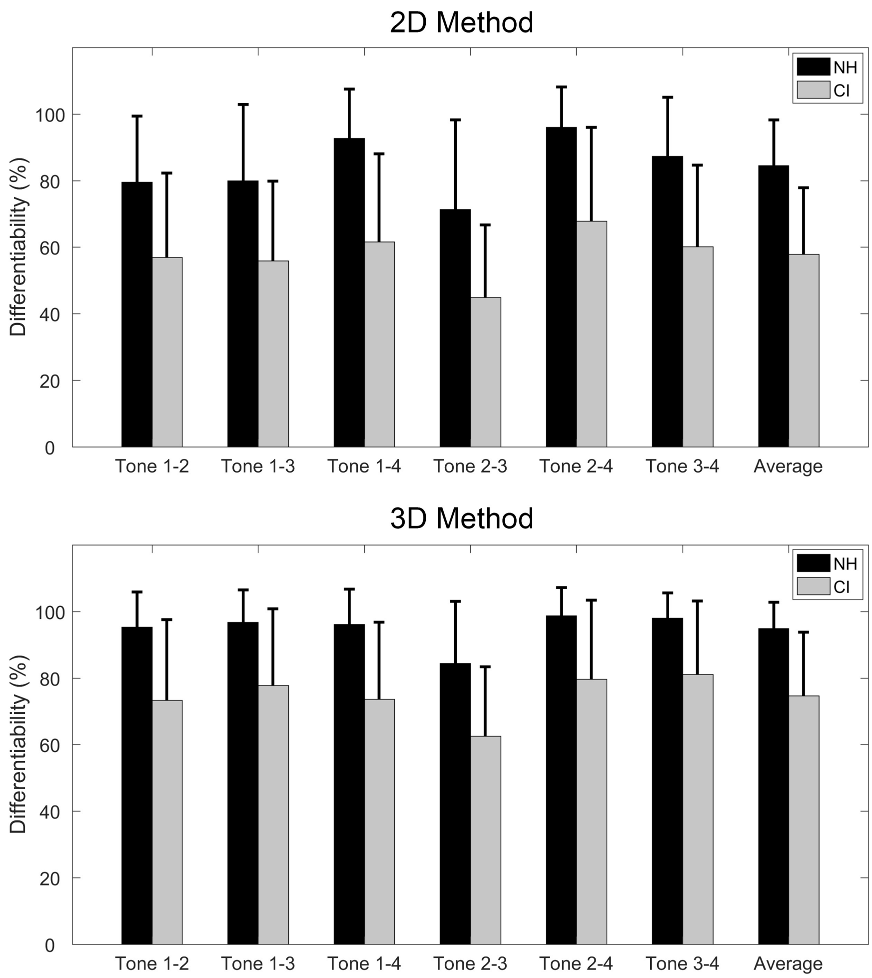

FIGURE 3 | Tone differentiability of the six tone contrasts, as well as the average differentiability, based on the 2D method **(upper panel)** and 3D method **(lower panel)**. Black and gray bars represent the NH and CI groups, respectively. The error bars stand for 1 standard deviation.

on tone production performance (tone differentiability and hit rate) under both 2D and 3D methods (all $p < 0.0001$). Although the age at implantation was not a significant predictor of tone performance, the interaction of these two factors did play a significant role in the tone production performance of the CI children. The R^2 in the GLM was between 0.154 and 0.189, with both ages at implantation, duration of CI use, and their interactions in the model. These results indicated that age at implantation and CI use duration could jointly explain approximately 15 to 19% of the total variance for the tone production outcomes.

DISCUSSION

The present study examined the acoustic properties of tone production in a large group of prelingually deafened children with CIs (N = 278). A large group of age-matched children with NH (N = 173) was also included as controls. Many acoustic features, such as duration, amplitude contour, and spectral envelope, are associated with lexical tones; however, F0 was the most important acoustic correlate for tones (Xu and Zhou, 2011; Yang et al., 2017b). Based on the F0 data, two methods (2D and 3D) were developed to quantify tone differentiability and tone hit rate of the tone production. Results showed that the 3D method produced consistently higher scores than the 2D method in both tone differentiability and tone hit rate. The children with CIs had much lower scores in tone differentiability and hit rate than the NH children. Tone differentiability in the children with CIs revealed that tone contrast in tone 2 vs. 3 yielded the lowest scores. Tone hit rate revealed that the production of tones 2 and 3 was most often confused with each other in the children with CIs. Both acoustic measures (i.e., average tone differentiability and tone hit rate) showed a weak correlation with duration of CI use, whereas the average tone differentiability showed a weak correlation with the age of implantation. Later, we compare the acoustic findings of the two methods (2D and 3D) and then discuss the tone production proficiency of prelingually deafened children with CIs related to the demographic factors.

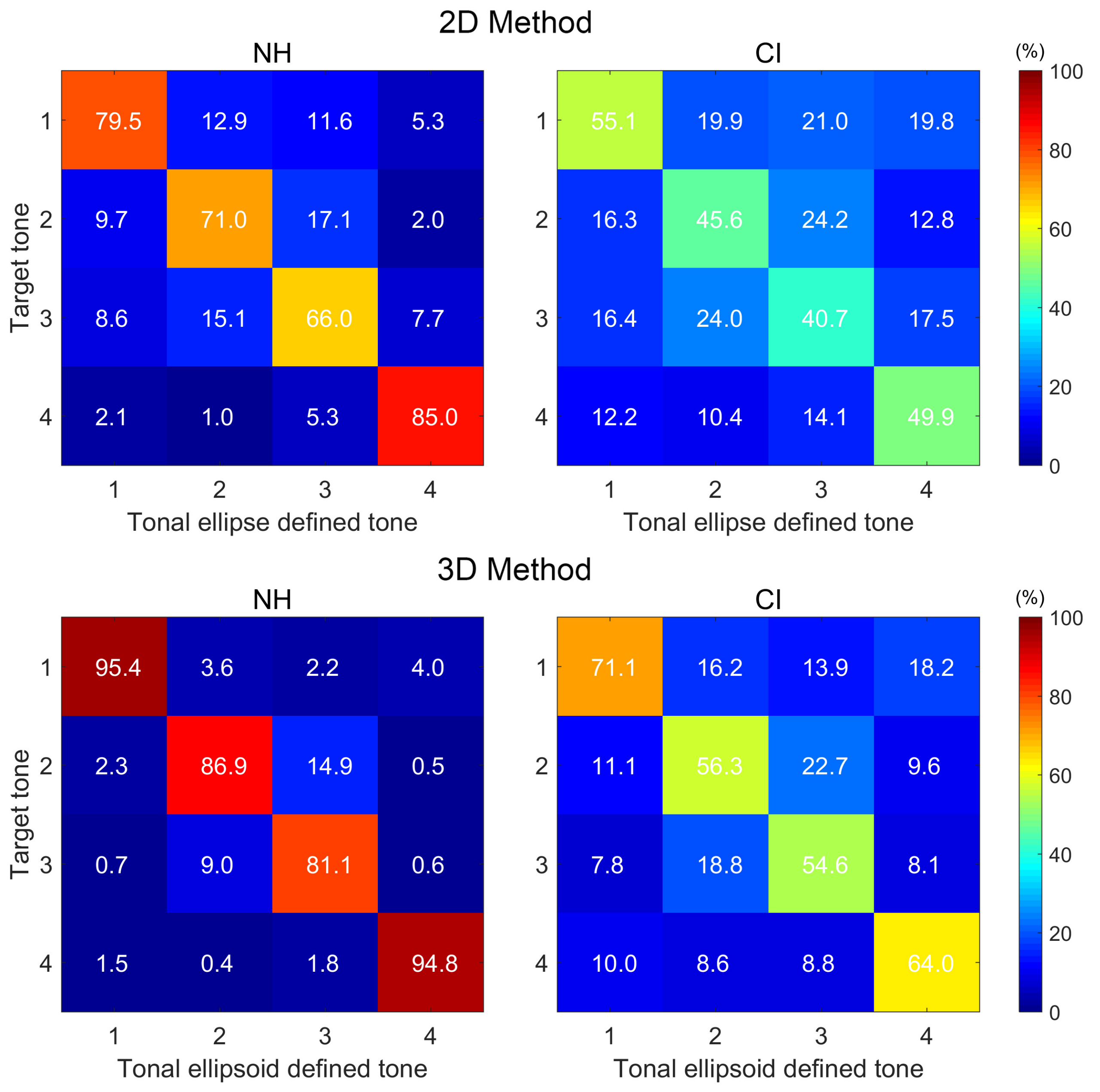

FIGURE 4 | Tone hit-rate confusion matrices of the two groups based on the 2D **(upper panels)** and 3D methods **(lower panels)**. Data were averaged values of all subjects in each group. For each panel of 4 × 4 cells, the columns represent the tone categories defined by the tonal ellipses or the tonal ellipsoids (i.e., within the tonal ellipses or ellipsoids), whereas the rows represent the target tone categories. Value in the cell of row *i* and column *j* is the proportion of target tone *i* inside tonal ellipse (or ellipsoid) *j* (*i* = 1, 2, 3, or 4; *j* = 1, 2, 3, or 4).

There are various ways to evaluate different aspects of Mandarin tone characteristics (e.g., Tupper et al., 2020). Zhou and Xu (2008) used the F0 onset and F0 offset of the F0 contour to analyze the produced tones acoustically. Their methods calculated a series of acoustic indices to assess the overall differentiability among tone types and averaged the hit rate of tones. The present study extended Zhou and Xu (2008) study by expanding the 2D dataset to a 3D dataset, modifying the algorithms of acoustic indices, and greatly enlarging the sample size of the subjects. In Zhou and Xu (2008) study, both Index 1 and Index 2 reflected the overall differentiability among the four tones and significantly correlated with each other (r = 0.94, p < 0.001). Therefore, there might be some degree of redundancy. In the present study, we modified the algorithm for tone differentiability. We scaled this index into percentage data (see *Methods* for details) to make the index more intuitive. Our modification also allowed us to derive tone contrast differentiability scores. For the tone hit rate, we basically followed

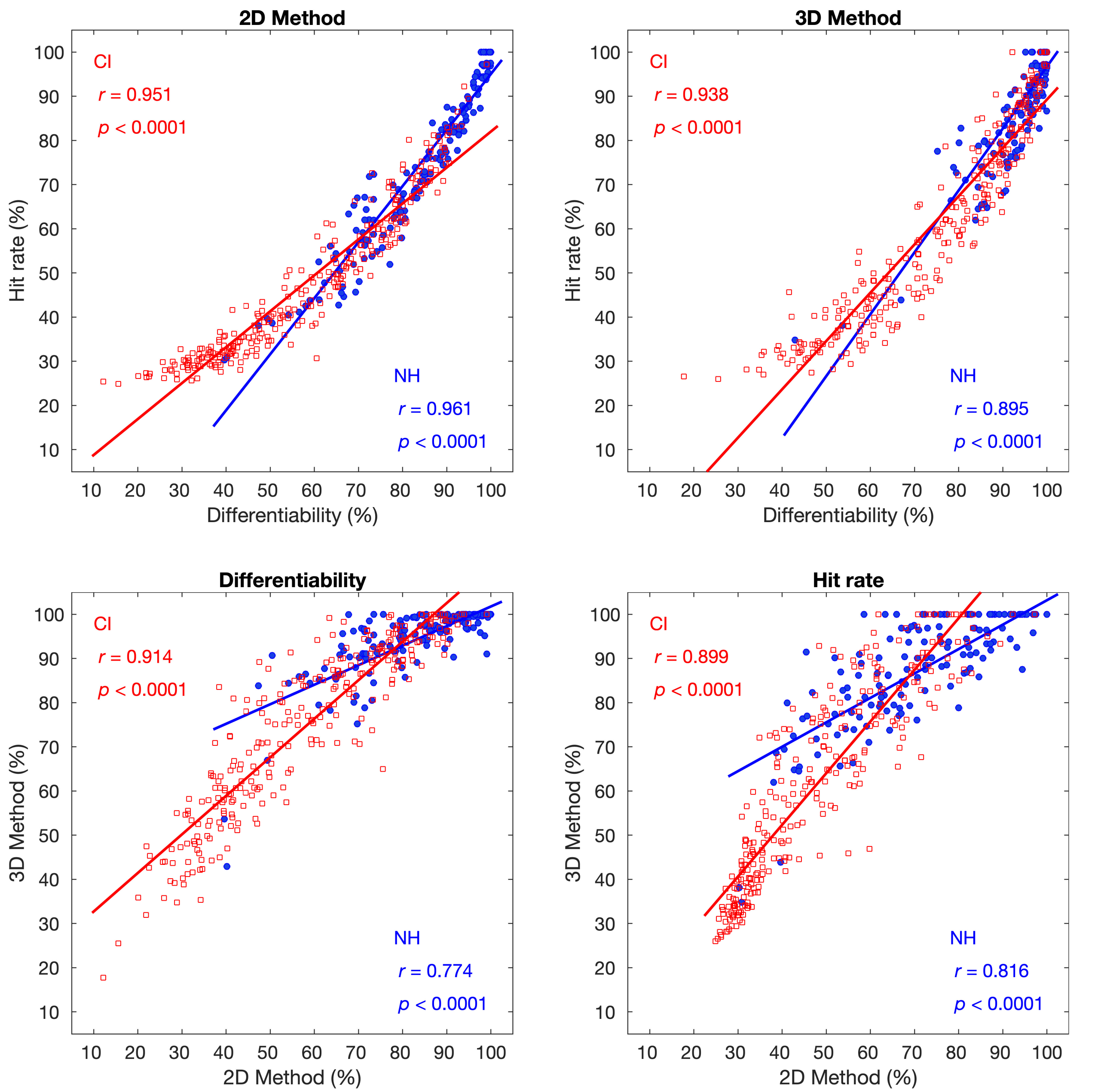

FIGURE 5 | Upper panels: Correlation between averaged tone differentiability and tone hit rate in the 2D **(left)** and 3D methods **(right)**. Lower panels: Correlation between the 2D and 3D methods in tone differentiability **(left)** and hit rate **(right)**. In each panel, each symbol represents one subject. Red open symbols represent subjects in the CI group, and the filled blue symbols represent subjects in the NH group. Red and blue lines are the linear fit of the CI and NH group data, respectively. Correlation coefficient *r* and the corresponding *p*-value for the CI group are shown on the upper left corner, and those for the NH group are shown in the lower right corner.

the Zhou and Xu (2008) algorithm but further measured the hit rate for each of the four tones rather than just the averaged hit rate. These results were shown in the form of confusion matrices. Another innovation of the present study was that the acoustic indices were calculated based on both 2D and 3D methods, thus making it possible to explore whether the 3D method would better highlight the distinctive characteristics of the tone production of CI children from their NH peers. We expected that the 3D method would improve the scores of the NH group but not the CI group compared with the 2D method, as pediatric CI users tended to produce flat tone contours (Xu et al., 2004; Tang et al., 2019), and the introduction of the middle point of F0 contour would not make any differences on flat tone contours. However, our results showed that the 3D method also improved the scores of the CI group similarly to the NH group (see **Figures 3**, **4**). Further analyses showed that, although the

TABLE 1 | The correlation coefficients *r* and *p* values under both Method conditions.

Method	Index	Age at implantation	Chronological age	Duration of CI use
2D	Differentiability	$r = -0.179$ $p = 0.014$*	$r = 0.105$ $p = 0.152$	$r = 0.291$ $p < 0.001$**
	Hit rate	$r = -0.125$ $p = 0.089$	$r = 0.176$ $p = 0.015$*	$r = 0.339$ $p < 0.001$**
3D	Differentiability	$r = -0.165$ $p = 0.023$*	$r = 0.121$ $p = 0.099$	$r = 0.295$ $p < 0.001$**
	Hit rate	$r = -0.108$ $p = 0.139$	$r = 0.193$ $p = 0.008$*	$r = 0.343$ $p < 0.001$**

** indicates a p of <0.05, ** indicates a p of <0.001.*

TABLE 2 | Results of the GLM analyses.

Method	Dependent variable	Independent variable	Coefficient (β)	*t*	*p*	R^2
2D		Age at implantation	0.006	0.631	0.529	0.163
	Differentiability	Duration of CI use	0.050	4.851	<0.0001**	
		A × D$^{\Delta}$	−0.007	−2.952	0.004*	
		Age at implantation	0.011	1.266	0.207	0.178
	Hit rate	Duration of CI use	0.046	5.372	<0.0001**	
		A × D$^{\Delta}$	−0.006	−3.199	0.002*	
3D		Age at implantation	0.010	0.862	0.390	0.154
	Differentiability	Duration of CI use	0.569	4.773	<0.0001**	
		A × D$^{\Delta}$	−0.007	−2.696	0.008*	
		Age at implantation	0.025	1.953	0.052	0.189
	Hit rate	Duration of CI use	0.076	5.833	<0.0001**	
		A × D$^{\Delta}$	−0.010	−3.602	0.0004**	

*$^{\Delta}$A × D stands for the interaction between Age at implantation and Duration of CI use. * indicates a p of <0.05, ** indicates a p of <0.001.*

scores of the 3D method were highly correlated with those of the 2D method, the amount of improvement of the 3D method over the 2D method varied greatly (**Figure 5**). Note that there were extreme cases in which a small proportion of subjects in the CI group either improved as much as >30 percentage points or decreased in scores with the 3D method compared with the 2D method. The two types of extreme cases indicated that errors in tone production might occur at different time segments in the F0 contour. Therefore, the 2D method, combined with the 3D method, might provide useful information to guide the tone rehabilitation process for the hearing-impaired children with CIs.

Through the two-way ANOVA, the two main effects [i.e., (1) NH and CI groups and (2) 2D and 3D methods] on tone differentiability were highly significant. These results indicated that the overall tone differentiability of NH children was significantly better than that of children with CIs and that the 3D method yielded significantly higher tone differentiability scores than the 2D method. It was likely that the 2D method might have underestimated the differentiability among tones, as only two data points (F0 onset and offset) of the F0 contour were used. In addition, we found that no matter what method it was based on, the differentiability of tone 2 vs. tone 3 was always the lowest, both for NH and CI groups. Interestingly, this finding was quite similar to the results of our previous tone-perception study in children with CIs (Mao and Xu, 2017); that is, the recognition rate of the tone 2 vs. tone 3 contrast was the lowest among the six tone contrasts. In addition, the differentiability scores derived from the 3D method were more similar to the absolute values of tone-recognition performance of the six tone contrasts for both NH and CI children. This finding implied that auditory perception was likely to play a decisive role in the acquisition of tone production and that the 3D method might be more precise in reflecting the true tone differentiability.

For tone hit rate, it was observed that the two main effects [i.e., (1) NH and CI groups and (2) 2D and 3D methods] on tone hit rate were also significant. These results illustrated that the averaged hit rate of the NH group was significantly higher than that of the CI group, and the averaged hit rate calculated based on the 3D method was significantly higher than that based on the 2D method. Under either method, the error pattern of tone production was similar to the tone-perception error pattern, except that the values of the diagonal in the tone-production matrices (**Figure 4**) were, in general, lower than that of the tone-perception matrices (Mao and Xu, 2017). Confusion of tone 2 and tone 3 with each other was the most prominent error for both NH and CI children. This was consistent with the findings for tone differentiability. Several earlier studies had found that the F0 contours of the four tones produced by pediatric CI users tended to be flat (Xu et al., 2004; Tang et al., 2019). This may be largely related to pitch perception deficits in CI users. Therefore, when CI children do not perceive the F0 contours of all tone types, their production tends to be flat with little pitch variation across the duration of the syllables. The flat pitch production was exacerbated when prelingually deafened children with CIs were asked to sing a song (Nakata et al., 2006; Xu et al., 2009; Mao et al., 2013), although some of them can achieve normal pitch production after rigorous, long-term training (Yang et al., 2019).

The present study analyzed the potential relationships within the acoustic indices. Not surprisingly, the average tone differentiability scores and the average hit rate were highly correlated for both groups (**Figure 5**, upper panels), similar to the findings among different indices in Zhou and Xu (2008) study. Although these two metrics were highly correlated, they provided insights into different aspects of tone production. Tone differentiability focused on the tone contrast, whereas the hit rate allowed us to construct the tone confusion matrix. For both acoustic measures, the scores between 2D and 3D methods were also highly correlated in both groups (**Figure 5**, lower panels). In the CI group, the 3D method yielded scores, on average, 16.6 and 13.6 percentage points higher than those of the 2D method for tone discrimination and hit rate, respectively. The higher scores produced by the 3D method might correspond more closely to the tone-perception outcomes of the CI children (Mao and Xu, 2017). However, the strong correlation between the scores of the 2D and 3D methods suggests that the simpler, 2D method might be efficient in clinical practice, whereas the 3D method might provide complementary information for an acoustic assessment of lexical tones.

There is abundant evidence showing that tonal ability of pediatric CI users is correlated with age at implantation or device use duration (Peng et al., 2004; Han et al., 2007, 2009; Lee et al., 2010; Xu et al., 2011; Zhou et al., 2013; Li et al., 2017; Mao and Xu, 2017; Tang et al., 2019), although the literature is not always consistent on the contributions of these two predictors. For example, Peng et al. (2004) and Tang et al. (2019) found that age at implantation was the only significant predictor for tone production, whereas Han et al. (2007) revealed that CI use duration also significantly predicted the tone production ability. For tone perception, several studies showed that age at implantation exerted a weak effect on CI users' tone perception ability, whereas the duration of CI use seemed to be a more robust predictor (Zhou et al., 2013; Li et al., 2017). An earlier study by Wong and Wong (2004) did not show any correlations of Cantonese tone identification with either implantation age or duration of CI use. However, more recent evidence encouraged early implantation for children with severe to profound sensorineural deafness for tonal ability rehabilitation (Peng et al., 2004; Han et al., 2007, 2009; Lee et al., 2010; Xu et al., 2011; Zhou et al., 2013; Li et al., 2017; Mao and Xu, 2017; Tang et al., 2019). Our results from the present study with a fairly large sample of subjects showed that duration of CI use was a significant predictor for both tone differentiability and hit rate, whereas age at implantation and chronological age seemed to be the weaker predictors, as they were only correlated with one of the indices (**Table 1**). Some of our CI participants who were implanted at an earlier age happened to be at a younger chronological age at the test, which could counteract part of the benefit of earlier implantation. The GLM analyses demonstrated that in the presence of duration of CI use, the effect of age at implantation was not manifested, but it had a significant interaction with a duration of CI use, and jointly, these two factors accounted for approximately 15–19% of the total variance of tone production performance of the CI children. Interestingly, for tone perception ability, earlier evidence had shown that these two variables could jointly explain approximately 50% of the outcome variance (Xu and Zhou, 2011). The difference of more than 30% of the variance interpretability between perception and production by the temporal factors might reside in the differences in the time course of development of perception and production, among other non-temporal factors. Generally, our results illustrated that persistent CI use might play a key role in developing vocal production of Mandarin tones for those implanted children. Besides lexical tone-related aspects, previous studies also supported the persistent use of CI devices for the music-related development of pediatric implantees (Yucel et al., 2009; Chen et al., 2010; Mao et al., 2013). The effects of duration of use could be attributed to their maturity, persistent training and learning, and increased experiences with time. Overall, our results supported early implantation and continual use of CI devices in the tone rehabilitation process of pediatric CI users. It is noteworthy that, compared with the NH peers, the children with CIs might still demonstrate deficits in tone production even after prolonged use of the devices despite the significant progress in CI technology in recent years.

CONCLUSION

The present study modified a previous 2D method and developed a new 3D method to assess lexical tone production in children with CIs. Two acoustic measures (i.e., tone differentiability and hit rate) were derived from the 2D and 3D methods. With a relatively large sample size, our results confirmed that the tone production ability of the CI children was significantly inferior to that of the normally developed children in both tone differentiability and hit rate. The scores obtained with the 2D and 3D methods were highly correlated, suggesting that the simpler, 2D method would be efficient in capturing the main acoustic characteristics of tone production and might be more practical for clinical assessment of lexical tone production. However, the 3D method might provide complementary information for the tone production deficits when combined with the 2D method. The tone differentiability and hit rate, although capturing different aspects of tone production, were also highly correlated with each other. Age at implantation and especially the duration of CI use were important predictors for tone-production ability but could only account for 15 to 19% of the variance. Other factors such as rehabilitation or training on tone production, mother's education, children's IQ, residual hearing, etc., should be explored in future studies of tone development in prelingually deafened children with CIs.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Institutional Review Board of Ohio University.

Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

YM and LX conceived the study, designed the analysis, and wrote the manuscript. LX collected the data. YM, SX, and HC obtained funding for this research. All authors discussed the results and implications and commented on the article at all stages.

FUNDING

The study was supported, in part, by the National Natural Science Foundation of China (81701847) and the Natural Science Foundation of Hunan Province (2017JJ3497).

REFERENCES

Amir, O., Amir, N., and Kishon-Rabin, L. (2003). The effect of superior auditory skills on vocal accuracy. *J. Acoust. Soc. Am.* 113, 1102–1108. doi: 10.1121/1.1536632

Barry, J. G., and Blamey, P. J. (2004). The acoustic analysis of tone differentiation as a means for assessing tone production in speakers of Cantonese. *J. Acoust. Soc. Am.* 116, 1739–1748. doi: 10.1121/1.1779272

Chen, J. K., Chuang, A. Y. C., McMahon, C., Hsieh, J., Tung, T., and Li, L. P. (2010). Music training improves pitch perception in prelingually deafened children with cochlear implants. *Pediatrics* 125, e793–e800. doi: 10.1542/peds.2008-3620

Chen, Y., and Wong, L. L. N. (2017). Speech perception in Mandarin-speaking children with cochlear implants: a systematic review. *Int. J. Audiol.* 56, S7–S16. doi: 10.1080/14992027.2017.1300694

Davidson, G. D. (1959). Sidetone delay and reading rate, articulation, and pitch. *J. Speech Hear. Res.* 2, 266–270. doi: 10.1044/jshr.0203.266

Deroche, M. L. D., Lu, H., Lin, Y., Chatterjee, M., and Peng, S. (2019). Processing of acoustic information in lexical tone production and perception by pediatric cochlear implant recipients. *Front. Neurosci.* 13:639. doi: 10.3389/fnins.2019.00639

Elman, J. L. (1981). Effects of frequency-shifted feedback on the pitch of vocal productions. *J. Acoust. Soc. Am.* 70, 45–50. doi: 10.1121/1.386580

Han, D., Liu, B., Zhou, N., Chen, X., Kong, Y., Liu, H., et al. (2009). Lexical tone perception with HiResolution and HiResolution 120 sound-processing strategies in pediatric Mandarin-speaking cochlear implant users. *Ear Hear.* 30, 169–177. doi: 10.1097/AUD.0b013e31819342cf

Han, D., Zhou, N., Li, Y., Chen, X., Zhao, X., and Xu, L. (2007). Tone production of Mandarin Chinese speaking children with cochlear implants. *Int. J. Pediatr. Otorhinolaryngol.* 71, 875–880. doi: 10.1016/j.ijporl.2007.02.008

Holt, C. M., Lee, K. Y. S., Dowell, R. C., and Vogel, A. P. (2018). Perception of Cantonese lexical tones by pediatric cochlear implant users. *J. Speech Lang. Hear. Res.* 61, 174–185. doi: 10.1044/2017_JSLHR-H-17-0027

Houde, J. F., and Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science* 279, 1213–1216. doi: 10.1126/science.279.5354.1213

Lane, H., and Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *J. Speech Hear. Res.* 14, 677–709. doi: 10.1044/jshr.1404.677

Leder, S. B., and Spitzer, J. B. (1993). Speaking fundamental frequency, intensity, and rate of adventitiously profoundly hearing-impaired adult women. *J. Acoust. Soc. Am.* 93, 2146–2151. doi: 10.1121/1.406677

Lee, K. Y. S., van Hasselt, C. A., Chiu, S. N., and Cheung, D. M. C. (2002). Cantonese tone perception ability of cochlear implant children in comparison with normal-hearing children. *Int. J. Pediatr. Otorhinolaryngol.* 63, 137–147. doi: 10.1016/s0165-5876(02)00005-8

Lee, K. Y. S., van Hasselt, C. A., and Tong, M. C. F. (2010). Age sensitivity in the acquisition of lexical tone production: evidence from children with profound congenital hearing impairment after cochlear implantation. *Ann. Otol. Rhinol. Laryngol.* 119, 258–265. doi: 10.1177/000348941011900409

Li, G., Soli, S. D., and Zheng, Y. (2017). Tone perception in Mandarin-speaking children with cochlear implants. *Int. J. Audiol.* 56, S49–S59. doi: 10.1080/14992027.2017.1324643

Limb, C. J., and Roy, A. T. (2014). Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hear. Res.* 308, 13–26. doi: 10.1016/j.heares.2013.04.009

Liu, D., Dai, G., Liu, C., Guo, Z., Xu, Z., Jones, J. A., et al. (2020). Top-down inhibitory mechanisms underlying auditory-motor integration for voice control: evidence by TMS. *Cereb. Cortex* 30, 4515–4527. doi: 10.1093/cercor/bhaa054

Liu, H., Peng, X., Zhao, Y., and Ni, X. (2017). The effectiveness of sound-processing strategies on tonal language cochlear implant users: a systematic review. *Pediatr. Investigation* 1, 32–39. doi: 10.1002/ped4.12011

Mao, Y., Chen, Z., and Xu, L. (2017). The application of neural network analysis on lexical tone rehabilitation of pediatric cochlear implant users. *Chin. J. Otorhinolaryngol. Head Neck Surg.* 52, 573–579. doi: 10.3760/cma.j.issn.1673-0860.2017.08.003

Mao, Y., and Xu, L. (2017). Lexical tone recognition in noise in normal-hearing children and prelingually deafened children with cochlear implants. *Int. J. Audiol.* 56, S23–S30. doi: 10.1080/14992027.2016.1219073

Mao, Y., Zhang, M., Nutter, H., Zhang, Y., Zhou, Q., Liu, Q., et al. (2013). Acoustic properties of vocal singing in prelingually-deafened children with cochlear implants or hearing aids. *Int. J. Pediatr. Otorhinolaryngol.* 77, 1833–1840. doi: 10.1016/j.ijporl.2013.08.022

Matthies, M. L., Svirsky, M. A., Lane, H. L., and Perkell, J. S. (1994). A preliminary study of the effects of cochlear implants on the production of sibilants. *J. Acoust. Soc. Am.* 96, 1367–1373. doi: 10.1121/1.410281

Mora, R., Crippa, B., Cervoni, E., Santomauro, V., and Guastini, L. (2012). Acoustic features of voice in patients with severe hearing loss. *J. Otolaryngol. Head Neck Surg.* 41, 8–13. doi: 10.1016/j.otohns.2010.06.111

Nakata, T., Trehub, S. E., Mitani, C., and Kanda, Y. (2006). Pitch and timing in the songs of deaf children with cochlear implants. *Music Percept* 24, 147–154. doi: 10.1525/mp.2006.24.2.147

Natke, U., and Kalveram, K. T. (2001). Effects of frequency-shifted auditory feedback on fundamental frequency of long stressed and unstressed syllables. *J. Speech Lang. Hear. Res.* 44, 577–584. doi: 10.1044/1092-4388(2001/045)

Nava, E., Bottari, D., Villwock, A., Fengler, I., Buchner, A., Lenarz, T., et al. (2014). Audio-tactile integration in congenitally and late deaf cochlear implant users. *PLoS One* 9:e99606. doi: 10.1371/journal.pone.0099606

Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N., Quittner, A. L., et al. (2010). Spoken language development in children following cochlear implantation. *JAMA* 303, 1498–1506. doi: 10.1001/jama.2010.451

Osberger, M. J., Maso, M., and Sam, L. K. (1993). Speech intelligibility of children with cochlear implants, tactile aids, or hearing aids. *J. Speech Hear. Res.* 36, 186–203. doi: 10.1044/jshr.3601.186

Peng, S. C., Tomblin, J. B., Cheung, H., Lin, Y. S., and Wang, L. S. (2004). Perception and production of mandarin tones in prelingually deaf children with cochlear implants. *Ear Hear.* 25, 251–264. doi: 10.1097/01.aud.0000130797.73809.40

Plant, G. (1993). The speech of adults with acquired profound hearing losses. I: a perceptual evaluation. *Eur. J. Disord. Commun.* 28, 273–288. doi: 10.3109/13682829309060041

Selleck, M. A., and Sataloff, R. T. (2014). The impact of the auditory system on phonation: a review. *J. Voice* 28, 688–693. doi: 10.1016/j.jvoice.2014.03.018

Song, L., Langfelder, P., and Horvath, S. (2013). Random generalized linear model: a highly accurate and interpretable ensemble predictor. *BMC Bioinformatics.* 14:5. doi: 10.1186/1471-2105-14-5

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *J. Speech Hear. Res.* 28, 455–462. doi: 10.1044/jshr.2803.455

Tan, J., Dowell, R., and Vogel, A. (2016). Mandarin lexical tone acquisition in cochlear implant users with prelingual deafness: a review. *Am. J. Audiol.* 25, 246–256. doi: 10.1044/2016_AJA-15-0069

Tang, P., Yuen, I., Xu, R. N., Gao, L., and Demuth, K. (2019). The acquisition of Mandarin tonal processes by children with cochlear implants. *J. Speech Lang. Hear. Res.* 62, 1309–1325. doi: 10.1044/2018_JSLHR-S-18-0304

Tobey, E. A., Pancamo, S., Staller, S. J., Brimacombe, J. A., and Beiter, A. L. (1991). Consonant production in children receiving a multichannel cochlear implant. *Ear Hear.* 12, 23–31. doi: 10.1097/00003446-199102000-00003

Tupper, P., Leung, K., Wang, Y., Jongman, A., and Sereno, J. A. (2020). Characterizing the distinctive acoustic cues of Mandarin tones. *J. Acoust. Soc. Am.* 147, 2570–2580. doi: 10.1121/10.0001024

Uchanski, R. M., and Geers, A. E. (2003). Acoustic characteristics of the speech of young cochlear implant users: a comparison with normal-hearing age-mates. *Ear Hear* 24, 90S–105S. doi: 10.1097/01.AUD.0000051744.24290.C1

Waldstein, R. S. (1990). Effects of postlingual deafness on speech production: implications for the role of auditory feedback. *J. Acoust. Soc. Am.* 88, 2099–2114. doi: 10.1121/1.400107

Wei, W. I., Wong, R., Hui, Y., Au, D. K., Wong, B. Y., Ho, W. K., et al. (2000). Chinese tonal language rehabilitation following cochlear implantation in children. *Acta Otolaryngol.* 120, 218–221. doi: 10.1080/000164800750000955

Wilson, B. S. (2019). The remarkable cochlear implant and possibilities for the next large step forward. *Acoustics Today* 15, 53–61. doi: 10.1121/AT.2019.15.1.55

Wilson, B. S., and Dorman, M. F. (2008). Cochlear implants: a remarkable past and a brilliant future. *Hear. Res.* 242, 3–21. doi: 10.1016/j.heares.2008.06.005

Wong, A. O. C., and Wong, L. L. N. (2004). Tone perception of Cantonese-speaking prelingually hearing-impaired children with cochlear implants. *Otolaryngol. Head Neck Surg.* 130, 751–758. doi: 10.1016/j.otohns.2003.09.037

Xu, L., Chen, X., Lu, H., Zhou, N., Wang, S., Liu, Q., et al. (2011). Tone perception and production in pediatric cochlear implants users. *Acta Otolaryngol.* 131, 395–398. doi: 10.3109/00016489.2010.536993

Xu, L., Chen, X., Zhou, N., Li, Y., Zhao, X., and Han, D. (2007). Recognition of lexical tone production of children with an artificial neural network. *Acta Otolaryngol.* 127, 365–369. doi: 10.1080/00016480601011477

Xu, L., Li, Y., Hao, J., Chen, X., Xue, S. A., and Han, D. (2004). Tone production in Mandarin-speaking children with cochlear implants: a preliminary study. *Acta Otolaryngol.* 124, 363–367. doi: 10.1080/00016480410016351

Xu, L., Zhang, W., Zhou, N., Lee, C.-Y., Li, Y., Chen, X., et al. (2006). Mandarin Chinese tone recognition with an artificial neural network. *J. Otol.* 1, 30–34. doi: 10.1016/S1672-2930(06)50005-4

Xu, L., and Zhou, N. (2011). "Tonal languages and cochlear implants," in *Auditory Prostheses: New Horizons*, eds F. Zeng, A. Popper, and R. Fay (New York, NY: Springer Science+Business Media, LLC), 341–364.

Xu, L., Zhou, N., Chen, X., Li, Y., Schultz, H. M., Zhao, X., et al. (2009). Vocal singing by prelingually-deafened children with cochlear implants. *Hear. Res.* 255, 129–134. doi: 10.1016/j.heares.2009.06.011

Yang, J., Brown, E., Fox, R. A., and Xu, L. (2015). Acoustic properties of vowel production in prelingually deafened Mandarin-speaking children with cochlear implants. *J. Acoust. Soc. Am.* 138, 2791–2799. doi: 10.1121/1.4932165

Yang, J., Liang, Q., Chen, H., Liu, Y., and Xu, L. (2019). Singing proficiency of members of a choir formed by prelingually deafened children with cochlear implants. *J. Speech Lang. Hear. Res.* 62, 1561–1573. doi: 10.1044/2019_JSLHR-H-18-0385

Yang, J., Vadlamudi, J., Yin, Z., Lee, C. Y., and Xu, L. (2017a). Production of word-initial fricatives of Mandarin Chinese in prelingually deafened children with cochlear implants. *Int. J. Speech Lang. Pathol.* 19, 153–164. doi: 10.3109/17549507.2016.1143972

Yang, J., and Xu, L. (2017). Mandarin compound vowels produced by prelingually deafened children with cochlear implants. *Int. J. Pediatr. Otorhinolaryngol.* 97, 143–149. doi: 10.1016/j.ijporl.2017.04.012

Yang, J., Zhang, Y., Li, A., and Xu, L. (2017b). "On the duration of Mandarin tones," in *Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech 2017)*, Stockholm, 1407–1411. doi: 10.21437/Interspeech.2017-29

Yucel, E., Sennaroglu, G., and Belgin, E. (2009). The family oriented musical training for children with cochlear implants: speech and musical perception results of two year follow-up. *Int. J. Pediatr. Otorhinolaryngol.* 73, 1043–1052. doi: 10.1016/j.ijporl.2009.04.009

Zhou, N., Huang, J., Chen, X., and Xu, L. (2013). Relationship between tone perception and production in prelingually deafened children with cochlear implants. *Otol. Neurotol.* 34, 499–506. doi: 10.1097/MAO.0b013e318287ca86

Zhou, N., and Xu, L. (2008). Development and evaluation of methods for assessing tone production skills in Mandarin-speaking children with cochlear implants. *J. Acoust. Soc. Am.* 123, 1653–1664. doi: 10.1121/1.2832623

Zhou, N., Zhang, W., Lee, C. Y., and Xu, L. (2008). Lexical tone recognition with an artificial neural network. *Ear Hear.* 29, 326–335. doi: 10.1097/AUD.0b013e3181662c42

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 14 December 2020
doi: 10.3389/fnins.2020.586119

Check for updates

Cochlear Implant and Hearing Aid: Objective Measures of Binaural Benefit

Tobias Balkenhol†, Elisabeth Wallhäusser-Franke*†, Nicole Rotter and Jérôme J. Servais

Department of Otorhinolaryngology Head and Neck Surgery, Medical Faculty Mannheim, University Medical Center Mannheim, Heidelberg University, Mannheim, Germany

Cochlear implants (CI) improve hearing for the severely hearing impaired. With an extension of implantation candidacy, today many CI listeners use a hearing aid on their contralateral ear, referred to as bimodal listening. It is uncertain, however, whether the brains of bimodal listeners can combine the electrical and acoustical sound information and how much CI experience is needed to achieve an improved performance with bimodal listening. Patients with bilateral sensorineural hearing loss undergoing implant surgery were tested in their ability to understand speech in quiet and in noise, before and again 3 and 6 months after provision of a CI. Results of these bimodal listeners were compared to age-matched, normal hearing controls (NH). The benefit of adding a contralateral hearing aid was calculated in terms of head shadow, binaural summation, binaural squelch, and spatial release from masking from the results of a sentence recognition test. Beyond that, bimodal benefit was estimated from the difference in amplitudes and latencies of the N1, P2, and N2 potentials of the brains' auditory evoked response (AEP) toward speech. Data of fifteen participants contributed to the results. CI provision resulted in significant improvement of speech recognition with the CI ear, and in taking advantage of the head shadow effect for understanding speech in noise. Some amount of binaural processing was suggested by a positive binaural summation effect 6 month post-implantation that correlated significantly with symmetry of pure tone thresholds. Moreover, a significant negative correlation existed between binaural summation and latency of the P2 potential. With CI experience, morphology of the N1 and P2 potentials in the AEP response approximated that of NH, whereas, N2 remained different. Significant AEP differences between monaural and binaural processing were shown for NH and for bimodal listeners 6 month post-implantation. Although the grand-averaged difference in N1 amplitude between monaural and binaural listening was similar for NH and the bimodal group, source localization showed group-dependent differences in auditory and speech-relevant cortex, suggesting different processing in the bimodal listeners.

Keywords: cochlear implant, hearing aid, electroencephalography, auditory evoked potentials, source localization, speech recognition, bimodal benefit, auditory rehabilitation

Edited by:
Fei Chen,
Southern University of Science and Technology, China

Reviewed by:
Richard Charles Dowell,
The University of Melbourne, Australia
Yonghee Oh,
University of Florida, United States

****Correspondence:***
Elisabeth Wallhäusser-Franke
elisabeth.wallhaeusser-franke@medma.uni-heidelberg.de

†These authors share first authorship

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *22 July 2020*
Accepted: *15 October 2020*
Published: *14 December 2020*

Citation:
Balkenhol T, Wallhäusser-Franke E, Rotter N and Servais JJ (2020) Cochlear Implant and Hearing Aid: Objective Measures of Binaural Benefit. Front. Neurosci. 14:586119. doi: 10.3389/fnins.2020.586119

INTRODUCTION

Cochlear implants (CI) are hearing prostheses that bypass defective sensory hair cells in the cochlea, allowing individuals with severe to profound sensorineural hearing loss to regain much of their hearing. As CI technology and surgical approaches have advanced, many patients with residual hearing in their opposite ear qualify for implantation. Thus today the bimodal group with electrically aided hearing in one ear and acoustically-aided hearing in the opposite ear represents the largest group of CI users (Holder et al., 2018). Beyond the fact that the better ear may change depending on the position of target and noise sources, and that bimodal fitting allows use of the ear that is best in any given situation, bimodal listening is expected to provide additional binaural benefits.

Binaural benefits are especially noticeable in challenging acoustic conditions, for instance when speech recognition is impeded by the presence of background noise. Normal hearing listeners (NH) are known to benefit from several binaural effects which have been well quantified in audiometric tests. These include: head shadow (HS), binaural summation (SU), binaural squelch (SQ), and spatial release from masking (SRM) (Bronkhorst and Plomp, 1988, 1989, 1990). Behavioral studies have investigated the amount of binaural benefit that exists in bimodal listeners, but results appear to be controversial (Schafer et al., 2011) and outcomes even include binaural interference, or worsening in comparison to hearing with the CI alone (Illg et al., 2014; Reiss et al., 2016). This may apply all the more so, since many CI listeners use hearing aids (HA) that are unsynchronized with and sometimes fitted independently of the CI. Thus, it is uncertain whether bimodal listeners benefit from a contralateral HA and which factors, either patient-based or provision-based, promote these benefits. In the current study, the CI was seen as the major channel for speech recognition, and we intended to explore whether addition of a HA posed a benefit. Therefore, all binaural benefits were calculated relative to monaural listening with the CI.

Audiometric binaural benefits have been investigated extensively in NH (Bronkhorst and Plomp, 1988, 1989, 1990), whereas objective measures are less well established, but should show as a difference in brain activity for conditions where a binaural effect on speech recognition is known to exist. Multichannel electrical recording (EEG) of auditory evoked potentials (AEP) can capture brain activity non-invasively. This method is compatible with CI use and time-sensitive enough to follow the rapid processing of speech signals (Balkenhol et al., 2020). Hence, responses evoked by monaural stimulus presentation can be directly compared to binaural presentation. Furthermore, comparing AEP traces and behavioral binaural effects for NH listeners, and potential discrepancies for bimodal listeners, may shed light on similarities as well as differences and on their behavioral relevance.

First aim of the current study was to describe AEP traces collected during monaural electrical and bimodal listening and to explore potential differences. In this context, AEP derived from NH listeners served as a template with which to compare the brain's response in bimodal listeners. Some studies have investigated the effects of monaural vs. binaural presentation of auditory stimuli for NH (Henkin et al., 2015; Papesh et al., 2015), and one group performed initial studies on pure tone reception for bimodal listeners (Sasaki et al., 2009). In the current study monosyllable words and their time-reversed acoustic traces were presented monaurally and binaurally within speech-shaped noise. A spatial signal-to-noise constellation, which is known to be associated with a brain-mediated binaural benefit, but at the same time is practicable with monaural CI listening was used. Speech was delivered from the front (S0) and the noise source faced the HA ear (NHA).

Secondly, a related question was to explore whether brain plasticity in the course of adaptation to bimodal hearing plays a role. Obtaining a binaural benefit in the spatial S0NHA constellation requires combining the information from both ears in the central auditory system (Schafer et al., 2011). Therefore changes in the AEP are expected during acclimatization to bimodal hearing. This should be evidenced by a change in the differences between monaural and binaural responses recorded shortly after switch-on of the CI compared to those recorded after an extended time of CI experience. As variability is large in the CI group, this comparison requires repeated measurements for the same subjects. We previously showed that the obligatory N1 and P2 deflections of the brain's AEP response approximated those of NH listeners within the first months of CI experience for binaural presentations, whereas the later event-related N2 potential did not show this effect (Balkenhol et al., 2020). Here we want to explore, whether differences exist between monaural and binaural responses, and whether these differences change with CI experience in the bimodal listeners.

Third aim of the study was to explore, whether monaural vs. binaural differences in the AEP correlate with binaural benefits evidenced by speech audiometry. If significant correlations exist, they could inform about aspects of the AEP response that may serve as an objective measure for binaural processing in bimodal listeners.

Taken together the present study explores whether bimodal listeners experience the same benefit that NH listeners experience, whether this needs time to develop, and whether potential differences in the AEP between monaural and binaural listening correlate with differences in behavioral performance.

MATERIALS AND METHODS

CI Participants

The study protocol was approved by the Institutional Review Board of the Medical Faculty of Mannheim at Heidelberg University (approval no. 2014-527N-MA). Prior to inclusion, each participant provided written consent for participation in the study, and in accordance with the Declaration of Helsinki. All participants were compensated for their visits.

Other aspects of the influence of CI experience in this group of CI users were described earlier (Servais et al., 2017; Wallhäusser-Franke et al., 2018; Balkenhol et al., 2020). Whereas previous reports focused on tinnitus (Servais et al., 2017), subjective perception of the improvement in auditory abilities

TABLE 1 | Participant characteristics.

	CI group (N_{CI} = 15)	NH group (N_{NH} = 14)
Age Mean ± SD (range) in years	57.67 ± 14.95 (27–78)	57.21 ± 13.69 (24–76)
Sex female/male	12/3	12/2
CI ear left/right	8/7	10/4
Lifetime with hearing impairment Mean ± SD in %	CI ear: 53.72 ± 39.01 HA ear: 24.21 ± 19.01	
HA use at future CI ear yes/no	12/3	

(Wallhäusser-Franke et al., 2018), and the development of bimodal hearing (Balkenhol et al., 2020), the current report focusses on the difference between monaural hearing with the CI and bimodal hearing.

Between 2014 and 2017, study participants were recruited from the patients of the CI Center at the University Medical Center Mannheim. Inclusion criteria comprised first-time unilateral CI provision, a HiRes 90K implant as chosen by the patient, continued HA use for the other ear, aged between 18 and 90 years, and speaking German as mother tongue. All patients who fulfilled these criteria were approached for inclusion. Exclusion criteria were assessed during an initial interview (T1) and included: more than mild cognitive deficit, as assessed by the DemTect Test (Kalbe et al., 2004), and presence of an internal stimulator apart from the CI. The initial interview, study inclusion (T1), and pre-surgery examination (T2) took place on the same day, usually the day before surgery. Patients received a CI on their poorer ear, while HA use was continued on the other ear. The CI was switched on 2–3 weeks following implantation. Post-implantation assessments T3 and T4 were scheduled for 3 and 6 months post-implantation, respectively. At each assessment, study participants went through audiometric tests, filled out standardized questionnaires, and underwent EEG recordings.

Twenty-seven patients with hearing loss for both ears, who planned to undergo unilateral CI provision were screened. One was excluded because of an exclusion criterion, while 26 were included in the study. Reasons for premature termination of the study were implantation of the contralateral ear (2 subjects), presence of an exclusion criterion that had not been disclosed at inclusion (1 subject), too much effort (1 subject), or reasons were not disclosed (2 subjects). Data of another 5 participants were excluded because of left-handedness (1 subject), missing AEP data (1 subject), or because of significant worsening of the HA ear during the study (3 subjects). This resulted in 15 participants who contributed data toward the AEP analysis. For demographic details of this group see **Table 1**. All study participants were right-handed native German speakers and used the NAIDA Q70 speech processor. Prior to implantation, 80% used a HA on both ears (**Table 1**), whereas post-implantation all non-implanted ears were aided by auditory amplification.

History of Hearing Loss

At inclusion, all CI participants could communicate verbally when using their HA. Six participants reported hearing problems since early childhood, while 9 had post-lingual onset of profound hearing impairment. On average, severe hearing impairment of the CI ear existed for half of the participants' lifetime, while the HA ear had a shorter duration of hearing loss (**Table 1**). Etiology was unknown for 73%, was due to sudden hearing loss in 2 cases, and one case each of Meniere's disease and Stickler Syndrome.

Normal Hearing Control Group

For each participant who completed the AEP measurement, a right-handed, age-, and sex-matched control with normal hearing was recruited. Control participants (NH) were recruited by word of mouth and from the employees of the University Medical Center Mannheim. Inclusion criteria were: German as native language, no past or present neurological, psychological or hearing problems and right handedness. NH underwent the same screening and undertook the same tests as the CI group. Data from one NH participant was not included because of poor AEP recording. Demographics for the 14 NH are presented in **Table 1**. Average hearing thresholds between 0.25 and 10 kHz for both ears of the 14 NH controls were 17.93 ± 10.32 dB.

Setup for Speech Audiometry and EEG Recordings

Experimental setup is described in detail in Balkenhol et al. (2020). Speech comprehension tests and EEG recordings were performed in a dimly lit sound booth shielded against electromagnetic interference (IAC Acoustics, North Aurora, IL, United States). Participants sat in a comfortable armchair and were observed via glass window and camera.

Speech stimuli were presented in soundfield from a loudspeaker (M-Audio Fast Track Ultra USB Audio Interface and a BX5 near field monitor loudspeaker by inMusic Brand, Cumberland, RI, United States) 1 meter in front of the participant (0° azimuth: S0). Noise came from the same loudspeaker or from one of the two loudspeakers (same brand) at ±90° azimuth. Sound pressure level was always calibrated before testing and with ±0.5 dB accuracy (Brüel & Kjær 2250 sound level meter, Naerum, Denmark) (Letowski and Champlin, 2014).

Speech Audiometry

Speech audiometry was performed as in Balkenhol et al. (2020). An overview on tests and listening conditions is given in **Table 2**. Tests were performed for the following monaural and binaural listening conditions: CI alone (monCI), HA alone (monHA), CI and HA in combination (binaural/bimodal). For all monCI, the HA was removed and the ear was masked with white noise at 65 dB SPL through an insert earphone (AKG K350; Harman International, Stamford, CT, United States; earplug: Grason-Stadler Inc., Eden Prairie, MN, United States). For monaural listening with the HA, the CI was removed. In NH the contralateral ear was masked in both monaural listening conditions in the same way as for monCI.

In all tests, speech was presented from the front (S0) by a male talker. Speech recognition in quiet was tested with the standard clinical German monosyllable test at 70 dB SPL (Freiburger Monosyllable Test or FBE: Hahlbrock, 1970; Löhler et al., 2014),

TABLE 2 | Experimental conditions.

	Test condition	Spatial arrangement	Test	Listening condition			HA ear muted for monCI
				T2	T3	T4	
	Quiet	S0	FBE	binaural (bin)	monCI, monHA, bimodal (bin)	monCI, monHA, bimodal (bin)	With white noise of 65 dB
		S0	OlSa				
Speech audiometry	OlSa noise	S0N0 S0NCI S0NHA	OlSa OlSa OlSa	binaural (bin)	monCI, monHA, bimodal (bin)	monCI, monHA, bimodal (bin)	With white noise of 65 dB
EEG	OlSa noise	S0NHA	30% monosyllable words, 70% time-reversed sound trace of monosyllable words	monCI (future CI ear), binaural (bin)	monCI, bimodal (bin)	monCI, bimodal (bin)	No

and the adaptive version of a sentence test (Oldenburg matrix sentence test or OlSa: Wagener et al., 1999a,b,c). Speech recognition in speech-modulated noise (OlSa noise) was tested with the OlSa with noise delivered from the front (N0), from the speaker facing the CI (NCI), or the HA (NHA). While noise was constant at 60 dB SPL, speech level was changed adaptively starting from +10 dB SNR (signal to noise ratio). Listeners verbally repeated the word (FBE), or each word in a sentence (OlSa) as understood, and the experimenter entered the correct words. No feedback was given, lists were not repeated within sessions. FBE results comprised two lists of 20 words per listening condition with higher percentage indicating better speech recognition. For each test condition, twenty OlSa sentences were presented with the average calculated from the last ten sentences for 50% speech recognition in quiet (dB SRT) or the SNR needed for 50% correct comprehension in noise (dB SNR). Sequence of tests and lists was constant between participants and assessments but listening conditions were varied at random. While in the FBE higher values indicate better speech recognition, lower values in the OlSa are indicative of better speech recognition. Because monaural speech tests were not possible before implantation monaural vs. binaural comparisons are available only for the post-implantation assessments T3 and T4, and for NH.

Ear Dominance and Bimodal Benefit

The better ear, post-operatively, was determined for each speech test by subtracting the values obtained with monHA from the respective values with monCI. For a difference of more than 10% in the FBE, or 3 dB in the OlSa tests, aided hearing was defined as asymmetric and the better ear was determined. The 10% boundary for the FBE was chosen according to Müller-Deile (2009), the 3 dB boundaries for OlSa tests were derived from work by Litovsky et al. (2006).

Binaural benefits were calculated from OlSa tests as head shadow (HS), binaural summation (SU), binaural squelch (SQ), and spatial release from masking (SRM). All benefits were calculated relative to monaural listening with the CI ear. Calculations were carried out in such a way that binaural benefits will produce a positive value while binaural interference, i.e., worsening in the binaural condition, has a negative leading sign. Because lower values represent better speech recognition in OlSa tests, calculations derived from OlSa results were inverted. Calculations for NH were performed alike for monaural vs. binaural listening.

The binaural benefits HS, SU, and SQ were calculated from OlSa results according to Schleich et al. (2004). HS was calculated as follows:

$$HS_{monCI} = S0NCI_{monCI} - S0NHA_{monCI}. \quad (1)$$

Binaural loudness summation (SU) was calculated for speech presented in quiet (SU_Q) and for speech presented with noise from the same source (SU_N) in the following way:

$$SU_Q = S0_{bin} - S0_{monCI}, \quad (2)$$
$$SU_N = S0N0_{bin} - S0N0_{monCI}. \quad (3)$$

Binaural SQ was calculated for the condition with lateral noise contralateral to the monaurally active ear (Schleich et al., 2004), here the CI ear:

$$SQ = S0NHA_{bin} - S0NHA_{monCI}. \quad (4)$$

This spatial signal to noise constellation is the same as the one used during EEG recordings (see section "EEG Recordings"). A measure of SRM was derived by subtracting speech recognition within lateral noise (S0NHA) from the condition of collocated speech and noise (S0N0) for monaural listening with the CI ear:

$$SRM_{monCI} = S0N0_{monCI} - S0NHA_{monCI} \quad (5)$$

and for binaural listening:

$$SRM_{bin} = S0N0_{bin} - S0NHA_{bin}. \quad (6)$$

Normal distributions of auditory outcomes were checked with the Shapiro-Wilk test (Shapiro and Wilk, 1965) and by inspection of outcome distributions. Monaural vs. binaural comparisons for speech tests in quiet (FBE, OlSa S0) were tested for significance by planned comparisons with parametric (*t* values) or non-parametric tests (*z* values) depending on normality. Statistical significance of differences for speech recognition in noise were determined for T3 and T4 assessments, and for NH with 3 spatial conditions (S0N0, S0NCI, S0NHA) × 2 listening conditions (monaural, binaural) with repeated-measures ANOVAs by MATLAB's Statistics and Machine Learning Toolbox (R2018a)

(Mathworks, Natick, MA, United States). Because of small sample size, Greenhouse-Geisser adjustment was used to correct against violations of sphericity. Given that a significant main effect existed, *post hoc* two-tailed paired samples *t* tests were performed and corrected for multiple comparisons according to Tukey-Kramer. Whether bimodal HS, SU, SQ, and SRM effects differed significantly from zero was determined with one-sample *t* tests. To correct for multiple testing, Bonferroni-corrected significance limens equivalent to the *p* value that indicates a trend ($^{+}p < 0.1$), a significant difference ($^{*}p < 0.05$), or a highly significant difference ($^{**}p < 0.01$) are given together with the uncorrected *p* value. Differences in HS, SU, SQ, and SRM between T3, T4, to NH were tested for significance with Dunnett's multiple comparison test (Dunnett, 1955; Dunlap et al., 1981). Group means (Mean) together with their standard deviations (SD) are used throughout the text if not indicated otherwise. Correlation analyses for audiometric measures were performed with SPSS25 (SPSS/IBM, Chicago, IL, United States).

EEG Recordings

As described in Balkenhol et al. (2020) EEG was continuously recorded from 62 active Ag/AgCl surface electrodes arranged in an elastic cap (g.LADYbird/g.GAMMAcap; g.tec Medical Engineering GmbH, Austria) according to the 10/10 system (Oostenveld and Praamstra, 2011), Fpz served as ground. Two active Ag/AgCl electrodes (g.GAMMAearclip; g.tec) were clipped to the earlobes. The electrooculogram (EOG) was monitored with 4 passive Ag/AgCl electrodes (Natus Europe GmbH, Germany) placed below and at the outer canthi of the eyes. Electrodes located above or close to CI or HA were not filled with gel [Mean ± SD (range): CI: 3 ± 1.1 (1–5); HA: 1 ± 0.5 (0–2)] and were interpolated during post-processing. Impedances were below 5 kOhm for passive electrodes, and below 30 kOhm for active electrodes. Sampling frequency was 512 Hz with 24-bit resolution (biosignal amplifier: g.HIamp; g.tec). Data acquisition and playback of the stimuli were controlled by MATLAB/Simulink R2010a (Mathworks, Natick, MA, United States) with custom MATLAB scripts. Real-time access to the soundcard was realized with the playrec toolbox[1]. A trigger box (g.TRIGbox; g.tec) was used to mark stimulus onsets and offsets and to record push button activity (see section "Task and Procedure").

Stimuli

Stimuli were German monosyllables from the FBE spoken by a male talker (Hahlbrock, 1970). Reversals were generated by time-reversing the audio tracks of these monosyllables. Only reversals that did not resemble a German word as judged by the lab members were used. In total, 269 words and 216 reversals, with a mean duration of 770 ± 98 ms (484–1,035 ms) were used. Lists were generated randomly from the complete set with 75 stimuli in a stimulation block. 30% of these stimuli were words and 70% were reversals. Lists were not repeated during an assessment. During all stimulation blocks OlSa noise at 60 dB SPL (Wagener et al., 1999a,b,c) was delivered toward the HA ear or the ear that was not active in the monaural condition in NH controls (azimuth ±90°: NHA).

[1]http://www.playrec.co.uk

Task and Procedure

Participants were instructed to face the loudspeaker in front where the signals originated (S0), to close their eyes, and not to move during recording. Their task was to respond to the infrequent words by pressing a button after hearing a signal sound (white noise, 75 dB SPL, 50 ms) that followed 1,000 ms after offset of each word and reversal. The button press served both to maintain alertness and to calculate the percentage of words identified within a stimulation block which was used to calculate binaural squelch (SQ_{AEP}) for this condition (for calculation see section "Bimodal Benefits"). Inter-stimulus intervals between the end of the signal sound and the start of the next stimulus were 1,900 ± 200 ms resulting in 75 stimuli per 5 minutes presentation block. During the entire presentation block continuous OlSa noise was played from the loudspeaker facing the HA ear (NHA). Each block was followed by a break during which participants could relax. All participants received the same randomized stimulus sequence within each block, whereas the sequence of monaural and binaural listening conditions varied. Overall, 297 ± 55 responses were recorded for monaural and 303 ± 61 for binaural listening conditions.

To avoid ceiling and floor effects, signal to noise ratio was individually set to achieve 70% correct detection of words (dB SNR) as ascertained in practice runs prior to recording. If rates deviated substantially from this criterion, the procedure was repeated with an adjusted presentation level. If button press occurred before the signal sound, that AEP was excluded from analysis. At T4, two familiarization blocks were performed using the same SNR as at T3.

EEG Pre-processing

EEG data were pre-processed offline with MATLAB R2018a (Mathworks, Natick, MA, United States) with the EEGLAB toolbox (version 13.3.2b) (Delorme and Makeig, 2004), and custom MATLAB scripts as described in Balkenhol et al. (2020). Raw data were: (1) re-referenced to linked earlobes, (2) low-pass filtered with 64 Hz cut-off and (3) high-pass filtered with 0.5 Hz cut-off using finite impulse response (FIR) filters, and (4) segmented into epochs from −300 to 2,200 ms relative to stimulus onset. Epochs with amplitudes exceeding ±150 μV in single channels or with non-stereotyped artifacts, classified by kurtosis and joint probability (threshold: 3 SD), were highlighted during visual inspection. Final rejection of epochs and the identification of poor electrode channels [CI group Mean ± SD (range): 0.8 ± 1.5 (0–7); NH group Mean ± SD (range): 0.9 ± 1.1 (0–3)] were performed by experienced lab members.

Next, EOG artifacts were removed automatically using a second-order blind identification (SOBI) and independent component analysis (ICA) (Molgedey and Schuster, 1994; Onton et al., 2006; Delorme et al., 2007), as described in Balkenhol et al. (2013).

The CI induced narrow- and wide-band EEG components above 25 Hz in response to words and reversals. These were removed with SOBI ICA using an automated artifact removal algorithm developed for this study which identifies artifacts in the independent components based on power distribution. While narrow-band artifacts were automatically detected by a spectral peak search algorithm, wide-band artifacts were identified by their average power in the high frequencies (40–256 Hz), relative to power in low frequencies (3–25 Hz). Components were removed if spectral power in the high-frequency interval exceeded power in the low-frequency interval (Balkenhol et al., 2020).

Then, muscle artifacts, heartbeat activity, and other sources of non-cerebral activity were visually identified on independent component scalp maps and their power spectra (Luck, 2014), and removed by back-projecting all but these components. Finally, unfilled and channels of poor quality were interpolated by spherical splines. On average, 14% of the AEP were removed while 256 ± 54 responses remained per participant and assessment.

EEG Data Analysis

Amplitudes and latencies were computed for the N1, P2, and N2 deflections for monaural and binaural listening conditions and for each stimulus category. Binaural-monaural differences were calculated.

As described in Balkenhol et al. (2020) data analysis was performed in MATLAB R2018a (Mathworks, Natick, MA, United States) with the fieldtrip toolbox (version 20170925[2]; Oostenveld et al., 2011) and custom MATLAB scripts. Computations are based on subject averages across all 62 electrodes, and separately for the categories "words" (all responses to word stimuli) and "reversals" (all responses to reversed stimuli). For baseline correction, the pre-stimulus mean from −150 to −50 ms was subtracted from each epoch. Differences in intensity rise times between stimuli were corrected by delaying the onset trigger to the first time point when a stimulus reached 50% of its maximal amplitude. Amplitudes were calculated for the time intervals from 80–180 ms (N1), 180–330 ms (P2), and 370–570 ms (N2) (Luck, 2014). N1, P2, and N2 latencies were quantified by the 50% area latency measure according to Liesefeld (2018) and as described in Balkenhol et al. (2020). In short, peak-to-peak amplitude distance to the preceding peak was determined, the baseline which divided amplitudes in half was identified, and the time point that splits this area in half was calculated.

Statistical analysis was performed with MATLAB's Statistics and Machine Learning Toolbox (R2018a) and custom scripts. Depending on distribution of the data, parametric or non-parametric tests were used. Amplitudes and area latencies for N1, P2, and N2 responses corresponding to "words" and "reversals" were subjected to separate Dunnett's multiple comparison procedures to compare CI group results at T2, T3, and T4 with the NH group for monCI and bimodal listening conditions (Dunnett, 1955; Dunlap et al., 1981). For comparisons with significant main effects, *post hoc t* or Wilcoxon tests were performed. A value of $p < 0.05$ was considered to be statistically significant, while $p < 0.1$ indicated a trend.

[2]http://www.ru.nl/fcdonders/fieldtrip

Source Localization

Source localization analysis for the N1 interval was performed with MATLAB's fieldtrip toolbox and time-domain based eLORETA (Pascual-Marqui, 2007, 2009) using the "colin27" head model (Holmes et al., 1998). Monte-Carlo estimates of probability were derived by non-parametric randomization tests (N_r = 1,000, two-sided). Leadfield resolution was 5 mm, statistical analysis was performed on dipole power, and a false discovery rate (FDR) was used to correct for multiple comparisons. A detailed description of this procedure is given in Balkenhol et al. (2020).

RESULTS FOR SPEECH AUDIOMETRY

For ease of reporting and interpretation, differences calculated from OlSa tests have been inverted such that binaural benefits will be reported with positive numbers while binaural interference is indicated through negative numbers. This is despite the fact that a lower score represents better performance for the OlSa tests.

Development of Speech Recognition in the CI-Aided Ear

Average scores are presented in **Figure 1** for speech recognition in quiet (FBE, OlSa S0), and with background noise from different directions (OlSa: S0N0, S0NCI, S0NHA). For the CI group, monaural speech comprehension tests were not performed pre-implantation due to the inability of many of the participants to complete these tests. In addition, one participant was not able to complete some OlSa tests with monaural CI-aided listening at T3 and T4.

Post-implantation, paired *t* tests (*t* values) and Wilcoxon tests in case of non-normality (*z* values) showed significant improvements of speech recognition with the CI ear between T3 and T4 for the OlSa S0 ($z = 3.296$, $^{**}p < 0.001$), S0N0 ($t = 3.300$, $^{**}p < 0.001$), and S0NHA ($z = 2.638$, $^{*}p < 0.009$) conditions when applying the Bonferroni-corrected significance limen for significant * ($p < 0.01$) and highly significant ** ($p < 0.002$) differences (**Figure 1**).

Ear Dominance

At the onset of the study, the CI ear was expected to become the better ear post-implantation. The study population included individuals with substantial amounts of aidable hearing on the HA side, however. So, at T4, hearing abilities were equally distributed across ears, with about one third each of the participants falling into the "symmetric", "better CI ear", or "better HA ear" categories according to average performance on all speech perception tests (**Table 3**). In contrast, at T3, the HA ear was the better ear for more than half of the participants, while about 30% had symmetric speech recognition, and the CI ear was the better ear for 16%. This distribution differed considerably between test conditions as can be seen in **Table 3**.

For comparison, about 76% of the NH listeners showed symmetric performance. Symmetry was not perfect, however, and mainly pertained to FBE and OlSa S0N0 (**Table 3**). As a right ear advantage has been reported for speech perception (Westerhausen et al., 2015), data of NH were additionally screened for right vs. left ear comparisons. Since there were no

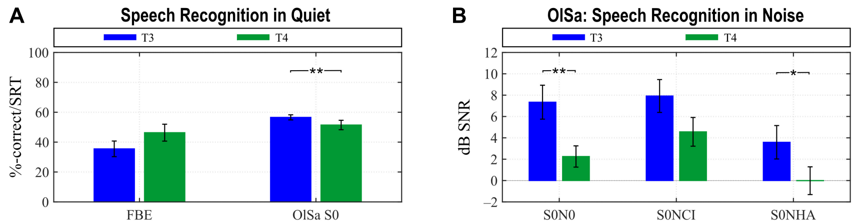

FIGURE 1 | Improvement of hearing with the CI ear between T3 and T4. Significant improvements are seen for various test constellations in quiet **(A)** and background noise **(B)**. In the FBE higher values (%-correct) indicate better performance, whereas in all OlSa tests, lower values indicate better performance. Group means with their standard errors are shown (**$p < 0.002$, *$p < 0.01$).

TABLE 3 | Better ear during speech recognition.

Speech recognition test	T3			T4			NH		
	Symmetric $N_{S,T3}$ (%)	CI ear better $N_{CI,T3}$ (%)	HA ear better $N_{HA,T3}$ (%)	Symmetric $N_{S,T4}$ (%)	CI ear better $N_{CI,T4}$ (%)	HA ear better $N_{HA,T4}$ (%)	Symmetric $N_{S,NH}$ (%)	Designated CI ear better $N_{CI,NH}$ (%)	Designated HA ear better $N_{HA,NH}$ (%)
FBE	5 (33.3)	2 (13.3)	8 (53.3)	2 (13.3)	5 (33.3)	8 (53.3)	14 (100)	0	0
OlSa S0	3 (20.0)	4 (26.7)	8 (53.3)	1 (6.7)	7 (46.7)	7 (46.7)	7 (50)	4 (28.6)	3 (21.4)
OlSa S0N0	6 (40.0)	2 (13.3)	7 (46.7)	8 (53.3)	5 (33.3)	2 (13.3)	14 (100)	0	0
OlSa $S0N_{ipsi}$	5 (33.3)	2 (13.3)	8 (53.3)	6 (40.0)	4 (26.7)	5 (33.3)	9 (64.3)	1 (7.1)	4 (28.6)
OlSa $S0N_{contra}$	4 (26.7)	2 (13.3)	9 (60.0)	8 (53.3)	3 (20.0)	4 (26.7)	9 (64.3)	2 (14.3)	3 (21.4)
Mean	4.6 (30.7)	2.4 (16.0)	8 (53.3)	5 (33.3)	4.8 (32.0)	5.2 (34.7)	10.6 (75.7)	1.4 (10.0)	2 (14.3)

significant differences for speech recognition achieved with either ear for any of the conditions tested in the current NH group, this issue was not pursued further.

The S0NHA constellation in the OlSa test was closest to the spatial distribution of speech and noise sources during EEG recordings. For this condition audiometric outcomes imply addition of a better HA ear in the bimodal listening condition for 60% of CI participants at T3, and addition of an equal ear for 53.3% at T4 which was closer to the situation in NH where symmetric speech recognition was found for 64.3% (**Table 3**).

Monaural vs. Binaural Comparisons

NH Group

NH listeners gained the largest binaural benefits, therefore results from this group are presented first, and results of the CI group are compared to them. Many NH performed the FBE with a ceiling effect (monCI: 95.5 ± 5.9%; binaural: 98.9 ± 1.6%), and the difference between listening conditions (z = –2.203, p = 0.028) failed the Bonferroni-corrected significance limen of $p < 0.025$. In contrast, recognition in the OlSa S0 test was significantly better for sentences presented binaurally (t = 4.806, $p < 0.0004$) (**Figure 2A**). A 3 × 2 (noise direction: S0N0, S0NCI, S0NHA × listening condition: monaural, binaural) repeated-measures ANOVA revealed a significant main effect for noise direction [$F(2,26) = 61.258$, $p < 3{\cdot}10^{-10}$] and listening condition [$F(1,13) = 64.599$, $p < 3{\cdot}10^{-6}$], with a significant interaction between these factors [$F(2,26) = 24.916$, $p < 2{\cdot}10^{-5}$]. *Post hoc* tests focused on listening condition and revealed significantly better results for binaural listening when noise was presented at the side of the monaurally active ear (S0NCI: t = 8.630, $p < 2{\cdot}10^{-5}$). The difference between monaural and binaural presentation remained insignificant for noise from the same source (S0N0: t = 1.925, p = 0.431), or for noise presented from the side of the monaurally inactive ear (S0NHA: t = 1.633, p = 0.593) (**Figure 2B**).

CI Group

At T3, the bimodal condition most often equaled addition of an equally or better performing HA ear (**Table 3**), and significant improvements for bimodal listening compared to monCI were evidenced for all speech comprehension tests. For tests in quiet, planned comparisons evidenced significant effects for FBE (t = –4.085, $p < 0.002$) and OlSa S0 (z = 3.296, $p < 0.001$), when applying the Bonferroni-corrected significance limen of $p < 0.025$ (**Figure 2C**). A 3 × 2 repeated-measures ANOVA for hearing in noise revealed a significant main effect for noise direction [$F(2,26) = 69.560$, $p < 0.009$] and listening condition [$F(1,13) = 14.538$, $p < 0.003$], as well as a significant interaction [$F(2,26) = 4.421$, $p < 0.041$]. *Post hoc* tests for monCI vs. bimodal hearing confirmed significant improvements with bimodal hearing for S0N0 (t = 3.385, $p < 0.045$) and S0NCI (t = 3.556, $p < 0.033$), whereas a trend was observed for S0NHA (t = 3.129, p = 0.069) (**Figure 2D**).

By the time of the T4 sessions, the distribution of performance between ears had changed (**Table 3**), and significant improvements between bimodal and monaural electric hearing existed, although not for all conditions (**Figures 2E,F**). In quiet, speech perception improved for bimodal hearing for OlSa S0

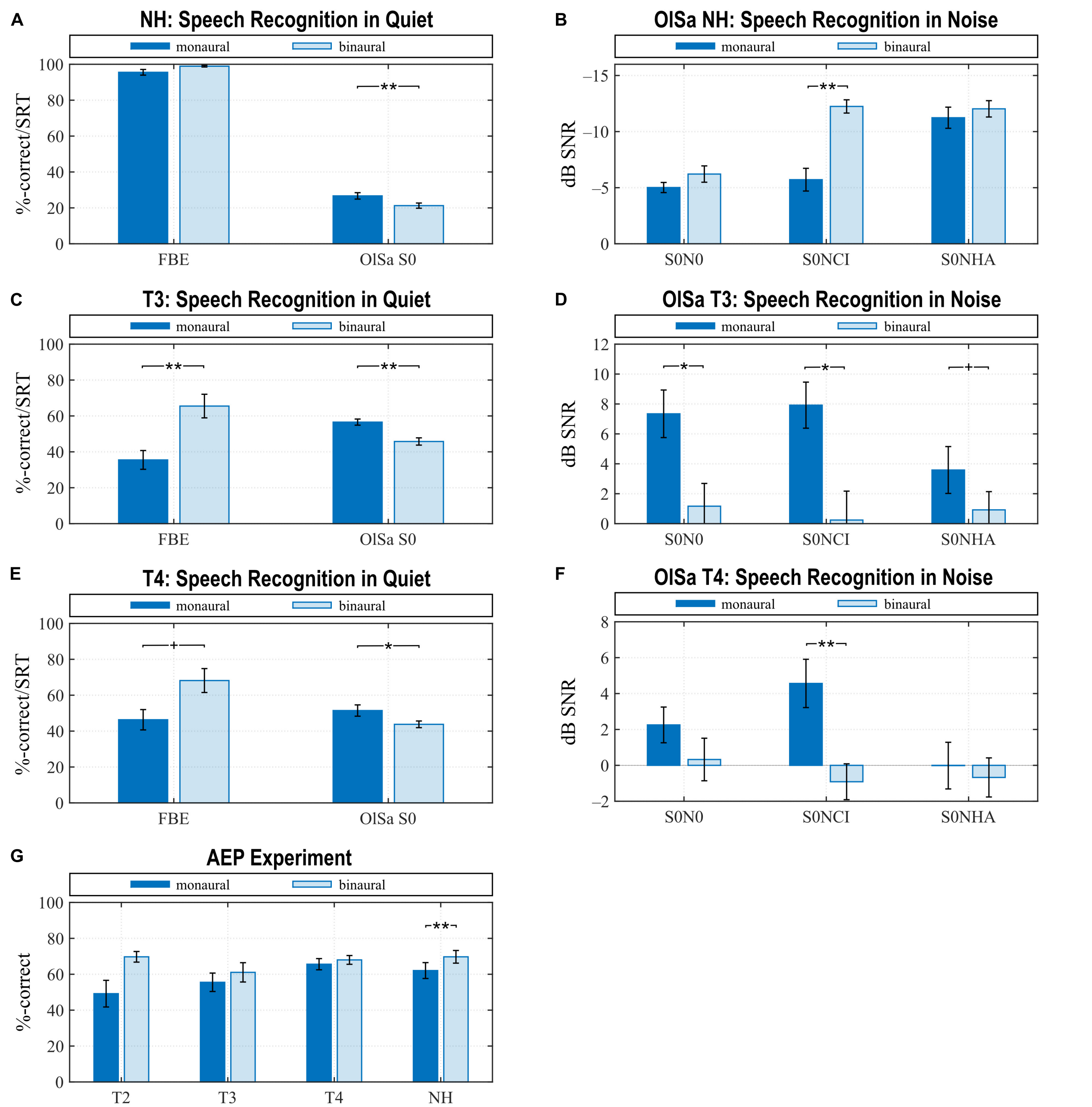

FIGURE 2 | Speech perception with FBE and OlSa tests **(A–F)** and assessed in the AEP experiment **(G)**. Higher values signal better speech recognition in the FBE and AEP conditions, whereas lower values indicate better speech recognition in OlSa tests. Note the reversed vertical scale in **(B)**. Perception is best in NH listeners **(A,B,G)**, worst shortly after CI provision **(C,D,G)** and improves with CI experience **(E–G)**. Statistically significant differences between monaural listening with the CI or the designated CI ear in NH and binaural speech recognition was observed for several test conditions. Due to insufficient monaural hearing, monaural data at T2 are only available for the AEP condition, where a significant difference existed between monaural and binaural listening. While behavioral results from the AEP experiment at T3 and T4 did not evidence a significant difference between listening conditions, behavioral tests showed significantly better bimodal speech recognition at T3 for all test conditions and at T4 for the S0NCI condition. Regarding the spatial arrangement of speech and noise sources, the AEP condition is closest to S0NHA. Means and their standard error are shown (**$p < 0.01$, *$p < 0.05$, trends $^{+}p < 0.1$).

($z = 2.727$, $p < 0.007$) but not for the FBE ($z = -2.047$, $p = 0.041$) when applying the Bonferroni-corrected significance limen of $p < 0.025$ (**Figure 2E**). Significant main effects for noise direction [$F(2,26) = 5.999$, $p < 0.017$] and listening condition [$F(1,13) = 7.877$, $p < 0.015$] were derived by a 3 × 2 repeated-measures ANOVA. The interaction effect was also statistically significant [$F(2,26) = 10.374$, $p < 0.002$]. *Post hoc* tests for monCI vs. bimodal showed significant improvements in the bimodal condition when noise was presented from the CI side (S0NCI: $t = 5.445$, $p < 0.002$), but no difference existed for S0N0 (t = 2.362, p = 0.282) or S0NHA (t = 0.669, p = 0.983) (**Figure 2F**).

Binaural Benefits

Binaural benefits were calculated for the addition of the HA ear relative to monaural listening with the CI as head shadow (HS), binaural summation in quiet (SU_Q) and noise (SU_N), as binaural squelch (SQ, SQ_{AEP}), and spatial release from masking (SRM) for monaural (SRM_{monCI}) and binaural listening (SRM_{bin}) (**Table 4**).

For NH, significant HS and SRM effects were estimated (HS: t = 7.670, $p < 4 \cdot 10^{-6}$; SRM_{monCI}: t = 10.567, $p < 10^{-7}$; SRM_{bin}: z = 3.170, $p < 0.002$). SRM_{monCI} and SRM_{bin} did not differ ($z = -0.699$, p = 0.485). SU_Q was higher than SU_N (z = 3.107, $p < 0.002$), and while SU_Q was significantly different from zero (z = 3.323, $p < 0.002$), SU_N failed significance after Bonferroni correction (z = 2.198, p = 0.028). Also, SQ calculated from OlSa S0NHA remained insignificant (t = 1.633, p = 0.126), whereas SQ_{AEP} derived from the button-press response during EEG recordings attained significance (t = 4.935, $p < 0.0003$). The corrected significance limen for all tests against zero was $p < 0.0071$.

In the CI group, a significant HS of similar magnitude as in NH was present at both post-CI assessments (Dunnett's test: F = 0.413, p = 0.665; T3 vs. NH: p = 0.600; T4 vs. NH: p = 0.719), and one-sample t tests against zero with a Bonferroni-corrected significance limen of $p < 0.0071$ evidenced its significance at T3 (t = 3.968, $p < 0.002$) and T4 (t = 4.290, $p < 0.001$).

A difference between NH and bimodal listeners existed regarding SRM_{monCI} and SRM_{bin} with CI listeners benefiting significantly less (Dunnett's test: SRM_{monCI}: F = 4.649, $p < 0.016$; SRM_{bin}: F = 9.235, $p < 0.0005$). Whereas SRM_{monCI} was significantly different from zero at T3 (t = 3.327, $p < 0.006$), significance did not survive Bonferroni correction at T4 (t = 2.793, p = 0.015), and SRM_{bin} was far from reaching significance at both assessments (T3: t = 0.241, p = 0.813; T4: t = 0.973, p = 0.347). The large reduction between SRM_{monCI} and SRM_{bin} for the bimodal listeners, especially at T3 (**Table 4**) did not attain significance when tested against NH where SRM_{monCI} and SRM_{bin} were similar (Dunnett's test: F = 1.430, p = 0.251).

Bimodal listeners benefited significantly from binaural summation at T3 (SU_Q: t = 4.098, $p < 0.002$; SU_N: t = 3.385, $p < 0.005$) but less so at T4 with SU_N losing significance when corrected for multiple comparisons (SU_Q: z = 2.723, $p < 0.007$; SU_N: t = 2.362, p = 0.033). Dunnett's test comparing T3 and T4 assessments with NH yielded a significant main effect for SU_N (F = 4.971, $p < 0.012$; T3 vs. NH: $p < 0.012$; T4 vs. NH: p = 0.872), but not for SU_Q (F = 0.819, p = 0.448). Similar to NH, at T4, SU dropped considerably between quiet and noise (T4: z = 2.101, $p < 0.036$; NH: z = 3.107, $p < 0.002$).

As in NH, the benefit derived from binaural SQ in the OlSa test did not attain the Bonferroni-corrected significance limen of $p < 0.0071$ at either assessment (T3: t = 3.129, $p < 0.008$; T4: t = 0.669, p = 0.516). Therefore, no further calculations were performed with this measure. In contrast, SQ_{AEP}, which attained significance in NH, was not different from zero pre- (T2: z = 2.273, p = 0.024) or post-implantation (T3: t = 1.229, p = 0.239; T4: t = 0.921; p = 0.373). Statistical comparisons between CI assessments and NH evidenced a significant main effect (Dunnett's test: F = 2.792, $p < 0.05$), while *post hoc* comparisons showed no significant differences to NH which may be a consequence of differences in SQ_{AEP} between assessments and heterogeneity of the bimodal listeners (**Table 4**).

Taken together, monaural intelligibility with the CI ear improved with CI experience, while evidence for binaural processing was limited to a positive SU_Q given that HS is essentially a monaural effect.

Correlations Between Audiometric Measures

Correlations were tested on an exploratory basis for NH and the T4 assessment and not corrected for multiple comparisons. Bivariate comparisons between the different binaural/bimodal effects, and between these effects and PTA-4 as well as with PTA-4 asymmetry were calculated. Only significant correlations are reported. Several significant positive as well as negative correlations were found (**Table 5**). Mostly, these were present for either the NH or the CI group, but not for both. The only exception was the significant inverse correlation between SU_Q and PTA-4 asymmetry, which in CI listeners represented the aided PTA-4 of the CI and HA ears. For both groups, the correlation had the same direction and magnitude. This indicated that, on one hand addition of a better HA ear produces a larger SU_Q, and that on the other hand, lower asymmetry between ears is associated with a larger SU_Q if the CI ear is an equal or the better ear.

RESULTS FOR EEG RECORDINGS

AEP: Monaural vs. Binaural and Words vs. Reversals Comparisons

Results from NH listeners are described first and then compared to those of the CI group.

In NH, most conspicuous differences between listening conditions and stimulus categories pertained to N1 with N1 amplitudes differing significantly for all comparisons (**Figures 3**, **4**). Only the monaural to binaural comparison following reversals failed full significance, but revealed a trend. N1 responses toward words were larger, and the most negative peak was seen following word presentation with binaural listening. A significant difference in N1 latency was observed between stimulus categories for monaural presentation with the response occurring earlier after words. In addition, a difference existed in response to reversals with the N1 response occurring significantly earlier with binaural listening. Whereas P2 amplitudes were similar for all conditions, P2 latencies differed between stimulus categories with shorter latencies following words. While this difference became significant for monaural listening, a trend toward significance was seen for the binaural response. N2 amplitude was low for all conditions.

TABLE 4 | Binaural Benefit when adding the HA ear. Shown are Mean ± SD (range).

Group	HS_{monCI} = $S0NCI_{monCI}$ – $S0NHA_{monCI}$ in dB SNR	SRM_{monCI} = $S0N0_{monCI}$ – $S0NHA_{monCI}$ in dB SNR	SRM_{bin} = $S0N0_{bin}$ – $S0NHA_{bin}$ in dB SNR	SU_Q = $S0_{bin}$ – $S0_{monCI}$ in –dB SRT	SU_N = $S0N0_{bin}$ – $S0N0_{monCI}$ in –dB SNR	SQ = $S0NHA_{bin}$ – $S0NHA_{monCI}$ in –dB SNR	SQ_{AEP} = $S0NHA_{bin,AEP}$ – $S0NHA_{monCI,AEP}$ in %-correct
CI T2	–	–	–	–	–	–	22.06 ± 33.37 (−14.1−86.2)
CI T3	4.34 ± 4.09 (−3.9−10.5)	3.76 ± 4.23 (−5.9−10.3)	0.25 ± 3.96 (−5.6−10.9)	10.26 ± 9.36 (0.7−33.4)	6.16 ± 6.81 (−4.1−17.8)	3.06 ± 3.67 (−3.2−9.7)	5.53 ± 17.43 (−18.5−35.9)
CI T4	4.58 ± 3.99 (−5.5−10.3)	2.44 ± 3.26 (−3.0−7.6)	1.00 ± 3.98 (−6.4−9.2)	7.68 ± 13.68 (−1.7−53.3)	1.93 ± 3.16 (−4.0−6.9)	0.96 ± 5.40 (−13.2−9.1)	2.39 ± 10.06 (−15−27)
NH	5.52 ± 2.69 (0−9.3)	6.22 ± 2.20 (1.1−9.3)	5.81 ± 3.27 (−2.9−9.7)	5.43 ± 4.23 (−1.1−16.8)	1.20 ± 2.33 (−1.4−8.4)	0.79 ± 1.82 (−3.0−2.7)	7.66 ± 5.80 (−3.3−19.6)

Grand average of the CI group for N1 and P2 was comparable to NH, but deviated for N2 (**Figures 3**, **4**). Pre-implantation, N1 amplitude did not differ between listening conditions or stimulus categories, while significant differences between listening conditions were observed post-implantation. Following words, N1 was larger with bimodal hearing. This difference extended with bimodal experience and became significant at T4 which paralleled the significant difference observed in NH in direction and magnitude. In contrast, the difference of N1 negativities between listening conditions following reversals peaked at T3 when this difference became statistically significant, and failed significance at T4. In contrast to NH, N1 amplitude did not differ between stimulus categories at either assessment, and N1 latencies to reversals were delayed relative to words in the binaural condition. This difference in delay was highly significant before implantation (T2), attained significance at T3, and reduced to a trend at T4. Thus, differences in the N1 response between listening conditions approximated those seen in NH within 6 months of CI experience, while absence of a difference between stimulus categories did not parallel the situation in NH.

Regarding P2, a significant difference in amplitude between monCI and bimodal listening following reversals was observed at T3, whereas further significant differences pertained to P2 latency at T4. At T4, significant latency differences existed between listening conditions, but in opposite directions for word and reversal stimulus categories. Whereas P2 latency following words was significantly shorter with bimodal listening, latency following reversals was significantly shorter for monCI. In addition, a highly significant difference existed between stimulus categories in the bimodal listening condition. Only the latter had a parallel in NH with a trend toward significance for the latency difference between stimulus categories with binaural listening and a later response to reversals.

While N2 was almost absent in NH, it was a prominent negative deflection in the CI group at all assessments and for all conditions. A trend toward a larger response to words than reversals existed at T2 for binaural listening, while significant differences between listening conditions were observed at T3 for both stimulus categories.

Source Localization of Monaural/Binaural Differences

Time domain eLORETA analyses were computed for the N1 response at T4 and in NH for monaural vs. binaural hearing. As for the majority of study participants the left ear was the ear that was stimulated in the monaural listening condition (**Table 1**), this analysis was performed for the 8 study participants with a CI on their left ear and the 10 NH with left ear monaural stimulation. Significant activation differences between monaural and binaural listening were observed (**Figure 5**), but locations differed between NH and CI groups. **Table 6** lists brain structures with a significant difference between listening conditions in at least 20% of their voxels.

In NH, activation between listening conditions differed significantly in left primary and secondary auditory cortices (Brodmann areas BA41, 42), i.e., ipsilateral to the side of monaural stimulation (**Table 6**). The positive *t* value indicated a more negative N1 with binaural listening. In addition, significantly increased negativity in the binaural listening condition was observed in left insula and postcentral gyrus, whereas negativity in the ventral frontal lobe was smaller in the right hemisphere with binaural hearing, indicated by negative *t* values. Affected areas belonged to inferior frontal gyrus (IFG) and orbital gyrus (OrG).

In contrast, differences between electric and bimodal hearing in the CI group affected auditory association areas in the temporal and parietal lobes that are involved in sensory aspects of speech processing (Ardila et al., 2016). Whereas negativity in left temporal areas (BA21, BA38) was smaller, increased negativity was observed in the parietal lobe with bimodal hearing. Affected areas were BA7 and BA39 in the left hemisphere and BA7 in the right. Furthermore, differential activation was observed in the left insula and cingulate gyrus with smaller negativities in the bimodal listening condition.

Correlations Between AEP and Audiometric Measures

For AEP measures that showed significant differences between listening conditions at T4 and in NH, differences between the binaural minus the monaural condition were calculated and bivariate correlation analyses were performed with these

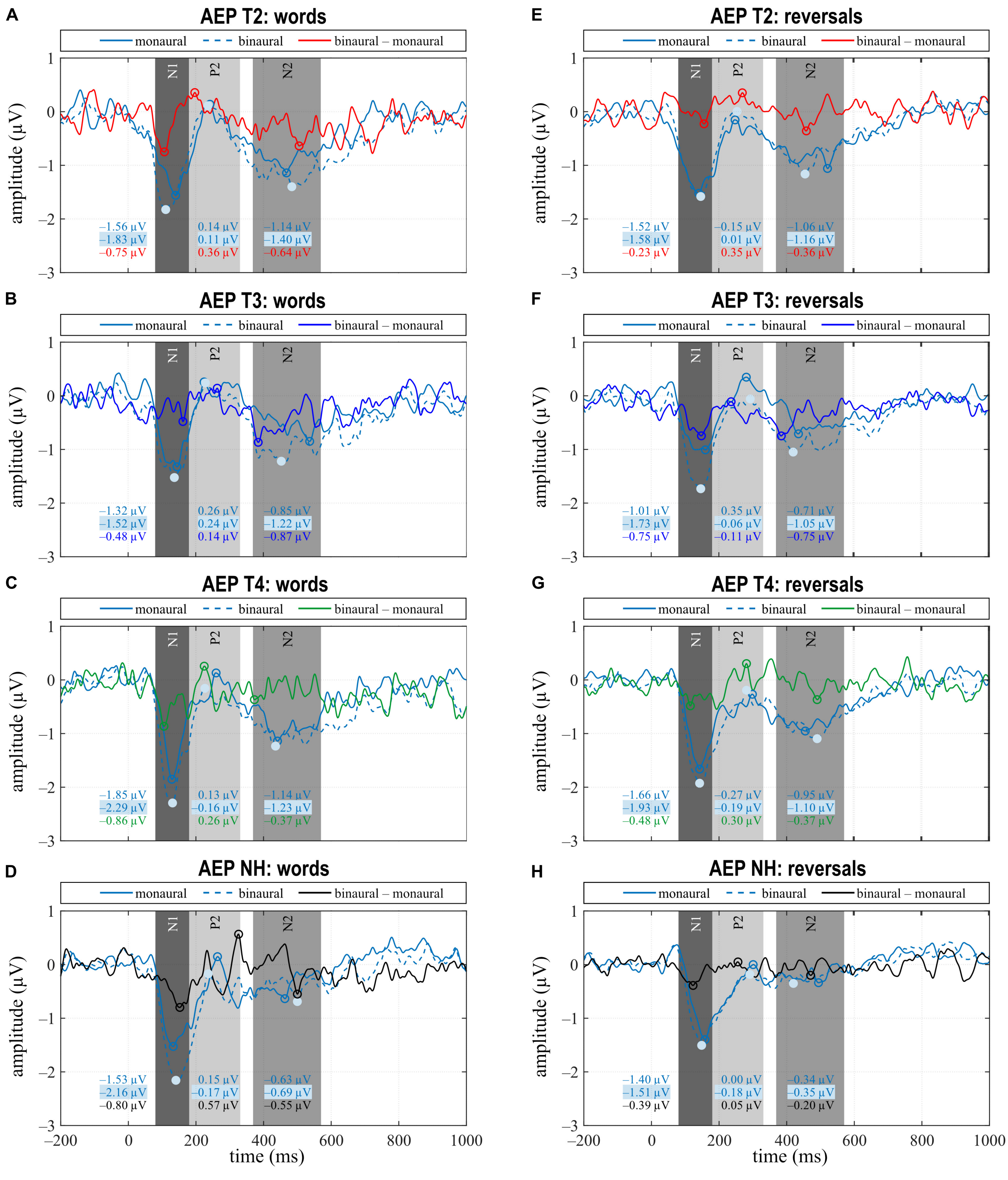

FIGURE 3 | Grand averages for monaural, binaural listening conditions, and the difference binaural – monaural for the categories "words" **(A–D)** and "reversals" **(E–H)** of the CI (T2–T4) and NH group. **(A–H)** Time intervals with N1, P2, and N2 responses are shaded in different grays.

differences and the binaural benefits (**Table 4** and section "Binaural Benefits"), and with PTA-4 asymmetry. For the CI group, these were the differences in N1 amplitude related to words, and P2 latency differences in response to words and reversals. For NH, correlations were computed with the difference in N1 amplitude and N1 latency. Only the differences in P2 latencies between monCI and bimodal condition of the CI users showed significant correlations.

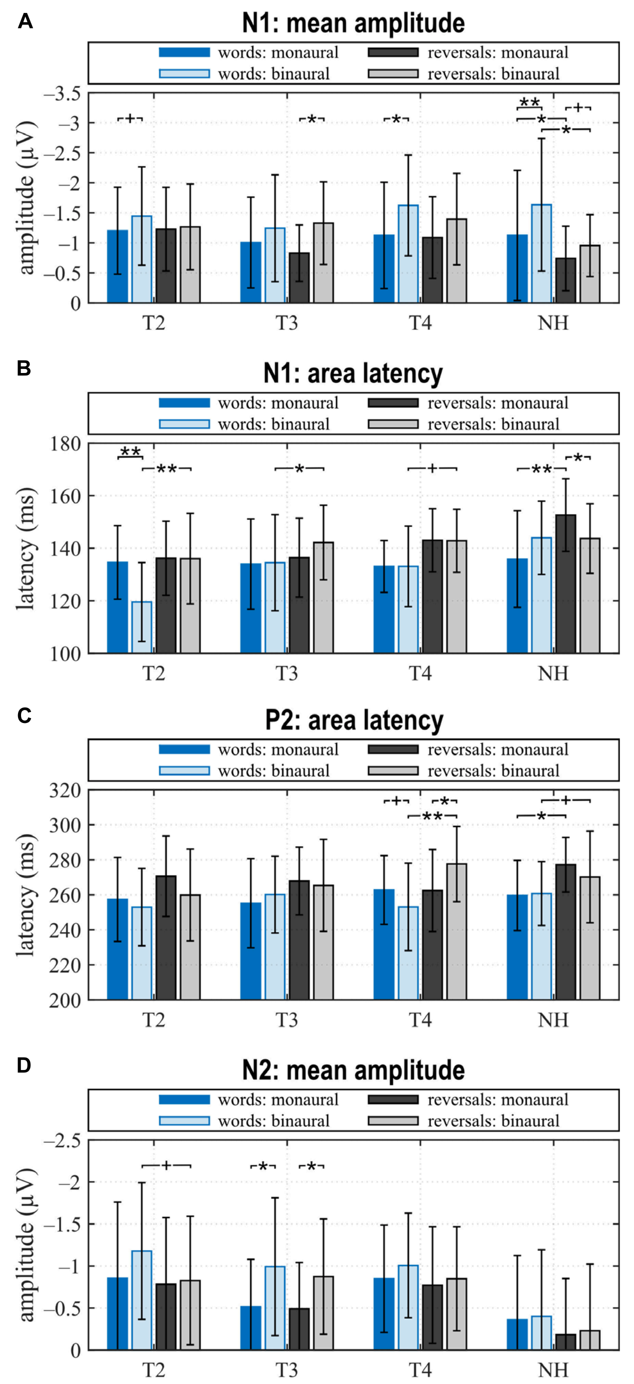

FIGURE 4 | Quantitative AEP results: **(A)** mean amplitude and **(B)** area latency of the N1, **(C)** area latency of P2, and **(D)** mean amplitude of N2 for the categories "words", "reversals", and listening conditions monCI and binaural. **(A–D)** Means with their standard deviations are shown; significant differences between stimulus categories and listening conditions are indicated (**$p < 0.001$, *$p < 0.05$, and trends $^{+}p < 0.1$).

The difference in P2 latency in response to words between monaural electric and bimodal hearing was significantly correlated with SU_N ($r = 0.541$, $p < 0.037$). A shorter bimodal latency and a larger difference with respect to monaural P2 latency correlated with a larger SU_N. Furthermore, SU_N became negative, which indicates binaural interference, if latency in the bimodal condition was longer than with monaural electric hearing.

In addition, the difference in P2 latency in response to reversals between monaural electric and bimodal hearing was significantly correlated with SU_N, but here P2 latency in the monaural LC was significantly shorter, and the binaural minus monaural difference in P2 latency showed a significant inverse correlation with SU_N ($r = -0.620$, $p < 0.014$), again indicating that a shorter P2 latency in the bimodal listening condition was associated with a larger SU_N. Thus, in the CI group and for both stimulus categories, a higher SU_N was associated with a shorter latency in the bimodal compared to the monaural condition, while a negative SU_N can be expected when monaural latency is shorter than latency after bimodal presentation.

DISCUSSION

Aim of the study was to investigate binaural interactions in bimodal listeners during the early phase of CI use evidenced by AEP and audiometric binaural benefits. With CI experience, the grand-averaged N1 amplitude became increasingly similar to the N1 of NH with an expansion of N1 amplitude in response to words and a reduction of the difference in N1 latency between stimulus categories with bimodal listening. In addition, P2 latency differences between stimulus categories increased for the bimodal condition. Several aspects remained different to NH, however, like the absence of a difference in N1 amplitude between stimulus categories, differences in the localization of brain activity during N1, and the large N2 irrespective of listening condition and stimulus category. The latter has been reported earlier for this group of CI users (Balkenhol et al., 2020). These results indicate that the N1 potential, which is related to the detection of an auditory stimulus, approximates the response seen in NH listeners in some aspects within 6 months of CI provision, including evidence for some binaural integration, albeit at significantly higher presentation levels. Grand average of the later N2 response that has been associated with the effort to understand speech in challenging acoustic situations (Balkenhol et al., 2020) remains different, suggesting continued problems with speech recognition for the bimodal listeners. These findings are in agreement with the CI literature (Sandmann et al., 2015; Finke et al., 2016).

In accordance with AEP results, speech tests at T4 evidenced some binaural integration in the form of a positive SU effect. Together with the increased N1 amplitude in the binaural/bimodal listening condition, this may have been due to an increase in perceived loudness with binaural/bimodal hearing as reported in the literature (Hawkins et al., 1987). Although, bimodal listeners also benefited from the HS effect to a similar extent to the age-matched NH group, this does not indicate central alignment of the electrically and acoustically mediated speech as HS is essentially a monaural effect (van Hoesel, 2012).

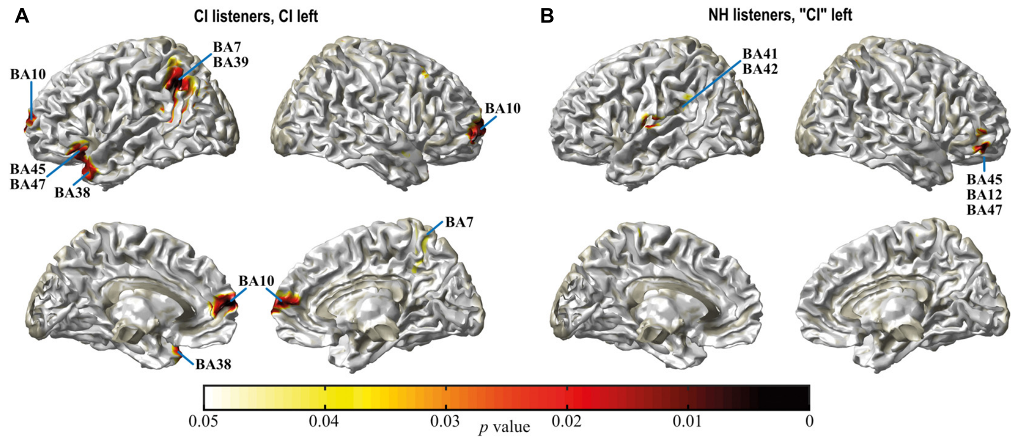

FIGURE 5 | Spatial spread of monaural vs. binaural activation in CI listeners at T4 **(A)** and in NH **(B)** during the N1 interval. Only differences in regions bordering the surface or the midline of the cortex are visible in this illustration. For a complete list of areas with differential activation (see **Table 6**). Differences are more widespread in CI listeners compared to NH. Whereas in NH differential activation located to primary and secondary auditory cortex (BA41, 42), it pertained to auditory association cortex related to speech processing (BA21, 38, 39) and a region related with these areas (BA7) in the bimodal listeners. Darkening of the color scale indicates decreasing *p* values or higher significance.

A second aim of the study was to explore whether bimodal benefit changes with CI experience. AEP data suggest some improvement in bimodal hearing. It is questionable, however, whether this translates to better speech recognition, in particular in view of the N2 that remains different from NH. Interpretation of audiometric results is more straightforward in this context. At the pre-implantation assessment monaural speech recognition tests were not possible. Whereas this evidences an improvement of speech recognition with CI provision, it prevented estimation of binaural results. At the 3-month interval, performance of the CI ear was worse than performance of the HA ear for a substantial number of the participants. Since adding a better ear in the bimodal condition inflates binaural benefits, these values may rather show a better ear effect. Some degree of binaural benefit was suggested through the significant SU_Q effect at the end of the study. While this indicates that the brain can combine the divergent signals transmitted via CI and HA, it requires assessments at a later time to decide whether binaural benefits improve with bimodal experience as suggested by a recent study (Devocht et al., 2017), for instance after performance with the CI ear has reached a stable plateau.

The third aim of the study was to find relevant correlations between binaural benefits and central processing. SU_N showed a significant correlation with latency of the P2 potential.

Binaural Benefits

Study participants continued to use their HA together with the CI, indicating that they accepted this form of hearing provision in their everyday life. Classic binaural benefits are the HS effect based on selection of the ear with better SNR, binaural SU derived from the information being available via two input channels, and the binaural SQ effect which requires central computation of interaural time (ITD) and intensity or level differences (ILD). In addition, speech recognition in noise is improved by spatial separation between signal and noise sources or SRM. Although HS, SU and SQ are largely ascertained for bimodal listeners, not all of them are significant in all published reports (Schafer et al., 2011; Illg et al., 2014; Devocht et al., 2017). Beyond individual capacities and the distribution of hearing ability across ears, the presence and magnitude of binaural effects depends on testing paradigm and material (Schafer et al., 2011), stimulus application (Epstein and Florentine, 2009; Finke et al., 2016), type of masking noise (Illg et al., 2014; Psychny et al., 2014), and the amount of CI experience (Eapen et al., 2009).

The present study group had considerable residual hearing at the HA side, which coincides with participant characteristics from a recent investigation (Devocht et al., 2017), but is distinct to those of earlier reports (Schafer et al., 2011; van Hoesel, 2012). Thus, quantitative comparisons of binaural benefits with those of the earlier reports are possible only to a limited extent. Therefore, bimodal results are mainly compared to the age-matched NH of the current study and to the results by Devocht et al. (2017). As the testing paradigm was similar and participants also used HiRes 90K implants, differences to the current study mainly pertained to longer CI experience (>1 year) and a higher percentage of better CI ears.

The head attenuates sounds at the ear that is shielded from the noise source. Utilization of this HS effect requires the ability to focus on input from the ear with better SNR (Schafer et al., 2011; van Hoesel, 2012). HS was around 4.5 dB SNR, it did not change between T3 and T4, and was not significantly different from the NH group, or experienced bimodal listeners (Devocht et al., 2017). This indicates that our bimodal listeners could exploit HS to a similar extent as NH, and that this does not depend on CI experience, which is consistent with previous findings (Schafer et al., 2011). Results suggest some binaural integration in the bimodal group evidenced by a positive SU. When speech and noise sources coincide in space, the identical signals presented

TABLE 5 | Correlation matrix between Binaural Benefits in the bimodal group at T4 and in NH controls.

Binaural Effect *r* and *p* value	**SRM_{monCI}**		**SU_Q**		**SU_N**		**SQ**		**SQ_{AEP}**		**PTA-4 CI**		**PTA-4 HA**		**PTA-4 asymmetry**	
	CI T4	**NH**	**CI T4**	**NH**	**CI T4**	**NH**	**CI T4**	**NH**	**CI T4**	**NH**	**CI T4**	**NH**	**CI T4**	**NH**	**CI T4**	**NH**
HS	0.819 *<0.0001*		−0.579 *0.030*													
SRM_{monCI}			−0.653 *0.011*					−0.604 *0.022*		−0.619 *0.018*		−0.542 *0.045*				
SRM_{bin}						−0.839 *<0.0001*										
SN_Q							0.608 *0.021*					0.756 *0.001*			−0.698 *0.005*	−0.655 *0.011*
SU_N							0.603 *0.023*									
SQ_{AEP}														0.927 *<0.0001*		
PTA-4 CI													0.926 *<0.001*		−0.723 *0.003*	

Pearson correlation coefficients r and corresponding p values (in italics) are shown.

to both ears lead to increased perceptual loudness and improved speech perception. This does not require the listener to use ITD or ILD, but relies on redundancy of the input (Hawkins et al., 1987; Bronkhorst and Plomp, 1990; Endrass et al., 2004; Schafer et al., 2011; Avan et al., 2015). Beyond that, the complementary nature of information transmitted via CI and HA is thought to be an important contributor to SU in bimodal listeners (van Hoesel, 2012). With a T4 group average of 7.7 dB SNR for SU_Q and 1.9 dB SNR for SU_N, SU was similar to that of NH, which in turn was similar to the SU_N reported previously for NH using a similar testing paradigm (Bronkhorst and Plomp, 1989). Similarly to the present study, Morera et al. (2005) have reported a reduction of SU between quiet and noise, and a significant SU for the quiet but not for the noise condition. This was for a group of bimodal listeners with about 6 months of CI experience. Furthermore, SU was found to be lower close to threshold (Morera et al., 2005), and in particular when tested at threshold using adaptive paradigms such as the one used in the present study (Schafer et al., 2011). As SU_N for the current NH group was as low as in bimodal listeners and SU appears to develop early after CI provision, at least in bilateral CI users (Eapen et al., 2009), the higher SU_N of 4.2 ± 0.9 dB SNR reported for experienced bimodal listeners (Devocht et al., 2017) may rather be the result of a difference in sample characteristics rather than more CI experience.

Bilateral symmetrical high-frequency hearing loss has little effect on SU (Hawkins et al., 1987), whereas asymmetry of hearing thresholds reduces SU_Q considerably in NH (Heil, 2014). In line with this, symmetry of hearing thresholds, here assessed via CI- and HA-aided PTA-4, correlated significantly with SU_Q in the CI group. In accordance, published CI literature suggests that SU is more affected by the interactions between CI and HA performance than by HA performance alone, with greater SU correlating with a smaller difference between CI and HA performance (van Hoesel, 2012; Yoon et al., 2015). In support, a CI simulation study found evidence for a significant binaural integration advantage when the CI simulation ear had a similar level of performance to the other ear (Ma et al., 2016).

In contrast, there was no benefit in spatial unmasking for our bimodal group. Monaural SRM_{monCI} was low and of similar magnitude as for experienced CI listeners (Devocht et al., 2017), while binaural SRM_{bin} was essentially absent (Devocht et al., 2017): 0.8 ± 1.0 dB SNR; current: 1.0 ± 4.0 dB SNR). In contrast, NH listeners of the present study benefited from SRM_{monCI} and SRM_{bin} of about 6 dB SNR each, a finding which is in line with previous work (Bronkhorst and Plomp, 1989). Part of the monaural SRM_{monCI} is attributed to the HS (Williges et al., 2015) and as suggested by the strong and highly significant correlation between these measures (**Table 5**), while binaural cues that promote SRM_{bin} are ITD and ILD (Papesh et al., 2017). Thus, absence of a binaural SRM effect is interpreted as an inability of the bimodal listeners to exploit ITD and ILD with current technology, due to differences of the temporal and spectral characteristics of sound information transmitted via CI and HA (van Hoesel, 2012).

In agreement with this interpretation, a significant SQ was not evidenced for the bimodal listeners of the current study. Binaural SQ describes the improvement of intelligibility in noise due to

TABLE 6 | Source localization results for subjects with CI on left ear.

		% significant (mean *t* values)	
Frontal lobe	**Voxel in ROI**	**CI listeners**	**NH listeners**
SFG, Superior Frontal Gyrus, medial area BA10	8,193	55.91 (−2.82)	
SFG, Superior Frontal Gyrus, medial area BA10	7,535	66.45 (−2.78)	
MFG, Middle Frontal Gyrus, area 46	6,299	25.16 (−2.19)	
MFG, Middle Frontal Gyrus, lateral area BA10	6,643	51.78 (−2.76)	
OrG, Orbital Gyrus, orbital area BA12/47	3,726	33.76 (−2.04)	
IFG, Inferior Frontal Gyrus, rostral BA45	2,971		34.26 (−1.90)
OrG, Orbital Gyrus, lateral area BA12/47	4,059	39.10 (−2.10)	
OrG, Orbital Gyrus, lateral area BA12/47	4,714		24.37 (−1.15)
Temporal lobe			
STG, Superior Temporal Gyrus, medial area BA38	5,294	33.47 (−2.15)	
STG, Superior Temporal Gyrus, lateral area BA38	2,166	41.92 (−2.31)	
STG, Superior Temporal Gyrus, area 41/42	1,489		49.93 (1.88)
MTG, Middle Temporal Gyrus, rostral area BA21	7,515	33.27 (−2.01)	
Parietal lobe			
SPL, Superior Parietal Lobule, rostral area BA7	3,178	32.35 (1.79)	
SPL, Superior Parietal Lobule, intraparietal BA7	3,590	78.38 (1.69)	
IPL, Inferior Parietal Lobule, caudal BA39	9,422	20.83 (1.30)	
IPL, Inferior Parietal Lobule, rostrodorsal BA39	7,928	76.80 (1.73)	
IPL, Inferior Parietal Lobule, rostroventral BA39	10,691	45.49 (1.58)	
POG, Postcentral Gyrus, area BA1/2/3	4,775		39.25 (1.61)
Insula			
INS, Insular Gyrus, ventral agranular insula	1,698	34.28 (−2.28)	
INS, Insular Gyrus, hypergranular insula	2,074		34.81 (1.66)
Cingulate Gyrus			
CG, Cingulate Gyrus, subgenual BA32	3,250	20.28 (−1.91)	

Brain areas with a significant activation difference between monaural and binaural listening condition in the N1 time interval in at least 20% of their voxels are listed. Red shading is used for differences in the right hemisphere, while blue shading represents differences in the left hemisphere. Darkest shading indicates a significant difference in at least 75% of the voxels, lightest label is used for differences in less than 25% of the voxels, and tones in between represent categories 50–75% and 25–49% of voxels with a significant monaural vs. binaural difference. Positive t values indicate a larger N1 with bimodal listening, while negative t values indicate a smaller N1 with bimodal hearing. If available, Brodmann areas (BA) are indicated. Note that the spatial extent of differences is larger and pertains to language-associated cortex in CI listeners.

addition of input at the contralateral ear with a poorer SNR than in the monaurally active ear (Schleich et al., 2004). It is seen as a binaural phenomenon based on computation of ITD and ILD in the central auditory system (van Hoesel, 2012). With their limited CI experience, our bimodal listeners were not able to exploit ITD and ILD, whereas Devocht et al. (2017) report a SQ of 2.6 dB SNR for experienced bimodal listeners.

Absence of a significant SQ in the OlSa S0NHA condition may have been a result of insensitivity of the adaptive listening paradigm, particularly as a significant SQ could not be shown for the NH group either. A meta-analysis supports this view and suggests that in contrast to supra-threshold testing at fixed SNR levels, the adaptive paradigm may be too insensitive to evidence a SQ, because it is conducted at threshold levels (Schafer et al., 2011). Absence of a significant SQ_{AEP} in the CI group, even at T4, which was tested using a fixed SNR, is in contrast to the highly significant SQ_{AEP} of the NH group, however, and suggests that at this early stage of bimodal experience, there may be no gain derived from SQ. In contrast, the investigation of Morera et al. (2005) on listeners with 6 months of CI experience and testing at a fixed SNR of +10 dB SNR evidenced a significant SQ effect, but here noise was presented at the side of the CI ear. While longitudinal studies investigating the development of the SQ effect in bimodal listeners do not exist as yet, a longitudinal study accompanying bilateral CI recipients over 4 years found SQ to arise at about 12 months after implantation and to continue to increase thereafter (Eapen et al., 2009). Others report significant SQ effects of between 1.9 and 2.9 dB SNR for listeners with more than 12 months of bimodal experience (Kokkinakis and Pak, 2014; Psychny et al., 2014; Devocht et al., 2017; van Loon et al., 2017). One study addressed the effect of adding a contralateral CI in participants with fairly good acoustic hearing (van Loon et al., 2017), and another tested intelligibility in the presence of a speech interferer (Kokkinakis and Pak, 2014) which increases binaural benefits in comparison to noise interferers (Psychny et al., 2014).

Overall, bimodal listeners were able to benefit from HS and SU effects, the latter despite the fact that the input from the two ears was dissimilar, but with limited CI experience of about 6 months they could not benefit from ITD and ILD evidenced by absence of SRM_{bin} and SQ.

AEP

Together with the studies by Sasaki et al. (2009) and Soshi et al. (2014), this is the only AEP study that addressed bimodal hearing. While Soshi et al. (2014) compared listening in quiet and noise in the bimodal condition and observed reduced N1-P2 amplitudes following speech syllables in noise compared to quiet, Sasaki et al. (2009) compared responses to pure tone stimuli between monaural and binaural listening in a mixed group of bimodal and bilateral CI users and reported shortened latencies of N2 and P3 potentials in the binaural conditions. In contrast, our earlier publication described changes in binaural processing of words with bimodal experience (Balkenhol et al., 2020), and the current investigation explored whether a bimodal benefit develops with CI experience. Bimodal benefit was estimated from differences of N1, P2, and N2 potentials between monaural electric and bimodal hearing in the spatial S0NHA constellation in response to monosyllabic words and their time-reversed sound tracks presented within speech-shaped noise.

NH listeners were expected to show maximal effects and served as a benchmark with which to compare the CI users. Grand average N1, P2, and N2 were present in CI and NH listeners for both listening conditions and in response to both stimulus categories. Several aspects of the N1 and P2 potentials differed between monaural and binaural listening and between stimulus categories in the CI and NH groups, while the late N2 potential differed between groups as reported previously (Balkenhol et al., 2020). As binaural benefits were inflated at T2 and T3 because of the high number of better HA ears, the discussion focusses on the T4 assessment.

A promising result from the present study is the approximation of the difference in N1 amplitude between electric and bimodal hearing to the monaural vs. binaural difference observed for NH listeners. In NH, N1 amplitude in response to words was significantly larger in the binaural condition. In the CI group, the difference in N1 amplitude increased between T2 and T4 and attained statistical significance for the T4 assessment. This finding is in line with the results of a previous study employing an auditory discrimination task to investigate monaural electric hearing (Sandmann et al., 2015), implicating early restoration of N1 amplitudes following CI provision.

Another significant difference between listening conditions pertained to a reduction of P2 latency with binaural listening. While this reduction was observed only in response to words for NH listeners, at T4, shorter P2 latencies in the bimodal condition were observed for both stimulus categories. Significant latency reductions between monaural and bimodal conditions were also reported by Sasaki et al. (2009) following stimulation with pure tones. In that study, latency differences concerned the later event-related N2 and P3 potentials, however. Moreover, Okusa et al. (1999) report longer P2 latencies with increasing task difficulty. Albeit not mentioned explicitly, this finding probably pertained to monaural electric hearing.

The current increase of N1 amplitude and the decrease in P2 latency for binaural/bimodal listening in NH and for the bimodal group at T4 is compatible with an increase in perceived loudness related to binaural loudness summation. The positive SU in NH and at T4 and the significant bivariate correlation between SU_N and the P2 latency difference between monaural and bimodal LC at T4 affirm this interpretation. Perceived loudness of a binaurally presented stimulus is louder than its monaural presentation (Hawkins et al., 1987), and there is ample evidence, that N1 and P2 amplitudes increase while their latencies decrease concomitant with the intensity of tones or speech syllables presented in quiet and within background noise (Firszt et al., 2002; Martin et al., 2008; Kim et al., 2012; Sharma et al., 2014; Prakash et al., 2016). Thus, results suggest that with a CI experience of about 6 months, the current sample of bimodal listeners could benefit from SU.

Further factors that influence processing of complex auditory stimuli are familiarity and attention. N1 amplitude was larger while N1 and P2 latencies were reduced in response to word stimuli, although at times pertaining to different comparisons in NH and at T4. Existing literature indicates a stronger response and a more rapid evaluation of familiar stimuli (Kuhl et al., 2007; Kuuluvainen et al., 2014) and suggests that reversed speech sounds are less easily classified within familiar phonetic categories (Binder et al., 2000). Endrass et al. (2004) interpret this as a neurophysiological manifestation of a bilateral redundancy gain, that improves processing of learned meaningful stimuli such as words, but not of complex, unfamiliar, or meaningless stimuli (Endrass et al., 2004). Noteworthy in this context is that significant differences, depending on stimulus categories, were found for more comparisons in NH than in bimodal listeners, suggesting that the familiar sound trace is processed more effectively in NH listeners.

Alternatively, or in addition, attention may have contributed to the difference in N1 amplitudes and P2 latencies related to words and reversals, as suggested previously (Lange, 2013). Our participants were instructed to respond to words but not to reversals. Attention increases amplitude of the N1 for target sounds in CI listeners, but not for distractor sounds (Paredes-Gallardo et al., 2018). Moreover, a study investigating the neural dynamics in the auditory cortex for attending and ignoring showed that responses to the to-be attended stimuli were enhanced around 100 ms post-onset, whereas ignoring led to a decrease in this response (Chait et al., 2010). While a significant difference in N1 amplitude, depending on stimulus category, emerged in NH listeners, current CI listeners did not show such a difference. As CI listeners become more effective at selectively listening to a target stream over time (Paredes-Gallardo et al., 2018), the limited CI experience of the current group may not have been sufficient to produce this effect. Taken together, the most likely interpretation of increased N1 amplitude and shorter P2 latencies in the binaural condition in NH and at T4 appear to be a result of loudness summation. The significant inverse correlation between P2 latency and SU_N in CI users affirms this interpretation.

Source Localization

The grand average of the N1 showed similar monaural vs. binaural amplitude differences in NH and CI groups. N1 consists of several subcomponents with spatially and temporally overlapping neural generators (Näätänen and Picton, 1987), however, leaving the possibility that activation contributing to the N1 response may differ. Based on LORETA source localization, Zhang et al. (2011), suggest spatial and temporal involvement of the following N1 contributors. Upon auditory stimulation, a pre-attentive mechanism in the frontal lobe is activated. If attention is involved, the attention-driven detection of the stimulus is then transmitted to temporal and parietal areas, with involvement of the parietal lobe probably reflecting the matching of sensory information to memory templates. To explore potential similarities and discrepancies between bimodal and NH listeners, source localization analyses were performed for the difference in activation between monaural vs. binaural LCs in the N1 interval. Taking this approach, activity in brain areas that are active to the same extent in monaural and binaural listening does not show, while areas with differential activation are highlighted.

The side of monaural stimulation may influence the N1 response (Gilmore et al., 2009; Hanss et al., 2009) and consequently also the difference between monaural and binaural activation. Therefore, source localization analyses were performed for the subgroups with monaural stimulation of left ears in both groups. Several brain areas exhibited differential activation between listening conditions. Localization differed between CI and NH groups and differences were more widespread in the CI listeners.

For NH, differential activation was found in the left auditory cortex (BA41, 42) with an augmented N1 response in the binaural condition. Organization of the ascending auditory pathways (Malmierca and Hackett, 2009) provides strong evidence for a contra-laterality effect in the N1 interval over the auditory cortex as a function of ear of stimulation and balanced bilateral activation with binaural stimulation (Gilmore et al., 2009). The difference in ipsilateral auditory cortex in NH is therefore interpreted to result from bilateral activation, with binaural presentation leading to a relatively stronger N1 response ipsilateral to the ear that was active in the monaural condition. The increased negativity in left insula in the binaural condition is compatible with the known connections between insula and ipsilateral auditory cortex (Augustine, 1996; Hackett, 2015), and suggests bilateral activation of the insula with binaural listening. The insula is functionally complex and highly connected, and parts of it are seen as a core region of the language system, interfacing sensory and motor language-associated areas (Augustine, 1996; Ardila et al., 2016). Negativity in the binaural condition was reduced in right-hemispheric IFG and OrG, including BA45, which on the left side is regarded as the core of Broca's area involved in language production (Ardila et al., 2016). Given that increased listening effort has been related to increased activation in the ventral frontal lobe as discussed in Balkenhol et al. (2020), reduced negativity may suggest less listening effort during binaural listening for NH listeners.

The pattern of differential activity in bimodal listeners deviates substantially from that seen in NH. In particular, no difference existed in the primary or secondary auditory cortex (BA41, 42). Possible reasons for this finding are suppression of one ear in the bimodal condition and/or dismantling and reorganization of connections in the auditory system as a result of long-term hearing impairment. It remains to be seen, whether contra-laterality in the auditory pathways increases with continued bimodal hearing.

Differences between electric and bimodal hearing were observed in auditory association areas in the temporal and parietal lobes, mainly in the left hemisphere, and thus in areas that are involved in the sensory aspects of speech processing. Negativity in left temporal areas (BA21, BA38) was smaller with bimodal listening. Left BA21 is part of the core of Wernicke's area that is involved in sensory aspects of speech processing (Ardila et al., 2016), whereas left BA38 was shown to be sensitive to the acoustic-phonetic contents of human speech (Leaver and Rauschecker, 2010). Reduction of activation in temporal areas in the bimodal condition may suggest less, or less synchronized, activation in these areas with bimodal listening.

On the contrary, N1 negativity increased with bimodal hearing in the parietal lobe, suggesting enhanced processing in parietal areas. Affected areas were BA7 and BA39 (angular gyrus) in the left hemisphere and also BA7 in the right. BA39 is a part of the extended Wernicke area as defined by Ardila et al. (2016) and is thought to be involved in associating language with other types of information (Ardila et al., 2016). Left BA7, in the superior parietal lobe, interacts with regions of the extended Wernicke area, participating in language processing and temporal context recognition (Ardila et al., 2016).

Conversely, during bimodal listening activation was reduced bilaterally in several prefrontal areas including, left BA47, which is part of Broca's complex and thus involved in language production (Ardila et al., 2016). Activation during bimodal listening was also reduced in the left insula and cingulate gyrus (BA32) that play a coordinating role in interconnecting the perceptive and productive language system (insula), or are associated with cognitive and emotional aspects of language processing (BA32) (Ardila et al., 2016).

Differences between bimodal and NH listeners and the spatial extent of differences in activity, suggest that neuronal circuits differ considerably between groups and between electric and bimodal processing, at least during the initial period of CI use: the latter possibly due to discrepancies in the information conveyed via CI and HA. In general, results from bimodal listeners suggest that contralaterality in the primary auditory cortex is reduced and that a larger part of the cortex involved in language associations is occupied with speech processing during the N1 time window.

Advantages and Limitations

To the best of our knowledge, this is the first EEG study on bimodal CI users which uses a large set of monosyllabic words. We did this to create a more natural listening situation and to avoid habituation. We could show that this approach is successful in producing several separable AEP. In support of our study design, the present study's findings are consistent with several other studies investigating speech perception in NH and CI listeners.

Another advantage of our study is that it included an age-matched control group, which allows direct comparison of the amount of binaural benefit that is possible in the testing conditions. A further advantage, but also a potential limitation, is that our CI users used the same CI provision, both in terms of implant and speech processor model being used.

As in other EEG studies with CI users, the major limitation of our study is the small sample size and the heterogeneity of the CI group, which does not permit generalization of the results. Furthermore, advances in CI and HA technology and expansion of implant criteria limit the interpretation of results in relation to former studies, because both personal as well as technical conditions have changed.

EEG data offer high temporal resolution, which is mandatory for describing evolution of the brain's response to speech stimuli. Because of the inverse problem and the need to employ source localization techniques, there is however, no unambiguous localization of the underlying sources. Therefore, localization data should be interpreted with caution.

Later follow-up would be worthwhile, although this increases the potential problem of worsening of hearing in the HA ear, which has been observed for the current sample and has been noted by others (Sanhueza et al., 2016; van Loon et al., 2017). As significant improvement of monaural electric hearing occurred during the study interval, which obscures the magnitude of benefit derived from binaural input, the study should be replicated with experienced CI listeners, when full employment of the CI ear is expected.

All participants were tested using their own devices with their clinical setting, because the purpose of the study was to evaluate the binaural benefits in everyday use, as opposed to the effect of optimal and synchronized CI and HA fitting. Results imply however, that to gain maximal binaural benefit, amplification in one or the other device may have to be reduced, as suggested by a reduction of SU with asymmetry of the PTA-4.

CONCLUSION

Major findings of the study are the following:

- With 6 months of CI experience, bimodal listeners were able to make use of the HS and SU effects but did not benefit from SQ or SRM, indicating insufficient alignment of electrically and acoustically transmitted auditory signals in the central auditory system.
- The significant correlation of binaural SU with the bimodal/monaural CI latency difference of the AEP response confirms its potential use as an objective measure for the quality of bimodal hearing.
- EEG results for the bimodal group demonstrated N1 responses that were similar to NH listeners in terms of magnitude and response characteristics.
- Source localization reveals distinct processing for bimodal listeners in the N1 interval, however, suggesting loss of lateralization in the auditory system and augmented associative processing in speech relevant areas. Therefore, it will not be sufficient to use an averaged N1 response to estimate the quality of bimodal processing.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because of ethical or legal reasons. Requests to access the datasets should be directed to EW-F, elisabeth.wallhaeusser-franke@medma.uni-heidelberg.de or TB, tobias.balkenhol@medma.uni-heidelberg.de.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board of the Medical Faculty of Mannheim at Heidelberg University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TB designed the computational framework, collected and analyzed the data, and wrote the manuscript. EW-F designed the study, collected and analyzed the data, and wrote the manuscript. NR did critical review. JS recruitment, collection of clinical data, and critical review. All authors contributed to the article and approved the submitted version.

FUNDING

This study was partly funded by Advanced Bionics AG, Staefa, Switzerland.

REFERENCES

Ardila, A., Bernal, B., and Rosselli, M. (2016). How localized are language brain areas? A review of Brodmann Areas involvement in oral language. *Arch. Clin. Neuropsychol.* 31, 112–122. doi: 10.1093/arclin/acv081

Augustine, J. R. (1996). Circuitry and functional aspects of the insular lobe in primates including humans. *Brain Res. Brain Res Rev.* 22, 229–244. doi: 10.1016/s0165-0173(96)00011-2

Avan, P., Giraudet, F., and Büki, B. (2015). Importance of binaural hearing. *Audiol. Neurootol.* 20(Suppl. 1), 3–6. doi: 10.1159/000380741

Balkenhol, T., Wallhäusser-Franke, E., and Delb, W. (2013). Psychoacoustic tinnitus loudness and tinnitus-related distress show different associations with oscillatory brain activity. *PLoS One* 8:e53180. doi: 10.1371/journal.pone.0053180

Balkenhol, T., Wallhäusser-Franke, E., Rotter, N., and Servais, J. J. (2020). Changes in speech-related brain activity during adaptation to electro-acoustic hearing. *Front. Neurol.* 11:161. doi: 10.3389/fneur.2020.00161

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–528. doi: 10.1093/cercor/10.5.512

Bronkhorst, A. W., and Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *J. Acoust. Soc. Am.* 83, 1508–1516. doi: 10.1121/1.395906

Bronkhorst, A. W., and Plomp, R. (1989). Binaural speech intelligibility in noise for hearing-impaired listeners. *J. Acoust. Soc. Am.* 86, 1374–1383. doi: 10.1121/1.398697

Bronkhorst, A. W., and Plomp, R. (1990). A clinical test for the assessment of binaural speech perception in noise. *Audiology* 29, 275–285. doi: 10.3109/00206099009072858

Chait, M., de Cheveigné, A., Poeppel, D., and Simon, J. Z. (2010). Neural dynamics of attending and ignoring in human auditory cortex. *Neuropsychologia* 48, 3262–3271. doi: 10.1016/j.neuropsychologia.2010.07.007

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Delorme, A., Sejnowski, T., and Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage* 34, 1443–1449. doi: 10.1016/j.neuroimage.2006.11.004

Devocht, E. M. J., Janssen, A. M. L., Chalupper, J., Stokroos, R. J., and George, E. L. J. (2017). The Benefits of bimodal aiding on extended dimensions of speech perception: intelligibility, listening effort, and sound quality. *Trends Hear.* 21:2331216517727900. doi: 10.1177/2331216517727900

Dunlap, W. P., Marx, M. S., and Agamy, G. J. (1981). FORTRAN IV functions for calculating probabilities associated with Dunnett's test. *Behav. Res. Meth. Instr.* 13, 363–366. doi: 10.3758/bf03202031

Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *J. Am. Stat. Assoc.* 50, 1096–1121. doi: 10.1080/01621459.1955.10501294

Eapen, R. J., Buss, E., Adunka, M. C., Pillsbury, H. C. III, and Buchman, C. A. (2009). Hearing-in-noise benefits after bilateral simultaneous cochlear implantation continue to improve 4 years after implantation. *Otol. Neurotol.* 30, 153–159. doi: 10.1097/mao.0b013e3181925025

Endrass, T., Mohr, B., and Pulvermüller, F. (2004). Enhanced mismatch negativity brain response after binaural word presentation. *Eur. J. Neurosci.* 19, 1653–1660. doi: 10.1111/j.1460-9568.2004.03247.x

Epstein, M., and Florentine, M. (2009). Binaural loudness summation for speech and tones presented via earphones and loudspeakers. *Ear. Hear.* 30, 234–237. doi: 10.1097/AUD.0b013e3181976993

Finke, M., Büchner, A., Ruigendijk, E., Meyer, M., and Sandmann, P. (2016). On the relationship between auditory cognition and speech intelligibility in cochlear implant users: an ERP study. *Neuropsychologia* 87, 169–181. doi: 10.1016/j.neuropsychologia.2016.05.019

Firszt, J. B., Chambers, R. D., Kraus, N., and Reeder, R. M. (2002). Neurophysiology of cochlear implant users I: effects of stimulus current level and electrode site on the electrical ABR, MLR, and N1-P2 response. *Ear. Hear.* 23, 502–515. doi: 10.1097/00003446-200212000-00002

Gilmore, C. S., Clementz, B. A., and Berg, P. (2009). Hemispheric differences in auditory oddball responses during monaural versus binaural stimulation. *Int. J. Psychophysiol.* 73, 326–333. doi: 10.1016/j.ijpsycho.2009.05.005

Hackett, T. A. (2015). Anatomic organization of the auditory cortex. *Handb. Clin Neurol.* 129, 27–53. doi: 10.1016/B978-0-444-62630-1.00002-0

Hahlbrock, K. H. (1970). *Sprachaudiometrie: Grundlagen und Praktische Anwendung Einer Sprachaudiometrie für Das Deutsche Sprachgebiet (German Edition)*, 2nd Edn, Stuttgart: Thieme.

Hanss, J., Veuillet, E., Adjout, K., Besle, J., Collet, L., and Thai-Van, H. (2009). The effect of long-term unilateral deafness on the activation pattern in the auditory cortices of French-native speakers: influence of deafness side. *BMC Neurosci.* 10:23. doi: 10.1186/1471-2202-10-23

Hawkins, D. B., Prosek, R. A., Walden, B. E., and Montgomery, A. A. (1987). Binaural loudness summation in the hearing impaired. *J. Speech Hear. Res.* 30, 37–43. doi: 10.1044/jshr.3001.37

Heil, P. (2014). Towards a unifying basis of auditory thresholds: binaural summation. *J. Assoc. Res. Otolaryngol.* 15, 219–234. doi: 10.1007/s10162-013-0432-x

Henkin, Y., Yaar-Soffer, Y., Givon, L., and Hildesheimer, M. (2015). Hearing with two ears: evidence for cortical binaural interaction during auditory processing. *J. Am. Acad. Audiol.* 26, 384–392. doi: 10.3766/jaaa.26.4.6

Holder, J. T., Reynolds, S. M., Sunderhaus, L. W., and Gifford, R. H. (2018). Current profile of adults presenting for preoperative cochlear implant evaluation. *Trends Hear.* 22, 1–16. doi: 10.1177/2331216518755288

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., and Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. *J. Comput. Assisted Tomogr.* 22, 324–333. doi: 10.1097/00004728-199803000-00032

Illg, A., Bojanowicz, M., Lesinski-Schiedat, A., Lenarz, T., and Büchner, A. (2014). Evaluation of the bimodal benefit in a large cohort of cochlear implant subjects using a contralateral hearing aid. *Otol. Neurotol.* 35, e240–e244. doi: 10.1097/MAO.0000000000000529

Kalbe, E., Kessler, J., Calabrese, P., Smith, R., Passmore, A. P., Brand, M., et al. (2004). DemTect: a new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. *Int. J. Geriatr. Psychiatry* 19, 136–143. doi: 10.1002/gps.1042

Kim, J. R., Ahn, S. Y., Jeong, S. W., Kim, L. S., Park, J. S., Chung, S. H., et al. (2012). Cortical auditory evoked potential in aging: effects of stimulus intensity and noise. *Otol. Neurotol.* 33, 1105–1112. doi: 10.1097/MAO.0b013e3182659b1e

Kokkinakis, K., and Pak, N. J. (2014). Binaural advantages in users of bimodal and bilateral cochlear implant devices. *Acoust. Soc. Am.* 135, EL47–EL53. doi: 10.1121/1.4831955

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., and Nelson, T. (2007). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 979–1000. doi: 10.1098/rstb.2007.2154

Kuuluvainen, S., Nevalainen, P., Sorokin, A., Mittag, M., Partanen, E., Putkinen, V., et al. (2014). The neural basis of sublexical speech and corresponding nonspeech processing: a combined EEG-MEG study. *Brain Lang.* 130, 19–32. doi: 10.1016/j.bandl.2014.01.008

Lange, K. (2013). The ups and downs of temporal orienting: a review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory N1 with opposite effects of attention and prediction. *Front. Hum. Neurosci.* 7:263. doi: 10.3389/fnhum.2013.00263

Leaver, A. M., and Rauschecker, J. P. (2010). Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. *J. Neurosci.* 30, 7604–7612. doi: 10.1523/JNEUROSCI.0296-10.2010

Letowski, T., and Champlin, C. (2014). Audiometric calibration: air conduction. *Semin. Hear.* 35, 312–328. doi: 10.1055/s-0034-1390161

Liesefeld, H. R. (2018). Estimating the timing of cognitive operations with MEG/EEG latency measures: a primer, a brief tutorial, and an implementation of various methods. *Front. Neurosci.* 12:765. doi: 10.3389/fnins.2018.00765

Litovsky, R., Parkinson, A., Arcaroli, J., and Sammeth, C. (2006). Simultaneous bilateral cochlear implantation in adults: a multicenter clinical study. *Ear. Hear.* 27, 714–731. doi: 10.1097/01.aud.0000246816.50820.42

Löhler, J., Akcicek, B., Wollenberg, B., Schönweiler, R., Verges, L., Langer, C., et al. (2014). Results in using the Freiburger monosyllabic speech test in noise without and with hearing aids. *Eur. Arch. Otorhinolaryngol.* 272, 2135–2142. doi: 10.1007/s00405-014-3039-x

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*, 2nd Edn, Cambridge, MA: The MIT Press.

Ma, N., Morris, S., and Kitterick, P. T. (2016). Benefits to speech perception in noise from the binaural integration of electric and acoustic signals in simulated unilateral deafness. *Ear. Hear.* 37, 248–259. doi: 10.1097/AUD.0000000000000252

Malmierca, M. S., and Hackett, T. A. (2009). "Structural organization of the ascending auditory pathway," in *The Oxford Handbook of Auditory Science. Vol. 2: The Auditory Brain*, eds A. Rees, and A. Palmer, (Oxford: Oxford University Press), 9–41.

Martin, B. A., Tremblay, K. L., and Korczak, P. (2008). Speech evoked potentials: from the laboratory to the clinic. *Ear. Hear.* 29, 285–313. doi: 10.1097/AUD.0b013e3181662c0e

Molgedey, L., and Schuster, H. G. (1994). Separation of a mixture of independent signals using time delayed correlations. *Phys. Rev. Lett.* 72, 3634–3637. doi: 10.1103/physrevlett.72.3634

Morera, C., Manrique, M., Ramos, A., Garcia-Ibanez, L., Cavalle, L., Huarte, A., et al. (2005). Advantages of binaural hearing provided through bimodal stimulation via a cochlear implant and a conventional hearing aid: a 6-month comparative study. *Acta Otolaryngol.* 125, 596–606. doi: 10.1080/00016480510027493

Müller-Deile, J. (2009). *Verfahren zur Anpassung und Evaluation von Cochlea-Implantat Sprachprozessoren*. Heidelberg: Median Verlag.

Näätänen, R., and Picton, T. (1987). The N1 wave of human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24, 375–425. doi: 10.1111/j.1469-8986.1987.tb00311.x

Okusa, M., Shiraishi, T., Kubo, T., and Nageishi, Y. (1999). Effects of discrimination difficulty on cognitive event-related brain potentials in patients with cochlear implants. *Otolaryngol. Head Neck Surg.* 121, 610–615. doi: 10.1016/s0194-5998(99)70067-7

Onton, J., Westerfield, M., Townsend, J., and Makeig, S. (2006). Imaging human EEG dynamics using independent component analysis. *Neurosci. Biobehav. Rev.* 30, 808–822. doi: 10.1016/j.neubiorev.2006.06.007

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011, 1–9. doi: 10.1155/2011/156869

Oostenveld, R., and Praamstra, P. (2011). The five percent electrode system for high-resolution EEG and ERP measurements. *Clin. Neurophysiol.* 112, 713–719. doi: 10.1016/s1388-2457(00)00527-7

Papesh, M. A., Billings, C. J., and Baltzell, L. S. (2015). Background noise can enhance cortical auditory evoked potentials under certain conditions. *Clin. Neurophysiol.* 126, 1319–1330. doi: 10.1016/j.clinph.2014.10.017

Papesh, M. A., Folmer, R. L., and Gallun, F. J. (2017). Cortical measures of binaural processing predict spatial release from masking performance. *Front. Hum. Neurosci.* 11:124. doi: 10.3389/fnhum.2017.00124

Paredes-Gallardo, A., Innes-Brown, H., Madsen, S. M. K., Dau, T., and Marozeau, J. (2018). Auditory stream segregation and selective attention for cochlear implant listeners: evidence from behavioral measures and event-related potentials. *Front. Neurosci.* 12:581. doi: 10.3389/fnins.2018.00581

Pascual-Marqui, R. D. (2007). Discrete, 3D distributed, linear imaging methods of electric neuronal activity. Part 1: exact, zero error localization. *arXiv* [Preprint], Available online at: http://arxiv.org/pdf/0710.3341 (accessed April 6, 2020).

Pascual-Marqui, R. D. (2009). "Theory of the EEG inverse problem," in *Quantitative EEG Analysis: Methods and Clinical Applications*, eds S. Tong, and N. V. Thakor, (Norwood, MA: Artech House), 121–140.

Prakash, H., Abraham, A., Rajashekar, B., and Yerraguntla, K. (2016). The effect of intensity on the speech evoked auditory late latency response in normal hearing individuals. *J. Int. Adv. Otol.* 12, 67–71. doi: 10.5152/iao.2016.1776

Psychny, V., Landwehr, M., Hahn, M., Lang-Roth, R., Walger, M., and Meister, H. J. (2014). Head shadow, squelch, and summation effects with an energetic or informational masker in bilateral and bimodal CI users. *Speech Lang. Hear. Res.* 57, 1942–1960. doi: 10.1044/2014_JSLHR-H-13-0144

Reiss, L. A. J., Eggleston, J. L., Walker, E. P., and Oh, Y. (2016). Two ears are not always better than one: mandatory vowel fusion across spectrally mismatched ears in hearing-impaired listeners. *J. Assoc. Res. Otolaryngol.* 17, 341–356. doi: 10.1007/s10162-016-0570-z

Sandmann, P., Plotz, K., Hauthal, N., de Vos, M., Schönfeld, R., and Debener, S. (2015). Rapid bilateral improvement in auditory cortex activity in postlingually deafened adults following cochlear implantation. *Clin. Neurophysiol.* 126, 594–607. doi: 10.1016/j.clinph.2014.06.029

Sanhueza, I., Manrique, R., Huarte, A., Ruiz de Erenchun, I., and Manrique, M. (2016). Bimodal stimulation with cochlear implant and hearing aid in cases of highly asymmetrical hearing loss. *J. Int. Adv. Otol.* 12, 16–22. doi: 10.5152/iao.2016.2185

Sasaki, T., Yamamoto, K., Iwaki, T., and Kubo, T. (2009). Assessing binaural/bimodal advantages using auditory event-related potentials in subjects with cochlear implants. *Auris Nasus Larynx* 36, 541–546. doi: 10.1016/j.anl.2008.12.001

Schafer, E. C., Amlani, A. M., Paiva, D., Nozari, L., and Verret, S. (2011). A meta-analysis to compare speech recognition in noise with bilateral cochlear implants and bimodal stimulation. *Int. J. Audiol.* 50, 871–880. doi: 10.3109/14992027.2011.622300

Schleich, P., Nopp, P., and D'Haese, P. (2004). Head Shadow, squelch, and summation effects in bilateral users of the MED-EL COMBI 40/40+ cochlear implant. *Ear. Hear.* 25, 197–204. doi: 10.1097/01.aud.0000130792.43315.97

Servais, J. J., Hörmann, K., and Wallhäusser-Franke, E. (2017). Unilateral cochlear implantation reduces tinnitus loudness in bimodal hearing: a prospective study. *Front. Neurol.* 8:60. doi: 10.3389/fneur.2017.00060

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* 52, 591–611. doi: 10.1093/biomet/52.3-4.591

Sharma, M., Purdy, S. C., Munro, K. J., Sawaya, K., and Peter, V. (2014). Effects of broadband noise on cortical evoked auditory responses at different loudness levels in young adults. *Neuroreport* 255, 312–319. doi: 10.1097/WNR.0000000000000089

Soshi, T., Hisanaga, S., Kodama, N., Kanekama, Y., Samejima, Y., Yumoto, E., et al. (2014). Event-related potentials for better speech perception in noise by cochlear implant users. *Hear. Res.* 316, 110–121. doi: 10.1016/j.heares.2014.08.001

van Hoesel, R. J. (2012). Contrasting benefits from contralateral implants and hearing aids in cochlear implant users. *Hear. Res.* 288, 100–113. doi: 10.1016/j.heares.2011.11.014

van Loon, M. C., Smits, C., Smit, C. F., Hensen, E. F., and Merkus, P. (2017). Cochlear implantation in adults with asymmetric hearing loss: benefits of bimodal stimulation. *Otol. Neurotol.* 38, e100–e106. doi: 10.1097/MAO.0000000000001418

Wagener, K., Kollmeier, B., and Kühnel, V. (1999a). Entwicklung und evaluation Eines Satztests für die deutsche Sprache I: design des Oldenburger Satztests. *Z. Audiol.* 38, 4–15.

Wagener, K., Kollmeier, B., and Kühnel, V. (1999b). Entwicklung und evaluation Eines Satztests für die deutsche Sprache I: evaluation des Oldenburger Satztests. *Z. Audiol.* 38, 86–95.

Wagener, K., Kollmeier, B., and Kühnel, V. (1999c). Entwicklung und evaluation Eines Satztests für die deutsche Sprache Teil II: optimierung des Oldenburger Satztests. *Z. Audiol.* 38, 44–56.

Wallhäusser-Franke, E., Balkenhol, T., Hetjens, S., Rotter, N., and Servais, J. J. (2018). Patient benefit following bimodal CI-provision: self-reported abilities vs. hearing status. *Front. Neurol.* 9:753. doi: 10.3389/fneur.2018.00753

Westerhausen, R., Bless, J., and Kompus, K. (2015). Behavioral laterality and aging: the free-recall dichotic-listening right-ear advantage increases with age. *Dev. Neuropsychol.* 40, 313–327. doi: 10.1080/87565641.2015.1073291

Williges, B., Dietz, M., Hohmann, V., and Jürgens, T. (2015). Spatial release from masking in simulated cochlear implant users with and without access to low-frequency acoustic hearing. *Trends Hear.* 19:2331216515616940. doi: 10.1177/2331216515616940

Yoon, Y. S., Shin, Y. R., Gho, J. S., and Fu, Q. J. (2015). Bimodal benefit depends on the performance difference between a cochlear implant and a hearing aid. *Cochlear Implants Int.* 16, 159–167. doi: 10.1179/1754762814Y.0000000101

Zhang, F., Deshpande, A., Benson, C., Smith, M., Eliassen, J., and Fu, Q. J. (2011). The adaptive pattern of the auditory N1 peak revealed by standardized low-resolution brain electromagnetic tomography. *Brain Res.* 1400, 42–52. doi: 10.1016/j.brainres.2011.05.036

Conflict of Interest: The authors declare that this study was partly funded by Advanced Bionics, Staefa, Switzerland. Advanced Bionics AG manufactures the device under investigation in this study. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 09 March 2021
doi: 10.3389/fnins.2021.608156

Higher Right Hemisphere Gamma Band Lateralization and Suggestion of a Sensitive Period for Vocal Auditory Emotional Stimuli Recognition in Unilateral Cochlear Implant Children: An EEG Study

Giulia Cartocci[1,2], Andrea Giorgi[1,2], Bianca M. S. Inguscio[2,3], Alessandro Scorpecci[4], Sara Giannantonio[4], Antonietta De Lucia[5], Sabina Garofalo[5], Rosa Grassia[6], Carlo Antonio Leone[6], Patrizia Longo[7], Francesco Freni[7], Paolo Malerba[8] and Fabio Babiloni[1,2,9]*

[1] Laboratory of Industrial Neuroscience, Department of Molecular Medicine, Sapienza University of Rome, Rome, Italy, [2] BrainSigns Srl, Rome, Italy, [3] Cochlear Implant Unit, Department of Sensory Organs, Sapienza University of Rome, Rome, Italy, [4] Audiology and Otosurgery Unit, "Bambino Gesù" Pediatric Hospital and Research Institute, Rome, Italy, [5] Otology and Cochlear Implant Unit, Regional Referral Centre Children's Hospital "Santobono-Pausilipon", Naples, Italy, [6] Department of Otolaryngology/Head and Neck Surgery, Monaldi Hospital, Naples, Italy, [7] Department of Otorhinolaryngology, University of Messina, Messina, Italy, [8] Cochlear Italia Srl, Bologna, Italy, [9] Department of Computer Science and Technology, Hangzhou Dianzi University, Xiasha Higher Education Zone, Hangzhou, China

OPEN ACCESS

Edited by:
Marc Schönwiesner, Leipzig University, Germany

Reviewed by:
Sébastien Paquette, McGill University, Canada
Andrej Kral, Hannover Medical School, Germany

***Correspondence:**
Giulia Cartocci
giulia.cartocci@uniroma1.it

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: 19 September 2020
Accepted: 01 February 2021
Published: 09 March 2021

Citation:
Cartocci G, Giorgi A, Inguscio BMS, Scorpecci A, Giannantonio S, De Lucia A, Garofalo S, Grassia R, Leone CA, Longo P, Freni F, Malerba P and Babiloni F (2021) Higher Right Hemisphere Gamma Band Lateralization and Suggestion of a Sensitive Period for Vocal Auditory Emotional Stimuli Recognition in Unilateral Cochlear Implant Children: An EEG Study. Front. Neurosci. 15:608156. doi: 10.3389/fnins.2021.608156

In deaf children, huge emphasis was given to language; however, emotional cues decoding and production appear of pivotal importance for communication capabilities. Concerning neurophysiological correlates of emotional processing, the gamma band activity appears a useful tool adopted for emotion classification and related to the conscious elaboration of emotions. Starting from these considerations, the following items have been investigated: (i) whether emotional auditory stimuli processing differs between normal-hearing (NH) children and children using a cochlear implant (CI), given the non-physiological development of the auditory system in the latter group; (ii) whether the age at CI surgery influences emotion recognition capabilities; and (iii) in light of the right hemisphere hypothesis for emotional processing, whether the CI side influences the processing of emotional cues in unilateral CI (UCI) children. To answer these matters, 9 UCI (9.47 ± 2.33 years old) and 10 NH (10.95 ± 2.11 years old) children were asked to recognize nonverbal vocalizations belonging to three emotional states: positive (achievement, amusement, contentment, relief), negative (anger, disgust, fear, sadness), and neutral (neutral, surprise). Results showed better performances in NH than UCI children in emotional states recognition. The UCI group showed increased gamma activity lateralization index (LI) (relative higher right hemisphere activity) in comparison to the NH group in response to emotional auditory cues. Moreover, LI gamma values were negatively correlated with the percentage of correct responses in emotion recognition. Such observations could be explained by a deficit in UCI children in engaging the left hemisphere for more demanding emotional task, or alternatively by a higher conscious elaboration in UCI than NH children. Additionally, for the UCI group, there was no difference between the CI side and the contralateral side in gamma activity, but a higher gamma activity in the right in comparison to the left hemisphere was found. Therefore, the CI side did not appear to influence the physiologic hemispheric lateralization of emotional processing. Finally, a negative correlation was shown between the age at the CI surgery and the percentage of correct responses in emotion recognition and then suggesting the occurrence of a sensitive period for CI surgery for best emotion recognition skills development.

Keywords: lateralization index, right hemisphere emotion hypothesis, deafness, hearing loss, brain activity, length of cochlear implant use, sensitive period, auditory age

INTRODUCTION

Processing emotional expressions is fundamental for social interactions and communication; in fact, from a very young age, infants are able to detect visual and auditory information in faces and voices of people around them (Grossmann, 2010). Such capability would develop into the skill to recognize and discriminate emotions, thanks to the contribution of the experience and of the maturation of sensory and perceptual systems. This recognition involves a multisensory effect, evidenced by integration effects of facial and vocal information on cerebral activity, which are apparent both at the level of heteromodal cortical regions of convergence (e.g., bilateral posterior superior temporal sulcus), and at unimodal levels of sensory processing (Campanella and Belin, 2007; Davies-Thompson et al., 2019; Young et al., 2020).

In relation to such cross-sensorial and unisensorial effects, hearing impairment could compromise multisensory integration, in relation to its onset, etiology, and severity, leading the patient to rely only or predominantly on the visual modality in communication, including emotional perception and expression (Mildner and Koska, 2014). In fact, for 92% of children with cochlear implant (CI), perception was dominated by vision when visual and auditory speech information conflicted (Schorr et al., 2005). This statement is supported by the results of studies employing the McGurk effect on CI users, which requires the integration of auditory and visual sensory stimuli. For instance, children who received their CI prior to age 30 months accurately identified the incongruent auditory–visual stimuli, whereas children who received their CI after 30 months of age did not (Schorr, 2005). This evidence appears particularly worthy because differently from adults, who mainly prefer visual modality, infants and young children show auditory processing preference, but in children with congenital hearing impairment, such auditory dominance appears absent. Interestingly, in post-lingually deaf CI patients, such greater relying on visual information, indexed by higher speech-reading performances than normal-hearing (NH) individuals, led instead to an increased capacity of integrating visual and distorted speech signals, producing higher visuoauditory performances (Rouger et al., 2007). Furthermore, such evidence in post-lingual deaf patients was also supported by neurophysiological assessments, evidencing a positive correlation between visual activity and auditory speech recovery, suggesting a facilitating role for the visual modality in auditory words' perception during communicative situations (Strelnikov et al., 2013). With respect to general processing preferences, contrary to adults, who prefer the visual modality (Scherer, 2003), infants and young children exhibit auditory processing preference. Importantly, congenital hearing-impaired children who underwent auditory–verbal therapy (a therapy limiting visual cue in order to strengthen the auditory pathway for language learning) reported a behavior similar to NH children, which is an overall auditory preference in response to audiovisual stimuli, although responses did not significantly differ from chance (Zupan and Sussman, 2009). Contrary to NH individuals, those with hearing impairments do not benefit from the addition of the auditory cues to the visual mode (e.g., Most and Aviner, 2009). Although the accuracy of emotion perception among children with hearing loss (HL) was lower than that of NH children in auditory, visual, and auditory–visual conditions, in prelingually deaf very young children (about 4–6 years old), the combined auditory–visual mode significantly surpassed the auditory or visual modes alone, as in the NH group, supporting the use of auditory information for emotion perception, probably thanks to intensive rehabilitation (Most and Michaelis, 2012) and neuroplasticity. Such results strongly support the hypothesis of a sensitive period (Kral et al., 2001; Sharma et al., 2005; Gilley et al., 2010) for the establishment of the integration of auditory and visual stimuli.

Thanks to their activity of direct stimulation of the acoustic nerve, converting the auditory stimuli into electrical signals directed to the brain, CIs can successfully restore hearing in profoundly deaf individuals. After intensive rehabilitation, most CI users can reach a good level of speech comprehension. However, the acoustic signal provided by the device is severely degraded, resulting in a poor frequency resolution and deficits in pitch patterns (Gfeller et al., 2007; eHopyan et al., 2012) and pitch changes or direction discrimination (Gfeller et al., 2002) in comparison to NH controls.

Hearing-impaired children go through an early auditory development that is different from that of NH toddlers. This condition would affect their judgment of the emotional content of a stimulus, insofar as the auditory modality resulted as particularly important for the communication of emotions in young children (Baldwin and Moses, 1996; Akhtar and Gernsbacher, 2008). The study of such mechanisms appears of great impact since about 600,000 patients world-wide are CI users (The Ear Foundation, 2017), and many of them are children who were born deaf or lost their hearing within the first few years of life. CI children are a paradigmatic model for the study of emotion recognition skills, as due to the early acquisition of deafness, they learned language through the degraded input of the CI, which greatly affects harmonic pitch perception. This ability is strongly necessary for emotion recognition in voices, and its deficiency could have implications on how child CI users learn to produce vocal emotions (Damm et al., 2019). However, a very recent study provided evidence that also deaf people can develop skills for emotional vocalizations despite the presence of some differences in comparison to NH adults (Sauter et al., 2019). Using unilateral CI (UCI) in children, due to non-physiological development of their auditory system and to their asymmetry in receiving auditory inputs, represents a powerful model of investigation of the possible modulation of the hemispheric specialization and of auditory-related emotional skills development in relation to the restored hearing condition. Additionally, such participants would provide evidence of the possible modulation of the physiological processes of emotion recognition following the restoration of the auditory capabilities, of which the exact time of beginning is due to the CI surgery time. Children, 7–13 years of age, using UCIs perform more poorly than age- and gender-matched controls on the affective speech prosody task but as well as controls in tasks of facial affect perception (Hopyan-Misakyan et al., 2009), as measured by the DANVA-2 (Nowicki and Duke, 1994).

One of the few studies that investigated both auditory recognition and vocal production of emotions did not find any consistent advantage for age-matched NH participants in comparison to three prelingually, bilaterally, profoundly deaf children aged 6–7 years who received CIs before age 2 years; however, confusion matrices among three of the investigated emotions (anger, happiness, and fear) showed that children with and without hearing impairment may rely on different cues (Mildner and Koska, 2014).

With respect to emotional skills attainment and in relation to the hemispheric specialization for emotional processing (Gainotti, 2019), it is interesting to consider that patients enrolled in the present study were UCI users, that is, single-side deaf (SSD) patients. In fact, in SSD population, it was evidenced that the occurrence of a massive reorganization of aural preference in favor of the hearing ear is greater than the precocity of unilateral HL onset, therefore supporting the importance of a short time between the first and second implantation in children (Kral et al., 2013; Gordon et al., 2015; Gordon and Papsin, 2019).

Concerning neural correlates of emotion recognition, gamma band electroencephalogram (EEG) was found to be particularly sensitive for emotion classification (Li and Lu, 2009; Yang et al., 2020). Gamma band cerebral activity has been previously linked to facial emotion recognition processes; for instance, a right hemisphere dominance in gamma activity was found during emotional processing of faces in comparison to neutral ones (e.g., Balconi and Lucchiari, 2008). Such evidences are in accord to the right hemisphere hypothesis for emotion processing, that starting from observations on patients with single hemisphere lesions states the dominance of the right hemisphere for every kind of emotional response (Gainotti, 2019). With specific regard to emotional prosody processing and brain activity lateralization, Kotz and colleagues hypothesized that (i) differentially lateralized subprocesses underlie emotional prosody processing and (ii) the lateralization of emotional prosody can be modulated by methodological factors (Kotz et al., 2006). Furthermore, concerning verbal stimuli, in adult CI users, gamma band–induced activity was found to be higher in NH than in CI users, irrespectively of the valence of the emotions investigated (Agrawal et al., 2013).

On the base of the previous issues, the following experimental questions have been approached in a population of NH and UCI children: (i) Given the non-physiological development of the auditory system in deaf children who underwent hearing restoration through CI use, are the emotional auditory stimuli processed in a similar way than NH children? (ii) Is the auditory age, meant as the age at CI surgery, crucial in the capacity of recognizing emotions? (iii) In light of the evidence that the right hemisphere has a unique contribution in emotional processing – summarized in the right hemisphere emotion hypothesis – does the side of the CI influence the processing of emotional cues in UCI children, or is the "physiological right lateralization" respected?

MATERIALS AND METHODS

Participants

For the present study, 10 NH (6 female, 4 male; 10.95 ± 2.11 years old) and 9 UCI user (UCI; 5 female, 4 male; 9.47 ± 2.33 years old) children were enrolled. Six children had their CI in their right ear and three in their left ear; at the moment of the test, none of them wore any hearing aid in their contralateral ear. All participants were right-handed except for two children: one belonging to the NH and one to the UCI group. Further clinical details of the UCI group are reported in **Table 1**.

Protocol

The task consisted of the recognition of nonverbal vocalizations belonging to a database previously validated and employed in several studies (Sauter et al., 2006, 2010, 2013) and grouped into three emotional states: positive (achievement, amusement, contentment, relief), negative (anger, disgust, fear, sadness), and neutral (neutral, surprise), which participants were asked to match with the corresponding emotional picture (**Figure 1**). For each emotion, six different audio stimuli were reproduced, whereas there was a single corresponding emotional picture for each emotion. The emotional audio stimuli had a mean duration of $1,354.25 \pm 223.39$ ms and were delivered at 65 dB HL (Cartocci et al., 2015, 2018; Marsella et al., 2017; Piccioni et al., 2018) through two loudspeakers placed in front of and behind the participant at the distance of 1 m each, to meet CIs' best requirements for their use. Participants underwent training with the kind of emotional stimuli employed in the study

TABLE 1 | Demographics concerning the UCI group, in particular etiology of deafness, its onset, and duration of deafness before CI surgery.

Participants	Age (years)	Etiology	Onset of deafness	Period of Deafness (years)
PI	11,39	Unknown	Birth	1,38
P2	12,04	Unknown	3 years old	5,91
P3	11,66	Unknown	4 years old	2,25
P4	10,22	Homozygous mutation of the connexin-26 gene	Birth	1,11
P5	7,08	Congenital CMV infection	Birth	3,82
P6	9,99	Homozygous mutation of the connexin-26 gene	Birth	2,93
P7	9,24	Homozygous mutation of the connexin-26 gene	Birth	8,16
P8	12,57	Unknown	3,5 years old	6,41
P9	14,37	Unknown	Birth	13,18

and a familiarization with the experimental protocol. Once the researcher verified the comprehension of the emotional stimuli and the task by the participant, he/she was asked to carefully listen to the emotional audio and then to identify the emotion reproduced by the stimulus pressing one out of five buttons on a customized keyboard, corresponding to the target emotional picture. For instance, the participant heard a laugh, and he/she had to identify the corresponding picture, a smiling young lady, out of five options. There was no time limit set for such identification and matching with the target emotion. Each picture representing the target emotion was placed at least once (and maximum twice) in each of the five positions on the screen. The number of five pictures among which the participant had to identify the target stimulus was chosen in accordance with Orsini et al. (1987), who found for the range of age of the enrolled participants a digit span of more than 4.5 items for both males and females. Stimuli were delivered through E-prime software, in a pseudorandomized order so that it was not possible that two stimuli belonging to the same emotion were consecutive.

The study was carefully explained to all participants and to their parents, who signed an informed consent to the participation. The study was approved by the Bambino Gesù Pediatric Hospital Ethic Committee, protocol 705/FS, and was conducted according to the principles outlined in the Declaration of Helsinki of 1975, as revised in 2000.

EEG

A digital EEG system (BE plus EBNeuro, Italy) was used to record 16 EEG channels (Fp, Fz, F3, F4, F7, F8, T7, T8, P3, P4, P7, P8, O1, O2) according to the international 10/20 system, with a sampling frequency of 256 Hz. The impedances were maintained below 10 kΩ, and a 50-Hz notch filter was applied to remove the power interference. A ground electrode was placed on the forehead and reference electrodes on earlobes. The EEG signal was initially bandpass filtered with a fifth-order Butterworth filter (high-pass filter: cutoff frequency fc = 1 Hz; low-pass filter: cutoff frequency fc = 40 Hz). Because we could not apply independent component analysis because of the low number of EEG channels (i.e., 16), we used a regression-based method to identify and correct eye-blinks artifacts. In particular, the Fpz channel was used to identify and remove eye-blink artifacts by the REBLINCA algorithm (Di Flumeri et al., 2016). This method allows the EEG signal to be corrected without losing data. For other sources of artifacts (e.g., environmental noise, user movements, etc.), specific procedures of the EEGLAB toolbox were employed (Delorme and Makeig, 2004). In particular, the EEG dataset was first segmented into epochs of 2 s through moving windows shifted by 0.125 s. This windowing was chosen with the compromise of having a high number of observations, in comparison with the number of variables, and in order to respect the condition of stationarity of the EEG signal. This is in fact a necessary assumption in order to proceed with the spectral analysis of the signal. Successively, three criteria were applied to those EEG epochs (Aricò et al., 2017; Borghini et al., 2017): (i) threshold criterion (amplitudes exceeding $\pm$ 100 μV); (ii) trend criterion (slope higher than 10 μV/s); and (iii) sample-to-sample criterion (sample-to-sample amplitude difference $>$25 μV).

All EEG epochs marked as "artifact" were removed in order to have a clean EEG signal. In order to accurately define EEG bands of interest, for each participant the individual alpha frequency (IAF) was computed on a closed-eyes segment recorded prior to the experimental task. Thus, the EEG was filtered in the following frequency bands: theta [IAF − 6 ÷ IAF − 2 Hz], alpha [IAF − 2 ÷ IAF + 2 Hz], beta [IAF + 2 ÷ IAF + 16 Hz], and gamma [IAF + 16 ÷ IAF + 25 Hz] (Klimesch, 1999). EEG recordings were segmented into trials, corresponding to audio stimulus listening and target picture matching. The power spectrum density was calculated in correspondence of the different conditions with a frequency resolution of 0.5 Hz. Trials were normalized by subtracting the open-eyes activity recorded before the beginning of the experimental task.

Lateralization Index

The lateralization index (LI) was calculated in order to assess the relative asymmetry between the two cerebral hemispheres' activity during the task execution (audio stimuli perception and target visual stimuli matching), as the right hemisphere theory for emotion predicts a relative higher right activation during emotional stimuli processing.

The LI was calculated on the basis of the formula previously adopted by Vanvooren et al. (2015):

$$\mathbf{L}I = \frac{\mathbf{R{-}S}}{\mathbf{RS}}$$

where R stands for right hemisphere, and L for left hemisphere. The LI ranges from +1, for cortical activity entirely asymmetrical

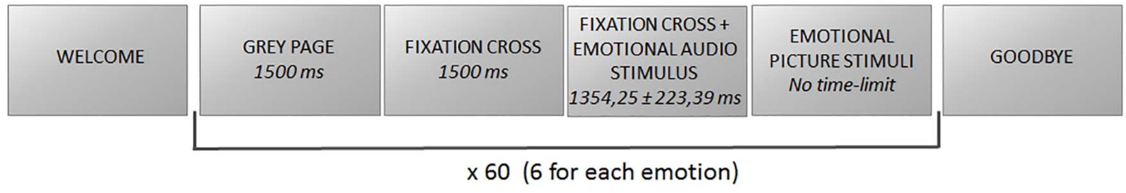

FIGURE 1 | Scheme of the experimental protocol.

to the right hemisphere, to zero for symmetrical cortical activity, and −1 for cortical activity entirely asymmetrical to the left hemisphere. For the right hemisphere activity calculation, the estimation from the following electrodes was averaged: F4, F8, T8, P4, P8, O2, whereas for the left hemisphere. It was averaged from the following ones: F3, F7, T7, P3, P7, O1. The LI was already employed on hearing-impaired children, in particular, SSD children, finding an asymmetry in cortical activity during the execution of a word in noise recognition task influenced by the direction of the background noise in SSD but not in NH children (Cartocci et al., 2019).

Statistical Analysis

Both the percentage of correct responses and LI data were compared between the NH and UCI groups through analysis of variance (ANOVA) with two factors: GROUP (2 levels: NH and UCI) and EMOTIONAL STATE (3 levels: positive, negative, and neutral). A simple regression analysis was performed for investigating the relation between (i) the percentage of correct responses and the LI values, (ii) between the percentage of correct responses and the age at the test execution, and (iii) between the percentage of correct responses and the age at CI surgery.

RESULTS

Behavioral results evidenced a higher percentage of correct responses provided by NH children in comparison to UCI children ($F = 18.898$, $p < 0.001$, partial $\eta^2 = 0.270$) (**Table 2**), but an effect of the emotional state was not seen ($F = 1.890$, $p = 0.161$, partial $\eta^2 = 0.069$), although for both groups the neutral cues were the most difficult to recognize. Neither the interaction between the variable group and emotional state ($F = 0.032$, $p = 0.968$, partial $\eta^2 = 0.001$) was observed (**Figure 2**).

ANOVA results showed higher LI values, indicating a higher activity in gamma band in the right in comparison to the left hemisphere, in UCI in comparison to NH children ($F = 58.656$, $p < 0.001$, partial $\eta^2 = 0.535$) (**Figure 3**), irrespectively of the emotional state (negative, neutral, and positive) ($F = 1.686$, $p = 0.195$, partial $\eta^2 = 0.062$). Additionally, any interaction between the variable groups and emotional state was not found ($F = 1.121$, $p = 0.333$, partial $\eta^2 = 0.042$).

A negative correlation was observed between LI gamma values and the percentage of correct responses ($F = 11.801$, $p = 0.001$, $r = -0.420$, partial $\eta^2 = 0.177$) (**Figure 4**).

Additionally, for the UCI group, any difference between the CI side and the deaf contralateral side in the gamma activity was not

TABLE 2 | Mean percentages of correct responses ± standard deviation for each group (UCI and NH) and for each emotional state.

Group	Negative	Neutral	Positive
NH	86,58% ± 9,82	78,33% ± 18,92	88,33% ± 10,17
UCI	65,05% ± 19,37	58,24% ± 22,17	69,67% ± 21,02

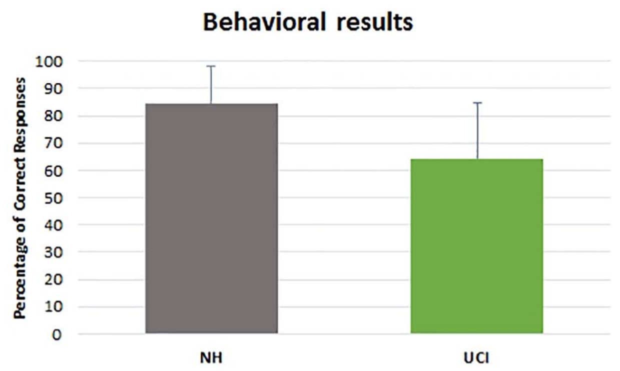

FIGURE 2 | Percentage of correct responses reported by NH and UCI children. Bars describe means, and error bars describe standard deviations.

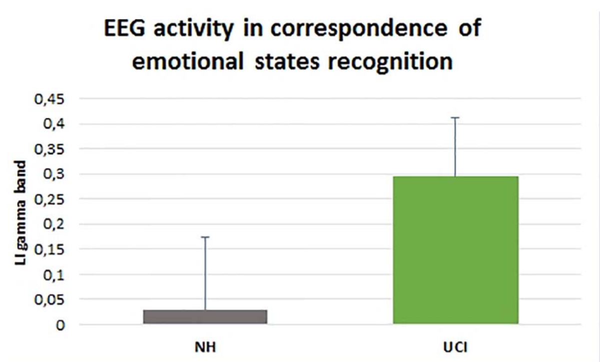

FIGURE 3 | Gamma band activity LI in NH and UCI children. Bars describe means, and error bars describe standard deviations.

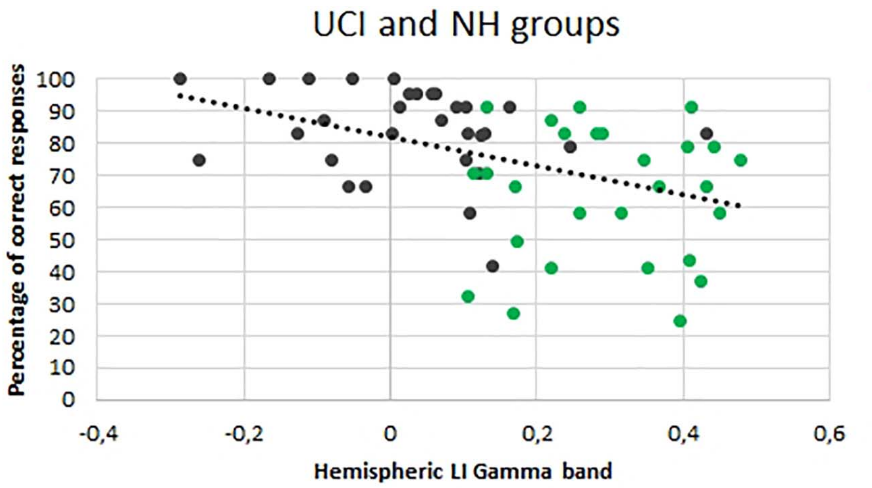

FIGURE 4 | Correlation between the lateralization (LI) gamma values and percentage of correct responses for both the UCI and NH groups. Dark dots represent NH values, and green dots represent UCI values.

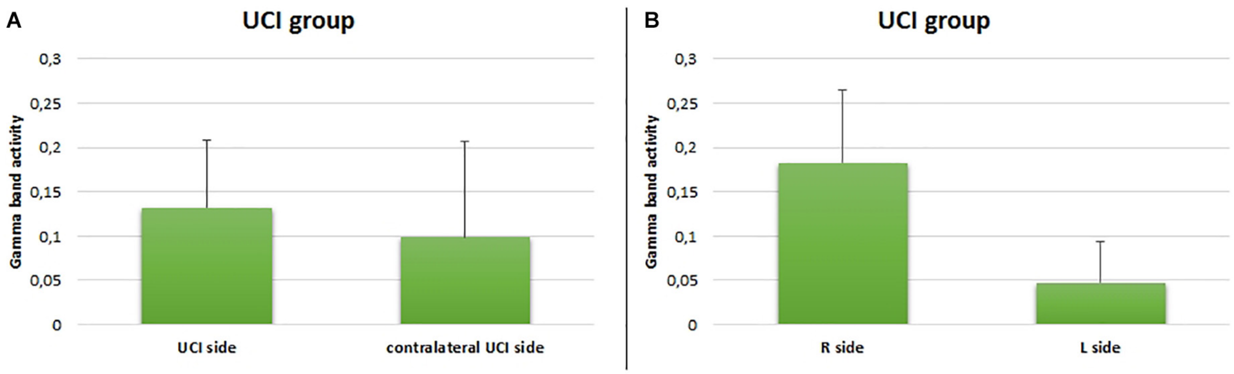

FIGURE 5 | Comparison between gamma activity in the UCI group with respect to the UCI side **(A)** and right or left side **(B)**. Bars describe means, and error bars describe standard deviations.

shown ($F = 0.598$, $p = 0.212$, partial $\eta^2 = 0.032$) (**Figure 5A**), but a higher gamma activity in the right in comparison to the left hemisphere was found ($F = 54.552$, $p < 0.001$, partial $\eta^2 = 0.532$) (**Figure 5B**).

For the UCI group, no correlation was found between the age of the UCI children at the moment of the experiment and the percentage of correct responses ($F = 0.052$, $p = 0.821$, $r = 0.046$, partial $\eta^2 = 0.002$), similarly to the NH children group ($F = 1.130$, $p = 0.297$, $r = 0.197$, partial $\eta^2 = 0.039$). Additionally, a negative correlation was shown between the age at the CI surgery and the percentage of correct response reported by UCI children ($F = 7.030$, $p = 0.014$, $r = 0.468$, partial $\eta^2 = 0.219$) (**Figure 6**). Finally, when calculating the mean of the correct responses for each participant, irrespective of the emotional states, despite the lack of significance ($F = 3.056$, $p = 0.124$, $r = -0.551$, partial $\eta^2 = 0.304$), a higher percentage of correct responses was highlighted, higher than 70%, only in early implanted children, that is, before 3.5 years of age (**Figure 6**, black dots).

DISCUSSION

According to literature, the lower percentage of correct responses provided by UCI children in comparison to NH children highlights their impairment in vocal emotion recognition skills (Agrawal et al., 2013; Wiefferink et al., 2013; Chatterjee et al., 2015; Jiam et al., 2017; Ahmed et al., 2018; Paquette et al., 2018). This would be strongly related to the preverbal and periverbal deafness acquisition. In fact, in a study employing emotional vocal stimuli in adult CI users, such performance difference was not shown (Deroche et al., 2019). Furthermore, there are evidences of different strategies implemented by CI and NH listeners for emotional stimuli recognition, more based on pitch range cues in the former and more relying on mean pitch in the latter group (Gilbers et al., 2015). In addition, such deficit in emotion recognition in UCI children in comparison to NH children appears strictly related to the matter of social interaction and social development (Jiam et al., 2017); in fact, a correlation between impairments in perception and production of voice

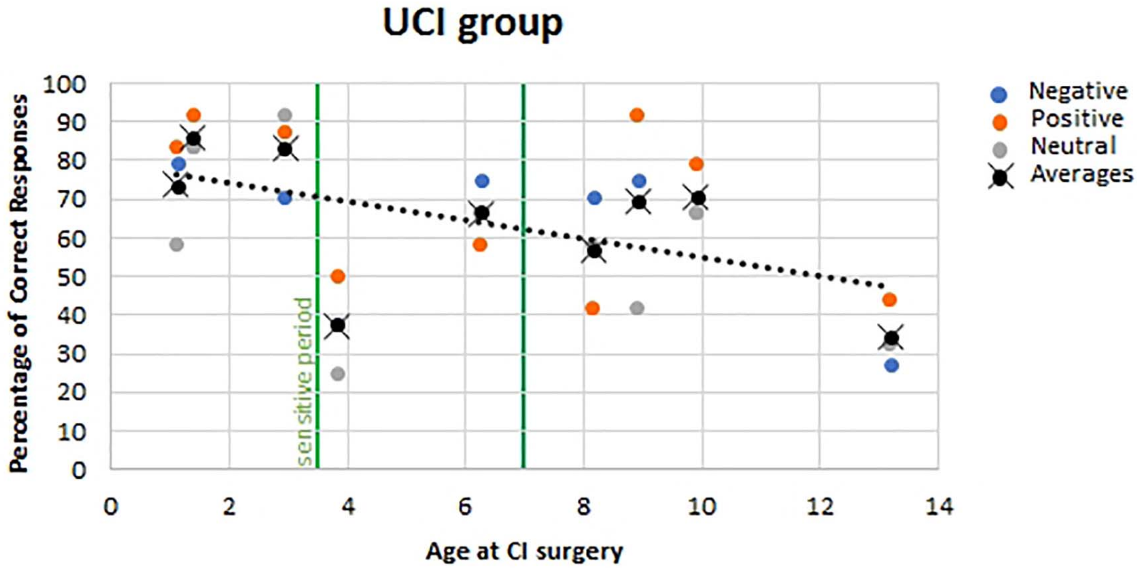

FIGURE 6 | Correlation between age at CI surgery and percentage of correct responses in UCI children. Orange dots stand for positive emotional states; blue dots stand for negative emotional states, and gray dots stand for neutral emotional states. Black dots stand for the mean of correct responses for each participant, irrespective of the emotional state. The vertical green lines represent the sensitive period threshold (3.5 and 7 years old) for the central auditory system development (Sharma et al., 2005).

emotion was found, like in the case of infant-directed speech, and in 5- to 13-year-old children who used CI (Nakata et al., 2012). It is interesting to note that a previous study employing vocal child-directed happy and sad speech stimuli reported higher performance in NH in comparison to CI using children; however, the percentage of recognition was higher than the one reported in the present study, probably due to the child-directed characteristic of the stimuli (Volkova et al., 2013).

Concerning the difference in gamma LI values observed in UCI in comparison to the NH group, it confirmed a difference in gamma band activity previously reported by Agrawal et al. (2013) in comparison between the same groups, therefore supporting the suitability of the study of gamma rhythms in the investigation of emotional messages conveyed by means of auditory stimuli. However, the previously mentioned study and the present study are not perfectly comparable because of the differences (i) in the sample – adults and children, respectively, – and therefore plausibly in the etiology of deafness; (ii) in the location of EEG activity acquisition, that is, Cz and multiple electrodes over the two hemispheres, respectively; and (iii) in the kind of emotional stimuli, that is, verbal stimuli pronounced with neutral, happy, and angry prosody in Agrawal and colleagues' study, while vocal nonverbal stimuli belonging to 10 emotions grouped into three emotional states in the present study. Moreover, the higher LI values reported for UCI in comparison to NH children would imply a more sustained conscious processing of the stimuli for the NH group in comparison to the UCI group and a higher processing of the emotional face stimuli – employed for the matching of the auditory stimuli for the identification of the target emotion – by the UCI group (Balconi and Lucchiari, 2008). In fact, McGurk studies showed a higher relying of UCI children on the visual sensation than on the auditory one in case of uncertainty (Schorr et al., 2005).

The correlation between higher right lateralization, as indexed by higher LI values, and the percentage of correct responses could be explained by the evidence of higher activation and asymmetry levels in poorer performers in emotion-in-voice recognition tasks than those of more proficient ones (Kislova and Rusalova, 2009). This possibly also reflects the poorer performance in emotion recognition obtained by UCI children, as well as their higher LI values in comparison to NH children. In fact, it was shown by studies on single hemisphere damage that although the right hemisphere is responsible for low-level discrimination and recognition of affective prosody, in case of higher task demands in terms of associational-cognitive requirements, the left hemisphere is engaged (Tompkins and Flowers, 1985). Thus, UCI children would present deficits in such engaging of the left hemisphere for more complex emotional processing tasks. This could be explained by the neuroimaging evidence that indeed areas appearing to be primarily involved in emotional prosodic processing, that is, posterior temporal (parietal) brain regions (Kotz et al., 2006), are the same areas presumably more involved by the neuroplastic changes that occurred after CI surgery (Giraud et al., 2000; Kang et al., 2004) and the following hearing sensation restoration.

The negative correlation between age of implantation and percentage of correct responses in emotion recognition is in accordance with previous studies (Mancini et al., 2016). On the contrary, in the Deroche and colleagues' study on adult CI users cited above, any effect of the age at implantation on the emotion recognition was not found, but this would be caused by the post-lingual acquisition of deafness in the majority of the sample (19 over 22 CI users) and by the type of emotions investigated, which is happy, sad, and neutral, whereas in the present study, 10 emotions were employed (Deroche et al., 2019). Furthermore, in Volkova et al.' (2013) study, employing child-directed emotional speech, performance of the children CI users was positively associated with duration of implant use. Such evidence could be compared to present results, given the almost overlap between age at CI surgery and length of CI use in the enrolled sample. In addition, the trend that better performances were obtained by children implanted before 3.5 years old suggests the influence of a sensitive period, identified through P1 cortical auditory-evoked potential trajectory post-CI development (Sharma et al., 2002, 2005; Sharma and Dorman, 2006; Kral and Sharma, 2012; Kral et al., 2019) also on emotion recognition skills development. Such phenomenon could be explained by the better auditory–visual integration achieved by children implanted before 3.5 years of age as shown by Miller's test of the race model inequality executed by early and late implanted children (Gilley et al., 2010). Such auditory–visual integration capability achievement is also witnessed by McGurk effect tests on CI children, showing that 38% of early implanted children – before the age of 2.5 years – but none of the late implanted children exhibited the bimodal fusion occurring in the McGurk effect, being instead biased toward the visual modality in contrast to the NH children who were biased toward the audio modality (Schorr et al., 2005). These evidences, with respect to the topic of emotion recognition skills development, are in accord to studies indicating that auditory and visual integration is necessary for the achievement of such capabilities (Campanella and Belin, 2007). In relation to this matter, there is also the evidence of a delay on facial emotion recognition in preschoolers using CI (and hearing aids) in comparison to NH mates, and interestingly, there was not any correlation between facial emotion recognition and language abilities (Wang et al., 2011). Differently, another study found a relation between better language skills and higher social competence, both in NH and CI children, although in the latter group, less adequate emotion-regulation strategies and less social competence than NH children were highlighted (Wiefferink et al., 2012). In addition, a study investigating both linguistic (recognition of monosyllabic words and of key words from sentences within background noise; repetition of non-words) and indexical (discrimination of across-gender and within-gender talkers; identification of emotional content from spoken sentences) properties in perceptual analysis of speech in CI children found an association between better performances in such feature recognition and a younger age at implantation (and use of more novel speech processor technology) (Geers et al., 2013).

Moreover, concerning the emotional communication, a suggestion of deficits also in the imitation of emotional (happy and sad) speech stimuli was found (Wang et al., 2013). Therefore, it sharply results in the vision and need of two targets of rehabilitation for children with CI that should be treated both conjointly and separately: language treatment and emotional intervention.

CONCLUSION

In light of the present results, in relation to the experimental questions previously declared, it is possible to conclude that (i) the processing of the emotional stimuli by deaf children using CI appears to be different from NH children, as suggested by the higher relative right hemisphere gamma band activity, possibly explained by the non-physiological development of the auditory system; (ii) on account of the inverse correlation between the age at the CI surgery and the percentage of correct responses, the precocity of performing the CI surgery for the attainment of best emotion recognition skills appears crucial, probably because of neuroplastic changes allowing a better processing and categorization of emotional stimuli; and (iii) the CI side does not appear to influence the processing of emotional stimuli, although interestingly the relative higher gamma band activity appears to be counterproductive in terms of emotion recognition performances; such aspect needs further investigation at the light of the possible particular implications of the right hemisphere hypothesis (Kotz et al., 2006).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Bambino Gesù Pediatric Hospital Ethics Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

GC conceived and conducted the study, performed the data analysis, and wrote the manuscript. AG and BI prepared the experimental protocol, conducted the study, and elaborated data. AS, SGi, AD, SGa, RG, CL, PL, and FF enrolled patients and organized experimental sessions. PM provided support for the organization and realization of the study. AS and FB edited the manuscript. FB supervised the entire experiment. All authors read and approved the final version of the article.

FUNDING

This research was funded by Cochlear Srl: IIR-1983.

ACKNOWLEDGMENTS

We acknowledge the contribution of participants and their families to the study, who kindly voluntarily granted their availability.

REFERENCES

Agrawal, D., Thorne, J. D., Viola, F. C., Timm, L., Debener, S., Büchner, A., et al. (2013). Electrophysiological responses to emotional prosody perception in cochlear implant users. *NeuroImage* 2, 229–238. doi: 10.1016/j.nicl.2013.01.001

Ahmed, D. G., Paquette, S., Zeitouni, A., and Lehmann, A. (2018). Neural processing of musical and vocal emotions through cochlear implants simulation. *Clin. EEG Neurosci.* 49, 143–151. doi: 10.1177/1550059417733386

Akhtar, N., and Gernsbacher, M. A. (2008). On privileging the role of gaze in infant social cognition. *Child Dev. Perspect.* 2, 59–65. doi: 10.1111/j.1750-8606.2008.00044.x

Aricò, P., Borghini, G., Di Flumeri, G., Sciaraffa, N., Colosimo, A., and Babiloni, F. (2017). Passive BCI in operational environments: insights, recent advances, and future trends. *IEEE Trans. Biomed. Eng.* 64, 1431–1436. doi: 10.1109/TBME.2017.2694856

Balconi, M., and Lucchiari, C. (2008). Consciousness and arousal effects on emotional face processing as revealed by brain oscillations. a gamma band analysis. *Int. J. Psychophysiol.* 67, 41–46. doi: 10.1016/j.ijpsycho.2007.10.002

Baldwin, D. A., and Moses, L. J. (1996). The ontogeny of social information gathering. *Child Dev.* 67, 1915–1939. doi: 10.1111/j.1467-8624.1996.tb01835.x

Borghini, G., Aricò, P., Di Flumeri, G., Cartocci, G., Colosimo, A., Bonelli, S., et al. (2017). EEG-Based cognitive control behaviour assessment: an ecological study with professional air traffic controllers. *Sci. Rep.* 7:547. doi: 10.1038/s41598-017-00633-7

Campanella, S., and Belin, P. (2007). Integrating face and voice in person perception. *Trends Cogn. Sci.* 11, 535–543. doi: 10.1016/j.tics.2007.10.001

Cartocci, G., Maglione, A. G., Rossi, D., Modica, E., Borghini, G., Malerba, P., et al. (2018). "Alpha and theta EEG variations as indices of listening effort to be implemented in neurofeedback among cochlear implant users," in *Symbiotic Interaction.* Symbiotic 2017. Lecture Notes in Computer Science, eds H. Jaap, A. Spagnolli, B. Blankertz, L. Gamberini, and G. Jacucci (Cham: Springer International Publishing), 30–41.

Cartocci, G., Maglione, A. G., Vecchiato, G., Di Flumeri, G., Colosimo, A., Scorpecci, A., et al. (2015). "Mental workload estimations in unilateral deafened children," in *Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, 1654–1657. doi: 10.1109/EMBC.2015.7318693

Cartocci, G., Scorpecci, A., Borghini, G., Maglione, A., Inguscio, B. M. S., Giannantonio, S., et al. (2019). EEG rhythms lateralization patterns in children with unilateral hearing loss are different from the patterns of normal hearing controls during speech-in-noise listening. *Hear. Res.* 379, 31–42. doi: 10.1016/j.heares.2019.04.011

Chatterjee, M., Zion, D., Deroche, M. L., Burianek, B., Limb, C., Goren, A., et al. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hear. Res.* 322, 151–162. doi: 10.1016/j.heares.2014.10.003

Damm, S. A., Sis, J. L., Kulkarni, A. M., and Chatterjee, M. (2019). How vocal emotions produced by children with cochlear implants are perceived by their hearing peers. *J. Speech Lang. Hear. Research* 62, 3728–3740. doi: 10.1044/2019_JSLHR-S-18-0497

Davies-Thompson, J., Elli, G. V., Rezk, M., Benetti, S., van Ackeren, M., and Collignon, O. (2019). Hierarchical brain network for face and voice integration of emotion expression. *Cereb. Cortex* 29, 3590–3605. doi: 10.1093/cercor/bhy240

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Deroche, M. L. D., Felezeu, M., Paquette, S., Zeitouni, A., and Lehmann, A. (2019). Neurophysiological differences in emotional processing by cochlear implant users, extending beyond the realm of speech. *Ear Hear.* 40, 1197–1209. doi: 10.1097/AUD.0000000000000701

Di Flumeri, G., Aricó, P., Borghini, G., Colosimo, A., and Babiloni, F. (2016). "A new regression-based method for the eye blinks artifacts correction in the EEG signal, without using any EOG channel," in *Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL, 3187–3190. doi: 10.1109/EMBC.2016.7591406

eHopyan, T., ePeretz, I., Chan, L. P., Papsin, B. C., and Gordon, K. A. (2012). Children using cochlear implants capitalize on acoustical hearing for music perception. *Front. Psychol.* 3:425. doi: 10.3389/fpsyg.2012.00425

Gainotti, G. (2019). Emotions and the right hemisphere: can new data clarify old models? *Neuroscientist* 25, 258–270. doi: 10.1177/1073858418785342

Geers, A. E., Davidson, L. S., Uchanski, R. M., and Nicholas, J. G. (2013). Interdependence of linguistic and indexical speech perception skills in school-age children with early cochlear implantation. *Ear Hear.* 34, 562–574. doi: 10.1097/AUD.0b013e31828d2bd6

Gfeller, K., Turner, C., Mehr, M., Woodworth, G., Fearn, R., Knutson, J. F., et al. (2002). Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implants Int.* 3, 29–53. doi: 10.1179/cim.2002.3.1.29

Gfeller, K., Turner, C., Oleson, J., Zhang, X., Gantz, B., Froman, R., et al. (2007). Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise. *Ear Hear.* 28, 412–423. doi: 10.1097/AUD.0b013e3180479318

Gilbers, S., Fuller, C., Gilbers, D., Broersma, M., Goudbeek, M., Free, R., et al. (2015). Normal-Hearing listeners' and cochlear implant users' perception of pitch cues in emotional speech. *IPerception* 6, 0301006615599139. doi: 10.1177/0301006615599139

Gilley, P. M., Sharma, A., Mitchell, T. V., and Dorman, M. F. (2010). The influence of a sensitive period for auditory-visual integration in children with cochlear implants. *Restor. Neurol. Neurosci.* 28, 207–218. doi: 10.3233/RNN-2010-0525

Giraud, A. L., Truy, E., Frackowiak, R. S., Grégoire, M. C., Pujol, J. F., and Collet, L. (2000). Differential recruitment of the speech processing system in healthy subjects and rehabilitated cochlear implant patients. *Brain* 123(Pt 7), 1391–1402. doi: 10.1093/brain/123.7.1391

Gordon, K., Henkin, Y., and Kral, A. (2015). Asymmetric hearing during development: the aural preference syndrome and treatment options. *Pediatrics* 136, 141–153. doi: 10.1542/peds.2014-3520

Gordon, K. A., and Papsin, B. C. (2019). Special edition on unilateral deafness and hearing loss: an introduction and overview. *Hear. Res.* 372, 1–2. doi: 10.1016/j.heares.2018.12.007

Grossmann, T. (2010). The development of emotion perception in face and voice during infancy. *Restor. Neurol. Neurosci.* 28, 219–236. doi: 10.3233/RNN-2010-0499

Hopyan-Misakyan, T. M., Gordon, K. A., Dennis, M., and Papsin, B. C. (2009). Recognition of affective speech prosody and facial affect in deaf children with unilateral right cochlear implants. *Child Neuropsychol.* 15, 136–146. doi: 10.1080/09297040802403682

Jiam, N. T., Caldwell, M., Deroche, M. L., Chatterjee, M., and Limb, C. J. (2017). Voice emotion perception and production in cochlear implant users. *Hear. Res.* 352, 30–39. doi: 10.1016/j.heares.2017.01.006

Kang, E., Lee, D. S., Kang, H., Lee, J. S., Oh, S. H., Lee, M. C., et al. (2004). Neural changes associated with speech learning in deaf children following cochlear implantation. *NeuroImage* 22, 1173–1181. doi: 10.1016/j.neuroimage.2004.02.036

Kislova, O. O., and Rusalova, M. N. (2009). EEG asymmetry in humans: relationship with success in recognizing emotions in the voice. *Neurosci. Behav. Physiol.* 39, 825–831. doi: 10.1007/s11055-009-9213-8

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Brain Res. Rev.* 29, 169–195.

Kotz, S. A., Meyer, M., and Paulmann, S. (2006). "Lateralization of emotional prosody in the brain: an overview and synopsis on the impact of study design," in *Progress in Brain Research. Understanding Emotions*, Vol. 156, eds S. Anders, G. Ende, M. Junghofer, J. Kissler, and D. Wildgruber (Amsterdam: Elsevier), 285–294. doi: 10.1016/S0079-6123(06)56015-7

Kral, A., Dorman, M. F., and Wilson, B. S. (2019). Neuronal development of hearing and language: cochlear implants and critical periods. *Annu. Rev. Neurosci.* 42, 47–65. doi: 10.1146/annurev-neuro-080317-061513

Kral, A., Hartmann, R., Tillein, J., Heid, S., and Klinke, R. (2001). Delayed maturation and sensitive periods in the auditory cortex. *Audiol. Neurotol.* 6, 346–362. doi: 10.1159/000046845

Kral, A., Hubka, P., Heid, S., and Tillein, J. (2013). Single-Sided deafness leads to unilateral aural preference within an early sensitive period. *Brain* 136(Pt 1), 180–193. doi: 10.1093/brain/aws305

Kral, A., and Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends Neurosci.* 35, 111–122. doi: 10.1016/j.tins.2011.09.004

Li, M., and Lu, B.-L. (2009). "Emotion classification based on gamma-band EEG," in *Proceedigs of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Minneapolis, MN, 1223–1226. doi: 10.1109/IEMBS.2009.5334139

Mancini, P., Giallini, I., Prosperini, L., D'alessandro, H. D., Guerzoni, L., Murri, A., et al. (2016). Level of emotion comprehension in children with mid to long term cochlear implant use: how basic and more complex emotion recognition relates to language and age at implantation. *Int. J. Pediatr. Otorhinolaryngol.* 87, 219–232. doi: 10.1016/j.ijporl.2016.06.033

Marsella, P., Scorpecci, A., Cartocci, G., Giannantonio, S., Maglione, A. G., Venuti, I., et al. (2017). EEG activity as an objective measure of cognitive load during effortful listening: a study on pediatric subjects with bilateral, asymmetric sensorineural hearing loss. *Int. J. Pediatr. Otorhinolaryngol.* 99, 1–7. doi: 10.1016/j.ijporl.2017.05.006

Mildner, V., and Koska, T. (2014). Recognition and production of emotions in children with cochlear implants. *Clin. Linguist. Phon.* 28, 543–554. doi: 10.3109/02699206.2014.927000

Most, T., and Aviner, C. (2009). Auditory, visual, and auditory-visual perception of emotions by individuals with cochlear implants, hearing AIDS, and normal hearing. *J. Deaf Stud. Deaf Educ.* 14, 449–464. doi: 10.1093/deafed/enp007

Most, T., and Michaelis, H. (2012). Auditory, visual, and auditory-visual perceptions of emotions by young children with hearing loss versus children with normal hearing. *J. Speech Lang. Hear. Res. JSLHR* 55, 1148–1162. doi: 10.1044/1092-4388(2011/11-0060)

Nakata, T., Trehub, S. E., and Kanda, Y. (2012). Effect of cochlear implants on children's perception and production of speech prosody. *J. Acoust. Soc. Am.* 131, 1307–1314. doi: 10.1121/1.3672697

Nowicki, S., and Duke, M. P. (1994). Individual differences in the nonverbal communication of affect: the diagnostic analysis of nonverbal accuracy scale. *J. Nonverbal Behav.* 18, 9–35. doi: 10.1007/BF02169077

Orsini, A., Grossi, D., Capitani, E., Laiacona, M., Papagno, C., and Vallar, G. (1987). Verbal and spatial immediate memory span: normative data from 1355 adults and 1112 children. *Ital. J. Neurol. Sci.* 8, 539–548. doi: 10.1007/BF02333660

Paquette, S., Ahmed, G. D., Goffi-Gomez, M. V., Hoshino, A. C. H., Peretz, I., and Lehmann, A. (2018). Musical and vocal emotion perception for cochlear implants users. *Hear. Res.* 370, 272–282. doi: 10.1016/j.heares.2018.08.009

Piccioni, L. O., Cartocci, G., Maglione, A. G., Modica, E., Rossi, D., Mancini, M., et al. (2018). EEG variations as estimators of listening effort during recognition of words in noise in unilateral and bilateral sequential adult cochlear implant users. *J. Hear. Sci.* 8, 116–116.

Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., and Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proc. Natl. Acad. Sci. U.S.A.* 104, 7295–7300. doi: 10.1073/pnas.0609419104

Sauter, D. A., Crasborn, O., Engels, T., Kamiloğlu, R. G., Sun, R., Eisner, F., et al. (2019). Human emotional vocalizations can develop in the absence of auditory learning. *Emotion* 20, 1435–1445.

Sauter, D. A., Eisner, F., Calder, A. J., and Scott, S. K. (2006). Perceptual cues in non-verbal vocal expressions of emotion. *Q. J. Exp. Psychol. (2006)* 63, 2251–2272. doi: 10.1080/17470211003721642

Sauter, D. A., Eisner, F., Ekman, P., and Scott, S. K. (2010). Cross-Cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2408–2412. doi: 10.1073/pnas.0908239106

Sauter, D. A., Panattoni, C., and Happé, F. (2013). Children's recognition of emotions from vocal cues. *Br. J. Dev. Psychol.* 31, 97–113. doi: 10.1111/j.2044-835X.2012.02081.x

Scherer, K. R. (2003). Vocal communication of emotion: a review of research paradigms. *Speech Commun.* 40, 227–256. doi: 10.1016/S0167-6393(02)00084-5

Schorr, E. (2005). *Social and Emotional Functioning of Children with Cochlear Implants*. Available online at: http://drum.lib.umd.edu/handle/1903/2408 (accessed December 23, 2020).

Schorr, E. A., Fox, N. A., van Wassenhove, V., and Knudsen, E. I. (2005). Auditory-Visual fusion in speech perception in children with cochlear implants. *Proc. Natl. Acad. Sci. U.S.A.* 102, 18748–18750. doi: 10.1073/pnas.0508862102

Sharma, A., and Dorman, M. F. (2006). Central auditory development in children with cochlear implants: clinical implications. *Adv. Otorhinolaryngol.* 64, 66–88. doi: 10.1159/000094646

Sharma, A., Dorman, M. F., and Kral, A. (2005). The influence of a sensitive period on central auditory development in children with unilateral and bilateral cochlear implants. *Hear. Res.* 203, 134–143. doi: 10.1016/j.heares.2004.12.010

Sharma, A., Dorman, M. F., and Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear Hear.* 23, 532–539.

Strelnikov, K., Rouger, J., Demonet, J.-F., Lagleyre, S., Fraysse, B., Deguine, O., et al. (2013). Visual activity predicts auditory recovery from deafness after adult cochlear implantation. *Brain* 136, 3682–3695. doi: 10.1093/brain/awt274

The Ear Foundation (2017). *Cochlear Implant Information Sheet*. (accessed October 14, 2019)

Tompkins, C. A., and Flowers, C. R. (1985). Perception of emotional intonation by brain-damaged adults: the influence of task processing levels. *J. Speech Hear. Res.* 28, 527–538. doi: 10.1044/jshr.2804.527

Vanvooren, S., Hofmann, M., Poelmans, H., Ghesquière, P., and Wouters, J. (2015). Theta, beta and gamma rate modulations in the developing auditory system. *Hear. Res.* 327, 153–162. doi: 10.1016/j.heares.2015.06.011

Volkova, A., Trehub, S. E., Schellenberg, E. G., Papsin, B. C., and Gordon, K. A. (2013). Children with bilateral cochlear implants identify emotion in speech and music. *Cochlear Implants Int.* 14, 80–91. doi: 10.1179/1754762812Y.0000000004

Wang, D. J., Trehub, S. E., Volkova, A., and van Lieshout, P. (2013). Child implant users' imitation of happy- and sad-sounding speech. *Front. Psychol.* 4:351. doi: 10.3389/fpsyg.2013.00351

Wang, Y., Su, Y., Fang, P., and Zhou, Q. (2011). Facial expression recognition: can preschoolers with cochlear implants and hearing aids catch it? *Res. Dev. Disabil.* 32, 2583–2588. doi: 10.1016/j.ridd.2011.06.019

Wiefferink, C. H., Rieffe, C., Ketelaar, L., and Frijns, J. H. M. (2012). Predicting social functioning in children with a cochlear implant and in normal-hearing children: the role of emotion regulation. *Int. J. Pediatr. Otorhinolaryngol.* 76, 883–889. doi: 10.1016/j.ijporl.2012.02.065

Wiefferink, C. H., Rieffe, C., Ketelaar, L., Raeve, L. De, and Frijns, J. H. M. (2013). Emotion understanding in deaf children with a cochlear implant. *J. Deaf Stud. Deaf Educ.* 18, 175–186. doi: 10.1093/deafed/ens042

Yang, K., Tong, L., Shu, J., Zhuang, N., Yan, B., and Zeng, Y. (2020). High gamma band EEG closely related to emotion: evidence from functional network. *Front. Hum. Neurosci.* 14:89. doi: 10.3389/fnhum.2020.00089

Young, A. W., Frühholz, S., and Schweinberger, S. R. (2020). Face and voice perception: understanding commonalities and differences. *Trends Cogn. Sci.* 24, 398–410. doi: 10.1016/j.tics.2020.02.001

Zupan, B., and Sussman, J. E. (2009). Auditory preferences of young children with and without hearing loss for meaningful auditory–visual compound stimuli. *J. Commun. Disord.* 42, 381–396. doi: 10.1016/j.jcomdis.2009.04.002

Conflict of Interest: GC, AG, BI, and FB were employed by BrainSigns Srl. PM was employed by Cochlear Italia Srl.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

REVIEW
published: 21 March 2022
doi: 10.3389/fnins.2022.824389

Research Status and Future Development of Cochlear Reimplantation

Xinyi Yao[1†], Haotian Liu[1†], Jinyuan Si[2], Xiuyong Ding[2], Yu Zhao[1*‡] and Yun Zheng[1*‡]

[1] Department of Otolaryngology-Head and Neck Surgery, West China Hospital of Sichuan University, Chengdu, China,
[2] Department of Otolaryngology-Head and Neck Surgery, XuanWu Hospital of Capital Medical University, Beijing, China

OPEN ACCESS

Edited by:
Xin Luo,
Arizona State University, United States

Reviewed by:
Chun Liang,
Shenzhen Maternity and Child Healthcare Hospital, China
Amit Keshri,
Sanjay Gandhi Post Graduate Institute of Medical Sciences (SGPGI), India
Evelyn E. M. O. Davies-Venn,
University of Minnesota Twin Cities, United States

***Correspondence:**
Yun Zheng
1141679315@qq.com
Yu Zhao
yutzhao@163.com

†These authors have contributed equally to this work and share first authorship

‡These authors have contributed equally to this work

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: 29 November 2021
Accepted: 14 February 2022
Published: 21 March 2022

Citation:
Yao X, Liu H, Si J, Ding X, Zhao Y and Zheng Y (2022) Research Status and Future Development of Cochlear Reimplantation. Front. Neurosci. 16:824389. doi: 10.3389/fnins.2022.824389

Cochlear implants are the most successful sensory prostheses worldwide, and they can be useful for patients with severe and profound hearing impairment. However, various complications, including infection, pain, and device failure which is mainly due to falls and trauma, are associated with the use of cochlear implants. Reimplantation is required to replace the initial device in severe complications. Nevertheless, reimplantation can present certain surgical risks and may impose a significant economic and psychological burden on patients and their families; therefore, it requires greater attention and focus. This article presents a review of the literature on cochlear reimplantation and summarizes the current status, knowledge gaps, and future research directions on cochlear reimplantation. Since 1980s, cochlear reimplantation techniques can be considered to be relatively mature; however, some clinical and scientific problems remain unresolved, including the lack of a unified definition of cochlear reimplantation, non-standardized calculation of the reimplantation rat, and insufficient effect assessment. This review highlights the urgent need to establish an international consensus statement on cochlear reimplantation research to standardize the definition, calculation formulas of reimplantation rate, and follow-up systems.

Keywords: cochlear implants, reimplantation, literature review, revision surgery, reimplantation rate

INTRODUCTION

Cochlear implants (CIs), the most successful sensory prostheses worldwide, are small but complex electronic devices implanted in the cochlea that can help patients with severe or profound hearing impairment hear external sound (Büchner and Gärtner, 2017). Since the 1970s, continuous advancements have been improved the speech-coding strategies, electrodes, and materials used for developing CIs (Büchner and Gärtner, 2017; Carlyon and Goehring, 2021). The indications for CIs have also expanded gradually, and the number of patients with CIs has been increasing (Lailach et al., 2021).

Cochlear reimplantation has also attracted increasing public attention because of the various complications associated with the use of CIs, including skin flap infection, electrodes migration, and device failure which are mainly caused by falls and trauma (Weise et al., 2005; Terry et al., 2015; Dağkıran et al., 2020). Cochlear reimplantation surgery is an invasive operation, and CIs are very expensive. Thus, reimplantation may greatly increase the economic and psychological burden on patients and families, especially in those living in developing countries where CIs are not covered by medical insurance (Sorkin and Buchman, 2016; Qiu et al., 2017; Chen et al., 2021).

Moreover, some patients showed poorer hearing and speech recognition levels after reimplantation than those before initial implantation (Chung et al., 2010; Balakina et al., 2015; Manrique-Huarte et al., 2015), indicating the need for greater attention to this population. Nevertheless, only a few studies have summarized the current status of research on cochlear reimplantation, and there are no studies or reviews on the future research directions of cochlear reimplantation. This review aims to summarize the history and current state of knowledge of cochlear reimplantation, enumerate the research gaps and suggest directions for future research.

PART I: HISTORY AND CURRENT STATUS OF COCHLEAR REIMPLANTATION

In Hochmair-Desoyer and Burian (1985) first reported the use of cochlear reimplantation in two adults with post-lingual deafness in Austria (Desoyer and Burian, 1985). The effect of cochlear reimplantation was assessed using Bekesy threshold tracking, loudness scaling, and open lists of unknown single-syllable words and daily sentences, among other techniques. The patients showed no significant changes in hearing thresholds or speech recognition between pre- to post-reimplantation. In Burian and Eisenwort (1989) studied four prelingually deaf children who underwent reimplantation and found that speech discrimination levels remained the same or even improved after reimplantation. In the same year, Spillmann and Dillier (1989) reported a case of reimplantation for device failure, in which a single-channel CI was upgraded to a multichannel CI. The device upgrade resulted in a significant improvement in the auditory performance of the patient. These early studies provided preliminary evidence confirming the satisfactory effects of reimplantation.

Two CI surgeries were performed on the same ear of eight adult cats (Robert et al., 1989). They found that the initial electrodes could be removed easily and that new electrodes could be implanted successfully without damaging the peripheral cochlear nerve; however, the proliferation of granulation tissue in round windows and scala tympani may lead to difficulties in implanting new electrodes. In addition, Jackler et al. (1989) was the first to assess manufacturer-specific cochlear reimplantation rates, i.e., Cochlear Corp., 2.8%; Richards Corp., 3.3%; Storz Corp., 7%. In Webb et al. (1991) who collated cochlear implantation complication data from Germany, Australia, and the United States, reported the number of reimplanted CIs. In Maas et al. (1996) first proposed a new reimplantation strategy, in which the contralateral ear was chosen for reimplantation, and the postoperative performance on speech recognition tests remained the same or improved.

Since 2000, CI technology has been advanced by successive multicenter cohort studies. For example, Parisier et al. (2001) reviewed the operations and postoperative findings for initial implantation and reimplantation in 25 children and reported the postoperative complications of reimplantation, which demonstrated that reimplantation in children is feasible and effective. From 2005 to 2018, an increasing number of new and special cases of reimplantation have been reported, such as reimplantation for various flap infections, intractable facial nerve stimulation, and reimplantation in patients with inner ear malformations (Puri et al., 2005; Polak et al., 2006; Ahn et al., 2008; Incesulu et al., 2008). In addition, long-term large-scale cohort studies have been undertaken. In a study of 18 hospitals across the United States, Henson et al. tracked and collected the information on 22 patients who had undergone reimplantation and found that approximately 60% of the patients had better or similar auditory outcomes after reimplantation than after the initial implantation (Henson et al., 1999). Weder et al. (2020) conducted the longest study, to date, from 1982 to 2018, in which a tertiary referral hospital performed 4,600 initial cochlear implantations and 22 reimplantations due to infection. They found that speech recognition after reimplantation was comparable to that before reimplantation. However, the reimplantation rate reported by Weder's study was only 0.48%, which does not represent the overall reimplantation situation, as patients who underwent reimplantation for device failure and other reasons were not included. The largest sample size was reported in a multicenter study conducted by Hermann et al. (2020) who described the findings for 4,952 initial cochlear implantations and 99 reimplantations, with a 2% reimplantation rate. Of the 99 cases of reimplantation, only one had postoperative infection after reimplantation.

The number and quality of CIs in China have been gradually increasing since the first adult multichannel cochlear implantation was performed at the Peking Union Medical College Hospital in 1995 (Han, 2004; Xi et al., 2005). In 2010, more than 10,000 CIs were implanted in China, leading the rest of the world in terms of the number of procedures performed (Cao and Wei, 2010). In 2018, the total number of children with CIs exceeded 50,000 (Lu, 2018). Meanwhile, the benefits of CIs to Chinese speech recognition of implanted patients have been widely reported (Luo et al., 2008; Chang et al., 2016). The reimplantation surgery in China was first reported in a study of six patients, with a reimplantation rate of 16.67% (Yu et al., 2004). In Zhao et al. (2008) evaluated the effects of reimplantation from the perspective of effective working electrodes. To date, more than 40 studies on reimplantation in China have been published in Chinese, and six have been published in English. These studies have primarily focused on cause analysis, surgical discovery, and the management of complications.

PART II: CURRENT DEFICIENCIES AND FUTURE RESEARCH DIRECTIONS

Non-standard Definitions of Cochlear Reimplantation

Currently, the primary concern in cochlear reimplantation is the lack of an internationally unified and normative definition.

In Jackler et al. (1989) defined cochlear reimplantation as the removal of an indwelling CI electrodes followed by reinsertion of a new device. They noted that reimplantation is a maneuver of uncertain consequences to the cochlea and its surviving nerve (Robert et al., 1989).

However, according to a literature review, most researchers have not clearly defined cochlear reimplantation (Orús et al., 2010; Ciorba et al., 2012). Some articles did not attempt to define cochlear reimplantation at all (Bhadania et al., 2018; Batuk et al., 2019). "Reimplantation/re-implantation" are often used to mean CI reimplantation, and "reinsertion/replacement" are also used in the same context (Desoyer and Burian, 1985; Parisier et al., 1991; Holcomb et al., 2018; Lane et al., 2019). However, there existed confusion and genericity between "revision" and "reimplantation" in many studies (Lassig et al., 2005; Rivas et al., 2008; Hwang et al., 2019). For example, Marlowe et al. proposed that revision surgery is a means to deal with abnormal implantation sites or internal problems with the implanted device (Marlowe et al., 2010). Revision surgery has been defined as the removal of the old device and replacement with a prosthesis, which is similar to the definition of cochlear reimplantation.

However, the procedures performed to treat CI complications were classified as follows: (1) cochlear reimplantation, (2) other revision surgery, and (3) medical treatment (Lescanne et al., 2011; Tarkan et al., 2013). Therefore, we speculate that cochlear reimplantation is characterized by replacing the initial electrodes with brand-new devices, and that revision surgery comprises surgical operations performed to address CI complications. Thus, cochlear reimplantation is a part of revision surgery.

Overall, the first requirement for future studies is to formulate a scientific and rigorous definition of reimplantation to standardize relevant studies and enhance comparability. In this regard, it is imperative to establish a global committee comprising cochlear implant manufacturers, FDA authorities, and clinicians/academicians from a variety of settings to propose an international consensus on cochlear reimplantation to standardize its definition.

Unclear Range of Cochlear Reimplantation

The second important issue is the unclear range of cochlear reimplantation. In Jackler et al. (1989) were the first to systematically summarize the following reasons for reimplantation: device failure, flap infections, electrodes mis-insertion or compression, hematoma or trauma at the receiver site, and accidental displacement (Robert et al., 1989). The European Consensus Statement published in 2005 classified the reasons as device failure, medical reason, characteristic decrement, and performance decrement (No authors listed, 2005). Device failure can be divided into hard and soft failures based on whether the failure can be proven with *in vivo* tests (Balkany et al., 2005). This was the first unified document regarding cochlear reimplantation. However, the 2005 consensus did not clearly propose a definition and scope of cochlear reimplantation and lacked procedures and tools for screening and the classification of the reasons.

In Battmer et al. (2010) published the International Classification of Reliability for Implanted Cochlear Implant Receiver Stimulators, which adopted a similar framework as the 2005 European Consensus. This framework illustrated some details, such as the CI survival time, (reduced) clinical benefit, and specification (Battmer et al., 2010). In 2017, the Association for the Advancement of Medical Instrumentation (AAMI) updated and standardized the classification of explanted CIs (Zwolan and Verhof, 2017) and proposed the following four categories: medical reasons, non-medical reasons, inconclusive/no faults found, and combined reasons. Most studies have adopted these three classification categories (Hermann et al., 2020; Layfield et al., 2021). However, there is no unified research range for cochlear reimplantation.

According to a literature review, most studies regarded the following situations as reimplantation: (1) removal of the initial electrodes for various reasons and implantation of a new device on the ipsilateral or contralateral side, which was the most common scenario defined as reimplantation (Lassig et al., 2005); (2) replacement of the failed hybrid CI with short electroacoustic stimulation (EAS) electrodes or full-length electrodes (Kamat et al., 2011; Jayawardena et al., 2012; Dunn et al., 2015); (3) reinsertion of the initial electrodes into the cochlea in cases of device migration (Luo et al., 2020); (4) reinsertion of the initial electrodes into the cochlea on the operation day or within a few days of the operation due to the electrodes in incorrect insertion places, such as the internal auditory canal, eustachian tube, vestibul. (Gözen et al., 2019); and (5) simultaneous implantation on both sides when patients with unilateral CI accept reimplantation surgery (Tang et al., 2019).

However, these classifications are not entirely appropriate. We consider that the first kind falls in the range of cochlear reimplantation, which replaced the initial device with a brand new one. And the second one can be regarded as a sort of reimplantation as well, which is relatively rare. For the third and fourth definitions, no new devices were inserted into the cochlea; thus, they cannot be regarded as reimplantation. For the fifth definition, ipsilateral implantation with a new device can be considered as reimplantation; however, contralateral implantation was the first CI, not the reimplanted CI. In conclusion, future studies should propose a more precise scope and classification of cochlear reimplantation.

Non-uniform Calculation of Cochlear Reimplantation Rate

The third important issue is the calculation formula of CI reimplantation rate, which varies greatly across studies. In Battmer et al. (2010) proposed the definition and calculation of the cumulative survival rate (CSR), which was in accordance with the methodology outlined in ISO standard 5841-2:2000 and targeted device failure without accounting for the medical reasons. In 2017, AAMI published the definition and calculation of the cumulative removal percentage (CRP), which covered all explanted CIs (Zwolan and Verhof, 2017). However, the formulation of CRP is not applicable to cochlear reimplantation rate. Cochlear removal mainly refers to the explantation of existing devices due to device failure, medical reasons, and inconclusive defects. In contrast, cochlear reimplantation is defined as the explantation of the initial device, followed by implantation of new electrodes. However, not all patients undergoing explantation of ordinary devices received new devices, and in some cases, the failed device was left *in situ*, while the new device was implanted on the contralateral side. These issues highlight the importance of standardizing the calculation formulation for the CI rate based on the CRP.

According to a literature review, some studies have used the number of patients as the unit of measurement, regardless of whether they had undergone unilateral and bilateral implantation (Sterkers et al., 2015), Dotú et al. (2010), whereas others have used the CI number as the unit of measurement Googe and Carron (2016); Karamert et al. (2019). Regarding the formula used for calculation, some studies adopted either of the following formula: "rate = no. of reimplanted CIs/no. of total CIs" or "rate = no. of reimplanted CIs/no. of primary CIs," in which primary CIs refer to devices inserted into the cochlea for the first time. Thus, the reimplantation rate may be different even for the same batch of patients.

Based on published studies, the CI reimplantation rate ranged from 0.5 to 30% (Beadle et al., 2005; Qiu et al., 2010), and a few studies have reported reimplantation rates higher than 20%. This high reimplantation rate may be attributed to the early publishing time in which the CI surgeries and devices were immature, small sample size, data bias, and immature surgical technology (Hamzavi et al., 2003; Beadle et al., 2005; Kanchanalarp et al., 2005). Patients of different ages with CIs showed different reimplantation rates. The reimplantation rate in children ranged from 0.7 to 30.0% (Kanchanalarp et al., 2005; Sun, 2019), whereas that in adults ranged from 0.4 to 27.3% (Hamzavi et al., 2003; Dağkıran et al., 2020). Some studies have reported that the reimplantation rate in children was significantly higher than that in adults (Sunde et al., 2013; Dağkıran et al., 2020). Children are prone to falls, resulting in head trauma during rapid growth and development (Weise et al., 2005; Wang et al., 2014). The skull and mastoid of children are immature, and rapid growth of the skull can lead to electrode array migration. In addition, the high prevalence of various types of otitis media in children increases the risk of CI failure (Manrique-Huarte et al., 2015). However, some studies have demonstrated no significant differences in the revision rates due to infection complications and device failure rates between adults and children (Sunde et al., 2013; Distinguin et al., 2017). The number of studies that separately calculated the reimplantation rates for children and adults is relatively low; therefore, larger longitudinal cohort studies are required.

Effect Assessment of Cochlear Reimplantation

The fourth important issue is that the current research mainly focuses on cause analysis and treatment complications. Thus, studies on postoperative effect assessment and the related methods remain limited. Hochmair-Desoyer and Burian (1985) used Bekesy threshold tracking, loudness scaling, open lists of unknown single-syllable words and everyday sentences, and other tests to evaluate the effects of reimplantation in their initial report (Desoyer and Burian, 1985). Since then, almost half of the studies related to reimplantation have mentioned effect evaluations using assessments, such as the Bamford–Kowal–Bench (BKB) test, phonetically balanced kindergarten (PBK) test, Northwestern University number 6 (NU#6) test, common phrases test (CPT), lexical neighborhood test (LNT), categories of auditory performance (CAP), and the speech intelligibility rating (SIR) (Saeed et al., 1995; Beadle et al., 2005; Marlowe et al., 2010; Bhadania et al., 2018). However, these studies have primarily focused on hearing thresholds and speech recognition, and no new assessment methods for neuro-electrophysiological monitoring and evaluation of neurofunctional characteristics have been proposed.

Nevertheless, some new technologies have been demonstrated to be applicable to patients with hearing loss or CIs, such as functional near-infrared optical brain imaging (fNIRS) and electroencephalography (EEG). In Sevy et al. (2010) used NIRS and fMRI to examine the cortical activity in response to auditory stimuli in five children with CIs and five children with normal hearing in Boston Children's Hospital. In Wang et al. (2020) adopted EEG to evaluate the effect of CIs at three time points after surgery and showed that multiple EEG indices could be used to assess speech perception ability. Many studies have demonstrated that neurological imaging is a safe and feasible approach for the examination of children with CIs, and that it could be an effective method for assessing the effects of reimplanted CIs. Therefore, if these advanced tools could be adopted for the general evaluation of both the first and reimplanted CIs, a better comparison and evaluation of the effect and the regularity of the rehabilitation of cochlear reimplantation can be achieved. In addition, results can be compared between patients with reimplanted CIs and the normal-hearing population, and differences can be used to evaluate the rehabilitation effect.

Regarding the actual effects of reimplantation, Chung et al. (2010) reported that the pure tone audiometry (PTA) and speech recognition scores of all patients who underwent reimplantation for soft failure decreased after reimplantation. Manrique-Huarte et al. (2015) used aided free-field auditory tests. Aided PTA findings after reimplantation improved in 44.4% of patients, deteriorated in 44.4%, and showed no significant difference in 11.1%, whereas speech recognition scores improved in 63.6%, showed no significant change in 9.1%, and worsened in 27.3%. In a study on the performance of 56 children 18 months after cochlear reimplantation, 87% showed better speech perception, 10% reported similar results, and 3% showed worse speech perception after surgery (Marlowe et al., 2010). Younger children were more likely to achieve or exceed their previous best performance than older children. The age at reimplantation, interval between the initial implantation and reimplantation, auditory input during the interval, depth of the electrode array, device activation, and device upgrade also influenced the postoperative effects (Marlowe et al., 2010; Lenarz, 2017; Roßberg et al., 2021). However, no single study has systematically discussed these factors. In addition, only a few effect evaluations have been carried out, with effect evaluation analysis not performed in some large-scale studies. Therefore, future studies should utilize more advanced assessment tools, such as functional near-infrared spectroscopy and electroencephalography, and should take more factors into consideration to systematically assess the rehabilitation effect and explore the internal patterns.

Study Scale and the Influence of the Reimplantation Side

Most studies on cochlear reimplantation were single-center studies with a relatively small sample size, and multicenter studies only accounted for 6% (Chung et al., 2010; Hermann et al., 2020). Therefore, the use of rigorous and advanced statistical analysis methods to obtain more robust findings is difficult. Moreover, in the absence of a standard criterion for the selection of the reimplantation side, most studies selected the side based on infection, cochlear ossification, and deformity (Chung et al., 2010; Lu and Cao, 2014; Manrique-Huarte et al., 2015), without considering other influencing factors, such as the duration between the first implantation and reimplantation and continuous auditory input during the period. Thus, integration of the patient resources of several hospitals and multicenter studies should be performed to elevate the level of evidence on this topic.

CONCLUSION

As the use of CIs continues to increase worldwide and the service life of early implanted devices approaches its end, the number of reimplanted CIs has increased in recent years. However, CI reimplantation is associated with some limitations, such as non-standard definitions and calculation formula for the reimplantation rate and the absence of high-quality studies on rehabilitation effect. Thus, establishment of a standard definition and appropriate scope for future studies is important. Longitudinal and multicenter studies should be conducted using more advanced tools and after adjusting for more covariates to systematically assess the effects of reimplantation and develop an effective system for follow-up and evaluation.

AUTHOR CONTRIBUTIONS

XY and HL made substantial contributions to the study conception and design, literature research, and drafting and revision of the manuscript. JS provided critical insights and points regarding the conception of cochlear reimplantation and performed draft revision. XD provided substantial suggestions for the definition and status of cochlear reimplantation and critically revised the manuscript for important intellectual content. YZa made substantial contributions to this study, including raising on-the-mark suggestions to the definition and status of cochlear reimplantation, sharing the surgery details of cochlear implantation and reimplantation, and revising the manuscript critically for important intellectual content. YZe made substantial contributions to the conception and design of this study, especially in the effect assessment of cochlear reimplantation, and was involved in critically revising the manuscript for important intellectual content.

ACKNOWLEDGMENTS

We sincerely appreciate the conclusions of CI surgery from the Department of Otolaryngology-Head and Neck Surgery, West China Hospital of Sichuan University. We are also so grateful for the useful advice of the Department of Otolaryngology-Head and Neck Surgery, XuanWu Hospital of the Capital Medical University.

REFERENCES

Ahn, J. H., Chung, J. W., and Lee, K. S. (2008). Complications following cochlear implantation in patients with anomalous inner ears: experiences in Asan Medical Center. *Acta Otolaryngol.* 128, 38–42. doi: 10.1080/00016480701361988

Balakina, A., Litvak, M., and Starokha, A. (2015). Cochlear reimplantation: audiological outcomes and assessment of quality of life. *J. Int. Adv. Otol.* 23, 345–355. doi: 10.1159/000496191

Balkany, T. J., Hodges, A. V., Buchman, C. A., Luxford, W. M., Pillsbury, C. H., Roland, P. S., et al. (2005). Cochlear implant soft failures consensus development conference statement. *Otol. Neurotol.* 26, 815–818. doi: 10.1097/01.mao.0000178150.44505.52

Battmer, R., Backous, D. D., Balkany, T. J., Briggs, R. J. S., Gantz, B. J., van Hasselt, A., et al. (2010). International classification of reliability for implanted cochlear implant receiver stimulators. *Otol. Neurotol.* 31, 1190–1193. doi: 10.1097/MAO.0b013e3181d2798e

Batuk, M. O., Cinar, B. C., Yarali, M., Bajin, M. D., Sennaroglu, G., and Sennaroglu, L. (2019). Twenty years of experience in revision cochlear implant surgery: signs that indicate the need for revision surgery to audiologists. *J. Laryngol. Otol.* 133, 903–907. doi: 10.1017/S0022215119001919

Beadle, E. A., Mckinley, D. J., Nikolopoulos, T. P., Brough, J., O'Donoghue, G. M., and Archbold, S. M. (2005). Long-term functional outcomes and academic-occupational status in implanted children after 10 to 14 years of cochlear implant use. *Otol. Neurotol.* 26, 1152–1160. doi: 10.1097/01.mao.0000180483.16619.8f

Bhadania, S. R., Vishwakarma, R., and Keshri, A. (2018). Cochlear implant device failure in the postoperative period: an institutional analysis. *Asian J. Neurosurg.* 13, 1066–1070. doi: 10.4103/ajns.AJNS_93_17

Büchner, A., and Gärtner, L. (2017). [Technical advancements in cochlear implants : state of the art]. *HNO* 65, 276–289. doi: 10.1007/s00106-017-0339-7

Burian, K., and Eisenwort, B. (1989). [Experiences with the Vienna auditory prosthesis]. *Laryngorhinootologie* 68, 95–97. doi: 10.1055/s-2007-998293

Cao, K., and Wei, C. (2010). Evaluation of the effect of cochlear implantation in children. *Chin. J. Otorhinolaryngol. Head Neck Surg.* 45, 798–799.

Carlyon, R. P., and Goehring, T. (2021). Cochlear implant research and development in the twenty-first century: a critical update. *J. Assoc. Res. Otolaryngol.* 22, 481–508. doi: 10.1007/s10162-021-00811-5

Chang, Y. P., Chang, R. Y., Lin, C. Y., and Luo, X. (2016). Mandarin tone and vowel recognition in cochlear implant users: effects of talker variability and bimodal hearing. *Ear Hear.* 37, 271–281. doi: 10.1097/AUD.0000000000000265

Chen, J., Chen, B., Shi, Y., and Li, Y. (2021). A retrospective review of cochlear implant revision surgery: a 24-year experience in China. *Eur. Arch. Otorhinolaryngol.* Online ahead of print. doi: 10.1007/s00405-021-06745-1

Chung, D., Kim, A. H., Parisier, S., Linstrom, C., Alexiades, G., Hoffman, R., et al. (2010). Revision cochlear implant surgery in patients with suspected soft failures. *Otol. Neurotol.* 31, 1194–1198. doi: 10.1097/MAO.0b013e3181f0c631

Ciorba, A., Bovo, R., Trevisi, P., Rosignoli, M., Aimoni, C., Castiglione, A., et al. (2012). Postoperative complications in cochlear implants: a retrospective analysis of 438 consecutive cases. *Eur. Arch. Otorhinolaryngol.* 269, 1599–1603. doi: 10.1007/s00405-011-1818-1

Dağkıran, M., Tarkan, Ö, Sürmelioğlu, Ö, özdemir, S., Onan, E., Tuncer, Ü, et al. (2020). Management of complications in 1452 pediatric and adult cochlear implantations. *Turk Arch. Otorhinolaryngol.* 58, 16–23. doi: 10.5152/tao.2020.5025

Desoyer, H., and Burian, K. (1985). Reimplantation of a molded scala tympani electrode: impact on psychophysical and speech discrimination abilities. *Ann. Otol. Rhinol. Laryngol.* 94(1 Pt 1), 65–70. doi: 10.1177/000348948509400114

Distinguin, L., Blanchard, M., Rouillon, I., Parodi, M., and Loundon, N. (2017). Pediatric cochlear reimplantation: decision-tree efficacy. *Eur. Ann. Otorhinolaryngol. Head Neck Dis.* 135, 243–247. doi: 10.1016/j.anorl.2018.05.002

Dotú, C. O., del Prado, M., Pizarro, V., Beltrán, J. D. J., and Delago, M. D. J. (2010). Cochlear reimplantation in the same ear: findings, peculiarities of the surgical technique and complications. *Acta Otorrinolaringol. Esp.* 61, 106–117. doi: 10.1016/j.otorri.2009.10.014

Dunn, C. C., Etler, C., Hansen, M., and Gantz, B. J. (2015). Successful hearing preservation after reimplantation of a failed hybrid cochlear implant. *Otol. Neurotol.* 36, 1628–1632. doi: 10.1097/MAO.0000000000000867

Googe, B. J., and Carron, J. D. (2016). Analyzing complications of minimally invasive pediatric cochlear implantation: a review of 248 implantations. *Am. J. Otolaryng.* 37, 44–50. doi: 10.1016/j.amjoto.2015.09.002

Gözen, E. D., Tevetoğlu, F., Yener, H. M., Kara, E., Ataş, A., Sirolu, S., et al. (2019). Extra-Cochlear insertion in cochlear implantation: a potentially disastrous condition. *J. Int. Adv. Otol.* 15, 358–363. doi: 10.5152/iao.2019.6492

Hamzavi, J., Baumgartner, W. D., Pok, S. M., Franz, P., and Gstoettner, W. (2003). Variables affecting speech perception in postlingually deaf adults following cochlear implantation. *Acta Otolaryngol.* 123, 493–498. doi: 10.1080/0036554021000028120

Han, D. (2004). Multichannel cochlear implantation in China. *Chin. J. Otorhinolaryngol.* 39, 70–71.

Henson, A. M., Slattery, W. R., Luxford, W. M., and Mills, D. M. (1999). Cochlear implant performance after reimplantation: a multicenter study. *Am. J. Otol.* 20, 56–64.

Hermann, R., Coudert, A., Aubry, K., Bordure, P., Bozorg-Grayeli, A., Deguine, O., et al. (2020). The French National Cochlear Implant Registry (EPIIC): cochlear explantation and reimplantation. *Eur. Ann. Otorhinolaryngol. Head Neck Dis.* 137(Suppl. 1), S45–S49. doi: 10.1016/j.anorl.2020.07.006

Hochmair-Desoyer, I., and Burian, K. (1985). Reimplantation of a molded scala tympani electrode: impact on psychophysical and speech discrimination abilities. *Ann. Otol. Rhinol. Laryngol.* 94, 65–70. doi: 10.1177/000348948509400114

Holcomb, M. A., Burton, J. A., Dornhoffer, J. R., Camposeo, E. L., Meyer, T. A., and Mcrackan, T. R. (2018). When to replace legacy cochlear implants for technological upgrades: indications and outcomes. *Laryngoscope* 129, 748–753. doi: 10.1002/lary.27528

Hwang, K., Lee, J. Y., Oh, H. S., Lee, B. D., Jung, J., and Choi, J. Y. (2019). Feasibility of revision cochlear implant surgery for better speech comprehension. *J. Audiol. Otol.* 23, 112–117. doi: 10.7874/jao.2018.00430

Incesulu, A., Adapinar, B., and Kecik, C. (2008). Cochlear implantation in cases with incomplete partition type III (X-linked anomaly). *Eur. Arch. Otorhinolaryngol.* 265, 1425–1430. doi: 10.1007/s00405-008-0614-z

Jackler, R. K., Leake, P. A., and McKerrow, W. S. (1989). Cochlear implant revision: effects of reimplantation on the cochlea. *Ann. Otol. Rhinol. Laryngol.* 98, 813–820. doi: 10.1177/000348948909801012

Jayawardena, J., Kuthubutheen, J., and Rajan, G. (2012). Hearing preservation and hearing improvement after reimplantation of pediatric and adult patients with partial deafness: a retrospective case series review. *Otol. Neurotol.* 33, 740–744. doi: 10.1097/MAO.0b013e318255dd91

Kamat, A., Goldin, L., and Hoffman, R. A. (2011). Unusual electroacoustic device failure and electroacoustic reimplantation with hearing preservation. *Otol. Neurotol.* 32, 948–950. doi: 10.1097/MAO.0b013e31821ac88a

Kanchanalarp, C., Cheewaruangroj, W., Kasemsuwan, L., Thawin, C., and Sriwanyong, S. (2005). Pediatric cochlear implantation: experience in Thai patients. *J. Med. Assoc. Thai.* 88, 484–491.

Karamert, R., Düzlü, M., Tutar, H., Eravci, F. C., Türkcan, A. K., Zorlu, M. E., et al. (2019). Assessment of cochlear implant revision surgeries in a cohort of 802 patients. *Otol. Neurotol.* 40, 464–470. doi: 10.1097/MAO.0000000000002152

Lailach, S., Neudert, M., and Zahnert, T. (2021). [Update cochlear-implantation: indications and surgical aspects]. *Laryngorhinootologie* 100, 652–672. doi: 10.1055/a-1491-3426

Lane, C., Zimmerman, K., Agrawal, S., and Parnes, L. (2019). Cochlear implant failures and reimplantation: a 30-year analysis and literature review. *Laryngoscope* 130, 782–789. doi: 10.1002/lary.28071

Lassig, A. A., Zwolan, T. A., and Telian, S. A. (2005). Cochlear implant failures and revision. *Otol. Neurotol.* 26, 624–634. doi: 10.1097/01.mao.0000178123.35988.96

Layfield, E., Hwa, T. P., Naples, J., Maina, I., Brant, J. A., Eliades, S. J., et al. (2021). Failure and revision surgery after cochlear implantation in the adult population: a 10-year single-institution retrospective and systematic review of the literature. *Otol. Neurotol.* 42, 408–413. doi: 10.1097/MAO.0000000000002940

Lenarz, T. (2017). Cochlear implant – state of the art. *GMS Curr. Top. Otorhinolaryngol. Head Neck Surg.* 16:Doc04.

Lescanne, E., Al, Z. M., Bakhos, D., Robier, A., and Morinière, S. (2011). Revision surgeries and medical interventions in young cochlear implant recipients. *Int. J. Pediatr. Otorhinolaryngol.* 75, 1221–1224. doi: 10.1016/j.ijporl.2011.07.003

Lu, X. (2018). *The Study on Auditory and Language Development and Health-related Quality of Life in Mandarin-speaking Children with Cochlear Implants.* China: Zhengzhou University.

Lu, Y., and Cao, K. (2014). Cochlear implant operation to summarize and postoperative outcome. *J. Clin. Otorhinolarygol. Head Neck Surg. (China)* 28, 1768–1773.

Luo, J., Hou, X., Sun, J., and Sun, J. (2020). Causes of revision surgery after cochlear implantation. *Chinese J. Otol.* 18, 871–874.

Luo, X., Fu, Q. J., Wei, C. G., and Cao, K. L. (2008). Speech recognition and temporal amplitude modulation processing by Mandarin-speaking cochlear implant users. *Ear Hear.* 29, 957–970. doi: 10.1097/AUD.0b013e3181888f61

Maas, S., Bance, M., O'Driscoll, M., Mawman, D., and Ramsden, R. T. (1996). Explanatation of a nucleus multichannel cochlear implant and re-implantation into the contralateral ear. a case report of a new strategy. *J. Laryngol. Otol.* 110, 881–883. doi: 10.1017/s0022215100135236

Manrique-Huarte, R., Huarte, A., and Manrique, M. J. (2015). Surgical findings and auditory performance after cochlear implant revision surgery. *Eur. Arch. Otorhinolaryngol.* 273, 621–629. doi: 10.1007/s00405-015-3610-0

Marlowe, A. L., Chinnici, J. E., Rivas, A., Niparko, J. K., and Francis, H. W. (2010). Revision cochlear implant surgery in children: the johns hopkins experience. *Otol. Neurotol.* 31, 74–82. doi: 10.1097/MAO.0b013e3181c29fad

No authors listed (2005). European consensus statement on cochlear implant failures and explantations. *Otol. Neurotol.* 26, 1097–1099.

Orús, D. C., Venegas, P. M. P., De Juan, B. J., and De Juan, D. M. (2010). [Cochlear reimplantation in the same ear: findings, peculiarities of the surgical technique and complications]. *Acta Otorrinolaringol. Esp.* 61, 106–117. doi: 10.1016/j.otorri.2009.10.014

Parisier, S. C., Chute, P. M., Popp, A. L., and Suh, G. D. (2001). Outcome analysis of cochlear implant reimplantation in children. *Laryngoscope* 111, 26–32. doi: 10.1097/00005537-200101000-00005

Parisier, S. C., Chute, P. M., Weiss, M. H., Hellman, S. A., and Wang, R. C. (1991). Results of cochlear implant reinsertion. *Laryngoscope* 101, 1013–1015. doi: 10.1288/00005537-199109000-00015

Polak, M., Ulubil, S. A., Hodges, A. V., and Balkany, T. J. (2006). Revision cochlear implantation for facial nerve stimulation in otosclerosis. *Arch. Otolaryngol. Head Neck Surg.* 132, 398–404. doi: 10.1001/archotol.132.4.398

Puri, S., Dornhoffer, J. L., and North, P. E. (2005). Contact dermatitis to silicone after cochlear implantation. *Laryngoscope* 115, 1760–1762. doi: 10.1097/01.mlg.0000172202.58968.41

Qiu, J., Chen, Y., Tan, P., Han, Y., Gao, L., Lu, Y., et al. (2010). Complications of cochlear implantation and clinical analysis. *Chinese J. Otol.* 8, 235–239.

Qiu, J., Yu, C., Ariyaratne, T. V., Foteff, C., Ke, Z., Sun, Y., et al. (2017). Cost-Effectiveness of pediatric cochlear implantation in rural China. *Otol. Neurotol.* 38, e75–e84. doi: 10.1097/MAO.0000000000001389

Rivas, A., Marlowe, A. L., Chinnici, J. E., Niparko, J. K., and Francis, H. W. (2008). Revision cochlear implantation surgery in adults: indications and results. *Otol. Neurotol.* 29, 639–648. doi: 10.1097/MAO.0b013e31817e5d31

Robert, K. J., Patricia, A. L., and William, S. M. (1989). Cochlear implant revision: effects of reimplantation on the cochlea. *Ann. Otol. Rhinol. Laryngol.* 10, 813–820. doi: 10.1177/000348948909801012

Roßberg, W., Timm, M., Matin, F., Zanoni, A., Krüger, C., Giourgas, A., et al. (2021). First results of electrode reimplantation and its hypothetical dependence from artificial brain maturation. *Eur. Arch. Otorhinolaryngol.* 278, 951–958. doi: 10.1007/s00405-020-06125-1

Saeed, S. R., Ramsden, R. T., Hartley, C., Woolford, T. J., and Boyd, P. (1995). Cochlear reimplantation. *J. Laryngol. Otol.* 109, 980–985.

Sevy, A. B., Bortfeld, H., Huppert, T. J., Beauchamp, M. S., Tonini, R. E., and Oghalai, J. S. (2010). Neuroimaging with near-infrared spectroscopy demonstrates speech-evoked activity in the auditory cortex of deaf children following cochlear implantation. *Hear. Res.* 270, 39–47. doi: 10.1016/j.heares.2010.09.010

Sorkin, D. L., and Buchman, C. A. (2016). Cochlear implant access in six developed countries. *Otol. Neurotol.* 37, e161–e164. doi: 10.1097/MAO.0000000000000946

Spillmann, T., and Dillier, N. (1989). Comparison of single-channel extracochlear and multichannel intracochlear electrodes in the same patient. *Br. J. Audiol.* 23, 25–31. doi: 10.3109/03005368909077815

Sterkers, F., Merklen, F., Piron, J. P., Vieu, A., Venail, F., Uziel, A., et al. (2015). Outcomes after cochlear reimplantation in children. *Int. J. Pediatr. Otorhinolaryngol.* 79, 840–843. doi: 10.1016/j.ijporl.2015.03.015

Sun, C. (2019). Retrospective analysis of patients undergoing reoperation after cochlear implantation. doi: 10.1097/00005537-199903000-00002

Sunde, J., Webb, J. B., Moore, P. C., Gluth, M. B., and Dornhoffer, J. L. (2013). Cochlear implant failure, revision, and reimplantation. *Otol. Neurotol.* 34, 1670–1674. doi: 10.1097/MAO.0000000000000079

Tang, H., Hu, R., Li, Q., and Tang, Y. (2019). Clinical analysis of cochlear reimplantation in 23 pediatric patients. *Chinese J. Otorhinolaryngology-skull Base Surg.* 25, 466–469, 475. doi: 10.11798/j.issn.1007-1520.201905004

Tarkan, Ö, Tuncer, Ü, özdemir, S., Sürmelioğlu, Ö, çetik, F., Kıroğlu, M., et al. (2013). Surgical and medical management for complications in 475 consecutive pediatric cochlear implantations. *Int. J. Pediatr. Otorhinolaryngol.* 77, 473–479. doi: 10.1016/j.ijporl.2012.12.009

Terry, B., Kelt, R. E., and Jeyakumar, A. (2015). Delayed complications after cochlear implantation. *JAMA Otolaryngology–Head Neck Surg.* 141, 1012–1017.

Wang, J. T., Wang, A. Y., Psarros, C., and Da, C. M. (2014). Rates of revision and device failure in cochlear implant surgery: a 30-year experience. *Laryngoscope* 124, 2393–2399. doi: 10.1002/lary.24649

Wang, S., Lin, M., Sun, L., Chen, X., Fu, X., Yan, L., et al. (2020). Neural mechanisms of hearing recovery for cochlear-implanted patients: an electroencephalogram follow-up study. *Front. Neurosci.* 14:624484. doi: 10.3389/fnins.2020.624484

Webb, R. L., Laszig, R., Lehnhardt, E., Pyman, B. C., Clark, G. M., and Franz, B. K. H. G. (1991). Surgical complications with the cochlear multiple-channel intracochlear implant: experience at hannover and melbourne. *Ann. Otol. Rhinol. Laryngol.* 100, 131–136. doi: 10.1177/000348949110000208

Weder, S., Shaul, C., Wong, A., O'Leary, S., and Briggs, R. J. (2020). Management of severe cochlear implant infections-35 years clinical experience. *Otol. Neurotol.* 41, 1341–1349. doi: 10.1097/MAO.0000000000002783

Weise, J. B., Muller-Deile, J., Brademann, G., Meyer, J. E., Ambrosch, P., and Maune, S. (2005). Impact to the head increases cochlear implant reimplantation rate in children. *Auris Nasus Larynx* 32, 339–343. doi: 10.1016/j.anl.2005.03.017

Xi, X., Han, D., Ji, F., and Hong, M. (2005). New progress in otology-profile of multichannel cochlear implants in China. *Chinese J. Practical Med.* 021, 111–113.

Yu, L., Chen, F., Zheng, H., and Ma, X. (2004). Cochlear reimplantation. *Chinese J. Otorhi-nolaryngol.* 39, 598–601. doi: 10.3760/j.issn:1673-0860.2004.10.006

Zhao, S., Zheng, Y., Liu, X., and Chen, S. (2008). Cause and effect of cochlear reimplantation. *Chinese J. Ophthalmol. Otorhinolaryngol.* 8:180. doi: 10.3969/j.issn.1671-2420.2008.03.020

Zwolan, T., and Verhof, J. (2017). *AAMI CI86 Cochlear Implant Systems – Requirements for Safety, Functional Verification, Labeling, and Reliability Reporting*. Virginia: Association for the Advancement of Medical Instrumentation.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 23 March 2022
doi: 10.3389/fnins.2022.857855

Check for updates

Analysis of Long-Term Cochlear Implantation Outcomes and Correlation With Imaging Characteristics in Patients With Common Cavity Deformity

Lifang Zhang[1,2†], Biao Chen[1,2†], Ying Kong[2,3], Natalia Liau[1,2], Xingmei Wei[1,2], Ying Shi[1,2], Jingyuan Chen[1,2], Mengge Yang[1,2], Anandhan Dhanasingh[4] and Yongxin Li[1,2]*

[1] Department of Otolaryngology, Head and Neck Surgery, Beijing Tongren Hospital, Capital Medical University, Beijing, China, [2] Key Laboratory of Otolaryngology, Head and Neck Surgery, Ministry of Education, Capital Medical University, Beijing, China, [3] Beijing Institute of Otolaryngology, Ministry of Education, Capital Medical University, Beijing, China, [4] MED-EL Medical Electronics GmbH, Innsbruck, Austria

OPEN ACCESS

Edited by:
Fei Chen,
Southern University of Science and Technology, China

Reviewed by:
Mohan Kameswaran,
Madras ENT Research Foundation (P) Ltd., India
Davide Brotto,
University of Padua, Italy

****Correspondence:***
Yongxin Li
entlyx@ccmu.edu.cn

† These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *19 January 2022*
Accepted: *14 February 2022*
Published: *23 March 2022*

Citation:
Zhang L, Chen B, Kong Y, Liau N, Wei X, Shi Y, Chen J, Yang M, Dhanasingh A and Li Y (2022) Analysis of Long-Term Cochlear Implantation Outcomes and Correlation With Imaging Characteristics in Patients With Common Cavity Deformity. Front. Neurosci. 16:857855. doi: 10.3389/fnins.2022.857855

Object: To investigate the long-term development of auditory and speech in patients with common cavity deformity (CCD) after cochlear implantation (CI) and its relationship to imaging characteristics.

Methods: Twenty-three CCD patients and 59 age- and sex-matched CI children with normal inner ear structure were recruited. The auditory and speech development of these two groups were evaluated at 0, 1, 3, 6, 12, and 18 months after CI activation using four parent reports questionnaires [Categories of Auditory Performance (CAP), Speech Intelligibility Rating (SIR), Meaningful Auditory Integration Scale/Infant-Toddler Meaningful Auditory Integration Scale (MAIS/ITMAIS), and Meaningful Use of Speech Scale (MUSS)]. Computed tomography-based 3-dimensional reconstruction of the surgical side of 18 CCD children was performed, the volume and surface area were calculated. Correlation analysis was performed on the imaging performance and post-operative outcomes.

Results: The percentages of MAIS/IT-MAIS scores and CAP scores at different evaluation time points are significantly different ($p < 0.05$). When comparing SIR results across time points, significant growth was observed in most of the comparisons. In addition, significant differences ($p < 0.05$) are observed among the percentages of MUSS scores at different time points except the comparison between 0 and 1 month after CI activation. Patients in the CCD group had poorer auditory and speech performances at different stages after CI compared with those in the control group. According to the reconstruction of CCD patients, the volume ranged from 12.21 to 291.96 mm^3; the surface area ranged from 27.81 to 284.7 mm^2. When the lumen surface area was <190.45 mm^2 or the volume was <157.91 mm^3, the survival time for CCD children to achieve a CAP score of 4 after CI was significantly shorter.

Conclusion: Cochlear implantation are less effective in CCD patients than in patients with normal inner ear structures, but they can still achieve significant improvement post-operatively. The morphology and size of the inner ear vary in CCD patients, which reflects the degree of inner ear development influences the outcome after CI surgery.

Keywords: common cavity deformity, cochlear implantation, auditory development, speech development, 3D reconstruction

INTRODUCTION

Common cavity deformity (CCD) is characterized by the presence of an abnormally ovoid or round chamber formed by the cochlea and vestibule and is generally associated with profound sensorineural hearing loss (SNHL) (Sennaroğlu and Bajin, 2017). This condition occurs due to the arrest of otocyst development during the fourth week of embryonic development (Jackler et al., 1987). CCD is diagnosed primarily on the basis of computed tomography (CT) or magnetic resonance imaging (MRI), showing a single fluid-filled cavity of the cochlea and vestibule (Casselman et al., 2001). This malformation varies in size and shape as well as in the location of the internal auditory canal and its size is often assumed to be related to the arrest time of the cochlear vestibule development; that is, the larger the cystic cavity, the later the arrest (Brotto et al., 2019).

For CCD patients, we often choose a cochlear implant (CI) or an auditory brainstem implant (ABI) to help them restore their hearing. Since the ABI was first used in a patient with neurofibromatosis type 2 in 1979, its indications have been constantly updated. In recent years, ABI has been applied to patients with profound SNHL who suffer from conditions such as cochlear sclerosis and severe cochlear malformations (Colletti et al., 2005; Bozorg Grayeli et al., 2007; Sennaroğlu et al., 2016). Sennaroğlu et al. (2016) performed ABI surgeries on seven patients with CCD and found that these patients achieved better hearing threshold and language outcome scores compared with patients with other types of severe cochlear malformations. Nevertheless, given the complications following ABI, such as cerebrospinal fluid leakage, electrode displacement, and limited post-operative benefit (Toh and Luxford, 2008), cochlear implantation (CI) is currently the primary intervention for CCD.

In 1986, McElveen et al. (1997) reported the first CI in a patient with CCD using the transmastoid facial recess approach. Since then, an increasing number of CCD patients have received CI surgery, and the surgical approach and electrodes have continually improved (Molter et al., 1993; Tucci et al., 1995; McElveen et al., 1997; Beltrame et al., 2000, 2005, 2013; Sennaroglu et al., 2014). In 2017, our team proposed that custom-made electrodes could be implanted *via* the transmastoid slotted labyrinthotomy approach (TSLA) for CCD patients (Wei et al., 2018). Instead of conventional electrodes, we used custom electrodes made by MED-EL, with 12 electrodes in the middle of the electrode array and extension wires made of inert silicone carriers containing platinum wires. This strategy allows the electrodes to remain as attached to the lumen as possible, allowing them to stimulate a larger area, which may mean that more spiral ganglion cells are stimulated, resulting in better post-operative outcomes.

Common cavity deformity patients could have long-term benefits with CI (Al-mahboob et al., 2022). Although most studies have concluded that post-operative outcomes in CCD patients are worse than in children with normal inner ear structures and mild malformations, specific auditory speech rehabilitation outcomes are inconsistent (Ahmad et al., 2005; Papsin, 2005; Ahn et al., 2011). In addition, a previous study showed a correlation between the effect of CI implantation and imaging performance (Wei et al., 2017). However, to date, no study has investigated the relationship between post-operative CI outcomes and imaging performance in CCD patients. The present study aims to understand the post-operative auditory-speech performance and developmental patterns of the CCD children implanted with custom electrodes *via* the TSLA approach by analyzing the long-term post-operative outcomes of these children and their relationship with imaging characteristics.

MATERIALS AND METHODS

Participants

A total of 23 pediatric patients (12 female and 11 male) with CCD were recruited from April 2016 to January 2020 at the Department of Otolaryngology Head and Neck Surgery of our hospital. High-resolution computed tomography (HRCT) and inner ear MRI were performed before surgery in all cases, and the diagnosis of CCD was confirmed by two or more physicians from the Radiology and Otorhinolaryngology departments.

Inclusion criteria were as follows.

(1) Bilateral severe or profound SNHL patients diagnosed with CCD.
(2) Available for post-operative follow-up.

Exclusion criteria were as follows.

(1) Patients with contraindications to cochlear implantation.
(2) Patients with history of serious systemic diseases or intellectual disorders.

The age at implantation ranged from 0 to 7 years (mean: 27.65 months, standard deviation: 13.79 months).

A total of 59 congenitally severe or profound SNHL children who met the inclusion criteria and who had normal inner ear structure, matched for age and sex, were recruited as a control group. The age range was 0–8 years (mean, 29.00 months; standard deviation, 20.14 months).

In compliance with ethical standards for human subjects, written informed consent was obtained from the guardians of all participants before proceeding with the study procedures. This study was approved by the Institutional Review Board of our hospital.

Procedures

Routine Clinical Investigations

All participants underwent routine otorhinolaryngological examination, followed by audiological tests, CT, and MRI scans before the CI surgery. To investigate the audiological status in terms of hearing level, function of the central auditory system, and the development of the auditory system, the following audiological tests were performed: (1) Behavioral hearing assessment, (2) Auditory Steady State Response, (3) Auditory Brainstem Response, (4) Distorted Product Otoacoustic Emissions, (5) conventional low 226 Hz tympanometry, and (6) 40-Hz auditory-evoked related potential.

Surgical Approach

In the CCD group, all children were implanted with customized electrodes using a transmastoid slotted labyrinthotomy approach (TSLA) (Wei et al., 2018). In the TSLA, the bony wall of the cavity was exposed after mastoidectomy, and a slot was made in the area where the lateral semicircular canal is commonly situated away from the facial nerve. A customized electrode was placed in the cavity toward the cochlear side after the perilymph flow abated (Shi et al., 2019). The electrodes were fully implanted in all children except for one child who had two extra cochlear electrodes because of the small size of the common cavity. None of the children had post-operative complications, such as facial paralysis or cerebrospinal fluid leakage.

In the control group, electrodes were implanted using the conventional transmastoid facial recess approach. All electrodes were successfully implanted.

Post-operative Follow-Up Questionnaires

The assessment of the child's auditory and speech development was performed using the categories of auditory performance (CAP), Speech Intelligibility Rating (SIR), Meaningful Auditory Integration Scale/Infant-Toddler Meaningful Auditory Integration Scale (MAIS/ITMAIS), and Meaningful Use of Speech Scale (MUSS). The parents or guardians of the infants were asked to complete the questionnaires. The questionnaires were evaluated at the activation of CI and at 1, 3, 6, 9, 12, 18, 24, and 36 months after activation. For the MAIS/IT-MAIS and MUSS, we converted the actual scores into percentages as final statistics, and the result was expressed as a percentage using the following equation: total score/40 × 100%.

Imaging Evaluation

All patients underwent a temporal bone CT scan (GE 64-row helical CT, Boston, MA, United States) and inner ear MRI (Philips Ingenia 3.0 T MRI scan, Amsterdam, Netherlands) at our hospital before surgery. The E-3D digital medical design system (Liao, 2018) was applied to reconstruct the lumen on the CI side in the CCD group to obtain the volume and surface area of the cavity.

Data Analysis

Statistical analyses were performed using SPSS 22.0. The results of the MAIS/IT-MAIS and MUSS for both groups did not comply with the normal distribution by the Shapiro-Wilk test. The results of the non-normal distribution were described using the median (25th percentile, 75th percentile) and a non-parametric test was used to compare whether their differences were significant between the two groups and between the various assessment stages. Since the results of the cap questionnaire as well as the SIR questionnaire were rank data, the Wilcoxon test was used to compare whether there was a significant difference between the results of the two groups and between the various assessment stages. For lumen volume and surface area, Kaplan-Meier survival analysis was used to analyze their correlation with post-operative outcomes. Considering that patients with a CAP score of 4 could discriminate speech sounds without the aid of lip-reading, we defined a CAP score of 4 as the endpoint event. The post-operative time for a child to reach a CAP score of 4 was defined as the survival time. The median volume and median surface area were used as criteria for grouping. At the same time, $p < 0.05$ was considered a statistically significant difference.

RESULTS

- **Demographic information of the CCD group**

Table 1 summarizes the subject information, including gender, age at implantation, and length of follow-up. In the CCD group, there were 11 males and 12 females, with an implantation age of 0–7 years and a mean implantation age of 28 months, while the control group consisted of 31 males and 28 females, with an implantation age of 0–8 years and a mean implantation age of 29 months. The groups were not statistically different in terms of age and gender after non-parametric testing ($p < 0.05$).

- **Development of auditory ability in the CCD Group**

As shown in **Tables 2**, **3**, both the percentage of MAIS/IT-MAIS scores and the CAP scores were significantly different between each time point in the CCD group ($p < 0.05$).

- **Development of speech in the CCD Group**

For the percentage of MUSS scores (**Table 4**), there was no significant difference between the results of each time point except between 0 and 1 month after cochlear activation in the CCD group ($p < 0.05$). The SIR scores of the CCD group (**Table 5**) were significantly different between 0 and 6 months, 1 and 6 months, 0 and 12 months, 1 and 12 months, 3 and 12 months, 6 and 12 months, 0 and 18 months, 1 and 18 months, 3 and 18 months, and 6 and 18 months after cochlear activation ($p < 0.05$).

- **Characteristics of IT-MAIS/MAIS in the CCD and control groups**

TABLE 1 | Demographic information of the recruited patients.

Group	Gender		Mean implantation age ($\bar{x} \pm SD$)	Average length of follow-up (M)	Maximum length of follow-up (M)
	Male	Female			
CCD	11	12	27.65 ± 17.30	23	48
Control group	31	28	29.00 ± 20.41	28	48

Figure 1A shows the mean, median, and 25th and 75th percentiles of the MAIS/ITMAIS score percentages for the CCD and control groups. The median scores were significantly different between the two groups at 1, 6, 12, and 18 months after cochlear activation ($p < 0.05$). Furthermore, trend of the average MAIS/ITMAIS score percentages of the two groups as the follow-up time increased.

- **Characteristics of CAP in the CCD and control groups**

Similar results were found in the comparison of CAP scores. Significant differences between the two groups were observed at 1, 3, and 6 months after CI activation (**Figure 1B**). The average CAP scores of the two groups increased over the follow-up period, but the average scores of the control group were higher than those of the CCD group.

- **Characteristics of MUSS in the CCD and Control Groups**

As for the median percentage of MUSS scores, there was no significant difference between the two groups at each time point after CI activation (**Figure 1C**). Nevertheless, the average percentage of MUSS scores increased slowly over time in both groups, and the CCD group obtained lower scores than the normal group (**Figure 1C**).

- **Characteristics of SIR in the CCD and control groups**

After statistically analyzing the differences in median SIR scores between the CCD and control groups, we observed that the differences were significant at 1, 3, 6, 12, and 18 months after CI activation (**Figure 1D**). Similar to the other questionnaires, the mean SIR score in the CCD group increased gradually over time and was worse than that in the control group.

- **Correlations between Imaging Characteristics and CAP results**

As shown in **Figure 2**, we reconstructed CT images of the surgical side of the cavity in 18 CCD patients, calculated their volume and surface area (**Table 6**), and performed a Kaplan-Meier survival analysis with the CAP results (**Figure 3**). Considering that patients with a CAP score of 4 could discriminate speech sounds without the aid of lip-reading, we defined a CAP score of 4 as the endpoint event. The post-operative time for a child to reach a CAP score of 4 was defined as the survival time. When the lumen surface area was $\geq$190.45 mm^2, the mean survival time for CCD children to achieve a CAP score of 4 after surgery was 20.57 months, and the median survival time was 18.00 months; when the lumen surface area was <190.45 mm^2, the mean survival time for CCD children to reach a CAP score of 4 after surgery was 12 months, and the median survival time was 12.00 months, with a statistically significant difference between the two groups (p = 0.02) (**Figure 3A**). When the lumen volume was $\geq$157.91 mm^3, the mean survival time was 20.571 months, and the median time was 18.00 months; when it was <157.91 mm^3, the mean and median survival times were 13.142 and 12.00 months, respectively.

TABLE 2 | Comparison of the percentage of Meaningful Auditory Integration Scale/Infant-Toddler Meaningful Auditory Integration Scale (MAIS/IT-MAIS) scores for each time point in the common cavity deformity (CCD) group.

	0 m	1 m	3 m	6 m	12 m
1 m	0.012				
3 m	0.018	0.005			
6 m	0.018	0.008	0.036		
12 m	0.008	0.003	0.010	0.008	
18 m	0.043	0.043	0.012	0.025	0.027

TABLE 3 | Comparison of the categories of auditory performance (CAP) scores for each time point in the common cavity deformity (CCD) group.

	0 m	1 m	3 m	6 m	12 m
1 m	0.011				
3 m	0.007	0.034			
6 m	0.003	0.005	0.016		
12 m	0.002	0.002	0.001	0.015	
18 m	0.014	0.007	0.003	0.011	0.015

TABLE 4 | Comparison of the percentage of Meaningful Use of Speech Scale (MUSS) scores for each time point in the common cavity deformity (CCD) group.

	0 m	1 m	3 m	6 m	12 m
1 m	0.068				
3 m	0.005	0.016			
6 m	0.002	0.012	0.011		
12 m	0.001	0.003	0.002	0.013	
18 m	0.018	0.027	0.017	0.018	0.026

TABLE 5 | Comparison of the Speech Intelligibility Rating (SIR) scores for each time point in the common cavity deformity (CCD) group.

	0 m	1 m	3 m	6 m	12 m
1 m	1.000				
3 m	0.157	0.157			
6 m	0.046	0.046	0.317		
12 m	0.001	0.001	0.003	0.014	
18 m	0.006	0.006	0.005	0.030	0.257

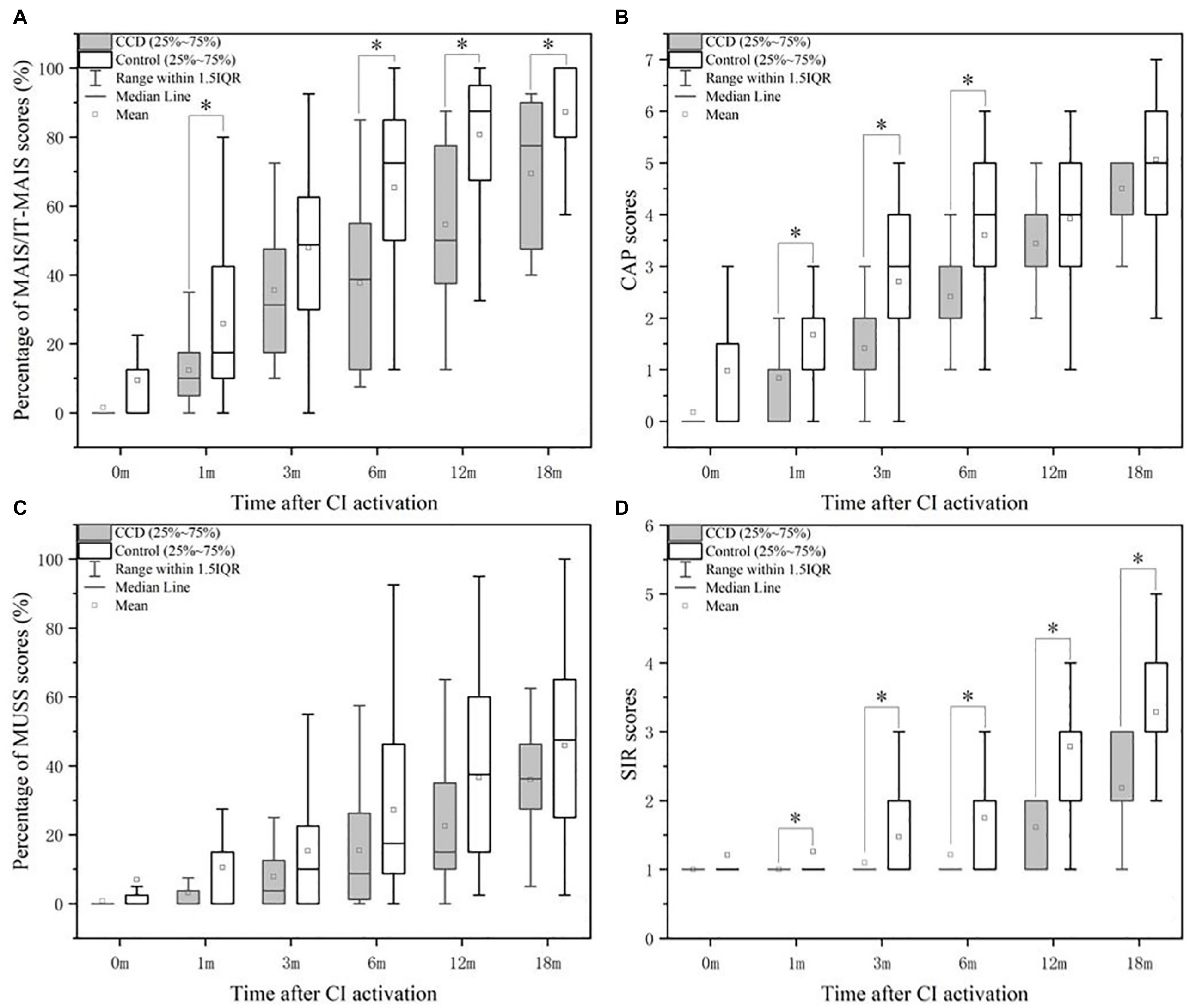

FIGURE 1 | Comparison of the performance in MAIS/IT-MAIS **(A)**, CAP **(B)**, MUSS **(C)**, and SIR **(D)** between the CCD and control groups. CAP, Categories of auditory performance; CCD, common cavity deformity; MAIS/IT-MAIS, Meaningful Auditory Integration Scale/Infant-Toddler Meaningful Auditory Integration Scale; MUSS, Meaningful Use of Speech Scale; SIR, Speech Intelligibility Rating. *indicates significant ($p < 0.05$).

There was a significant difference between the two groups ($p = 0.022$) (**Figure 3B**).

DISCUSSION

Common cavity deformity is regarded as a severe inner ear malformation, and the post-operative outcome of CI is generally considered to be worse in patients with CCD than in children with normal cochlear structures. In this study, 23 children with CCD were followed up using IT-MAIS/MAIS, MUSS, CAP, and SIR to assess their post-operative outcomes and correlations with imaging performance.

The MAIS is a common tool used to assess functional hearing in hearing-impaired children (Robbins et al., 1991). Each child was assessed by the answers provided by a parent or guardian familiar with the child's condition. It consists of 10 questions that assess the use of hearing aids and the ability to perceive and understand sounds. The CAP is a scale used to assess the auditory ability of pediatric cochlear implant users in their daily lives (Archbold et al., 1998). The above two questionnaires are effective for evaluating the development of auditory stimuli in children who received cochlear implants or hearing aids. In relation to this, stimulation of the remaining spiral ganglion cells in the cochlea has been found to activate hearing (Fayad and Linthicum, 2006) and it has been suggested that spiral ganglion cells are distributed in the cavity walls of patients with CCD. This is the foundation for hearing acquisition after CI in CCD patients (Brotto et al., 2019). ITMAIS/MAIS and CAP scores were lower in the CCD group than in the normal group after

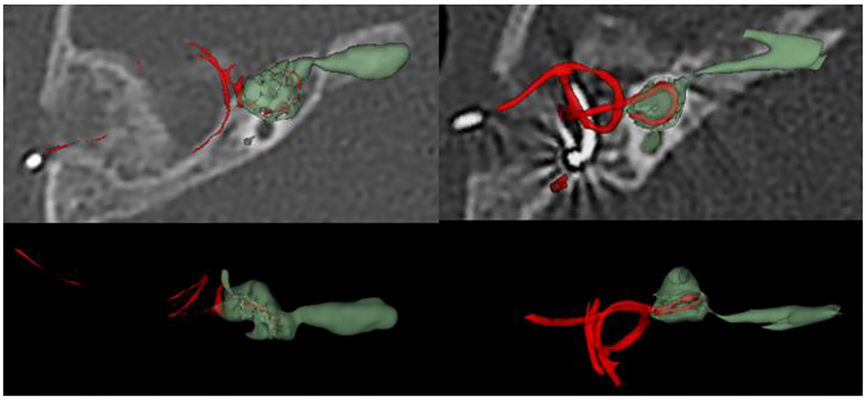

FIGURE 2 | Computed tomography (CT) reconstruction of surgical side in patients with common cavity deformity (CCD).

activation. Similar results were reported by Xia et al. (2015). According to **Tables 2**, **3**, the auditory development of CCD patients continuously improved up to 18 months after activation; that is, in the CCD group, there was no significant platform phase of auditory development in the 18 months after the activation of CI. Ahn et al. (2011) also observed an increase in the auditory performance of CCD patients with prolonged follow-up after 48 months of post-cochlear surgery evaluation, with a mean percentage MAIS score of 90.3 ± 18.1% and a mean CAP score of 4.9 ± 1.6. Although auditory development after CI is slower in CCD patients than in CI patients with normal cochlear structures, progress has been consistently made, suggesting the need for long-term post-operative rehabilitation in these patients.

For the assessment of speech development, MUSS and SIR are commonly used instruments. As shown in **Tables 4**, **5**, there were significant differences in both the percentage of MUSS score and the SIR score when comparing the assessment results at different time points. However, the SIR score did not show a significant difference compared to the previous assessment results until 1 year after activation, whereas the percentage of MUSS score showed a significant difference at 3 months after activation, indicating that CCD patients showed a faster increase in MUSS performance compared to SIR. This may be related to the various aspects of speech development assessed by the two questionnaires. Like the MAIS, the MUSS is a parental report scale. It assesses the use of speech and consists of 10 questions designed to evaluate three aspects of speech development: vocalizing behavior, oral communication skills, and oral clarification skills (Archbold et al., 1998; Fayad and Linthicum, 2006). The SIR has been regularly used to evaluate the intelligibility of spontaneous speech in patients with cochlear implants (Xia et al., 2015). That is, the MUSS questionnaire evaluates speech skills, while SIR rates the intelligibility of pronunciation. However, the young age of the children in our study, with a follow-up period of only 1.5 years, made it difficult to demonstrate significant improvements in speech intelligibility. Nevertheless, since the MUSS includes an evaluation of vocalizing behavior, the children were given the opportunity to improve their scores. Additionally, this study observed that the percentage of MUSS scores in children with CCD was higher at 18 months than at 12 months after the activation, and the difference was significant, indicating that speech ability was still improving at 1.5 year after surgery. Xia et al. (2015) also observed a sustained increase in SIR scores for 4 years after CI. Therefore, post-operative speech development in patients with CCD is slow and requires long-term rehabilitation training.

Based on the comparison between the CCD group and the control group, it is obvious that auditory and speech development after CI was poorer in the CCD group than in the CI children with normal cochlear structures. This may be related to the structure malformation, lack of sufficient spiral ganglion cells (Brotto et al., 2019), or developmental delays (Buchman et al., 2004).

However, we observed good outcomes in some children. By 18 months after the activation, 66.67% of children with CCD had a CAP score of 5, but 16.67% achieved a score of only 3. Among

TABLE 6 | Surface area and volume after three-dimensional reconstruction of the lumen in common cavity deformity (CCD) patients.

Case	CI side	Surface area (mm^2)	Volume (mm^3)
1	Left	224.80	213.96
2	Right	238.00	225.9
3	Left	145.60	128.22
4	Left	284.70	283.6
5	Right	155.60	106.1
6	Left	141.10	117.49
7	Left	238.00	225.9
8	Left	187.10	160.48
9	Right	193.80	182.78
10	Left	254.70	285.07
11	Left	226.90	155.34
12	Left	260.10	291.96
13	Right	97.42	77.77
14	Right	115.90	97.25
15	Left	138.30	113.07
16	Right	224.20	232.94
17	Right	67.85	41.51
18	Right	27.81	12.21

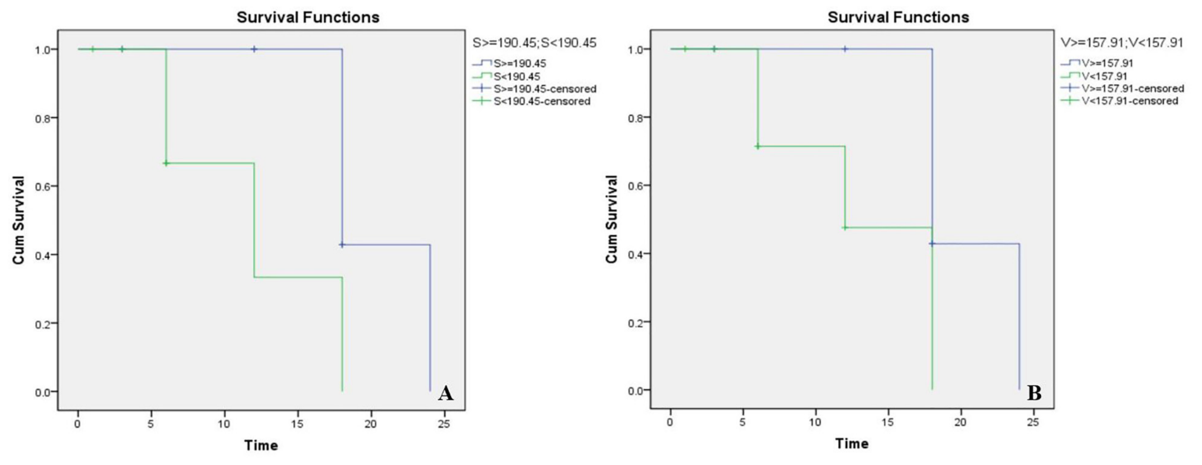

FIGURE 3 | Kaplan-Meier survival chart of different lumen surface areas **(A)** and volumes **(B)**. S, surface areas; V, volumes.

the four children with CCD who had been followed for 3 years, only one reached a score of 40 on the IT-MAIS at 3 years after CI activation. We suspect that this may be related to the distribution and number of spiral ganglion cells in the cavity and the location of the electrode. The higher the number of spiral ganglion cells the electrode can stimulate, the greater the benefit to the child.

Furthermore, the post-operative outcome of CCD patients is also related to cerebrospinal fluid gusher, partial electrode insertion, and fewer active electrodes and the contact of the electrode with the inner wall (Dettman et al., 2011; Bae et al., 2022). The present study explored the relationship of post-operative outcomes with the degree of inner ear development in these patients.

With the development of imaging technology, techniques to assess the development of inner ear structures have become more sophisticated (Skinner et al., 2002; Verbist et al., 2010). After evaluating 36 cochleae with inner ear malformations using volume-rendering technique reconstruction and MPR, Ma et al. (2008) concluded that cochlear development could be more clearly assessed using volume-rendering technique reconstruction and MPR. We calculated the lumen volume and the surface area of 18 CCD patients using three-dimensional reconstruction techniques. The volume ranged from 12.21 to 291.96 mm^3, with a mean volume of 163.98 $\pm$ 84.02 mm^3; the surface area ranged from 27.81 to 284.7 mm^2, with a mean value of 178.99 $\pm$ 71.85 mm^2. This result implies that patients with CCD have variable cochlear morphology and size differences, which is consistent with the findings of Dhanasingh et al. (2019). This difference may require serious consideration of the surgical approach and the choice of electrodes.

Furthermore, after analyzing the correlations between volume, surface area, and post-operative outcomes of CCD patients, we found that the smaller the lumen, the shorter the time to reach a 4-point post-operative CAP. We speculate that this may be due to the smaller lumen, whose spiral ganglion cells may be more densely distributed, thus providing a larger effective area of electrode stimulation, which results in a greater likelihood of stimulation to ganglion cells. Earlier achievement of 4 points in CAP indicates a faster speed of rehabilitation within a year and a half after surgery but is not indicative of higher scores in the distant future.

Therefore, further studies using larger sample sizes and longer follow-up periods are needed to explore the distribution of intracochlear spiral ganglion cells in conjunction with the post-operative electrode location. These studies will provide guidance in the selection of treatment strategies for CCD patients.

CONCLUSION

The present study investigated the imaging performance and long-term auditory speech outcomes of 23 children with CCD, who were found to have poorer auditory and speech development and slower progress after CI than the control group. However, CCD patients still showed improvement in auditory and speech abilities at 1.5 year after CI; hence, they required long-term rehabilitation. The reconstruction of the temporal bone CT showed that the size, volume, and morphology of the cavity in CCD patients varied widely, and a small lumen size is associated with shorter time needed to reach a 4-point post-operative CAP. Further studies should be conducted to verify these results and clarify the relationship and mechanism using a larger sample of patients with CCD.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Beijing Tongren Hospital, Capital Medical University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

LZ and BC performed experiments, acquired and analyzed the data, drafted and revised the manuscript. YK programmed the CI and critically revised the manuscript. NL performed experiments and revised the manuscript. XW and YS conceived and designed the study, revised the manuscript. JC and MY acquired data and revised the manuscript. AD reconstructed the CT image of CCD patients. YL supervised the study, interpreted the result and critically revised the manuscript. All authors contributed to the article and approved the final and submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant No. 81870716).

ACKNOWLEDGMENTS

We thank all the subjects for their participation in this study.

REFERENCES

Ahmad, R. L., Lokman, S., and Raja, A. A. K. (2005). Cochlear implantation in congenital cochlear abnormalities. *Med. J. Malaysia* 60, 379–382.

Ahn, J. H., Lim, H. W., and Lee, K. S. (2011). Hearing improvement after cochlear implantation in common cavity malformed cochleae: long-term follow-up results. *Acta Otolaryngol.* 131, 908–913. doi: 10.3109/00016489.2011.570786

Al-mahboob, A., Alhabib, S. F., Abdelsamad, Y., and Alzhrani, F. (2022). Cochlear implantation in common cavity deformity: a systematic review. *Eur. Arch. Oto-Rhino-Laryngol.* 279, 37–48. doi: 10.1007/s00405-021-06884-6885

Archbold, S., Lutman, M. E., and Nikolopoulos, T. (1998). Categories of auditory performance: inter-user reliability. *Br. J. Audiol.* 32, 7–12. doi: 10.3109/03005364000000045

Bae, S. H., Choi, J., and Choi, J. Y. (2022). Cochlear implants for patients with common cavity deformities and the impact of electrode positioning. *Clin. Exp. Otorhinolaryngol.* Online ahead of print. doi: 10.21053/ceo.2021.00745

Beltrame, M. A., Birman, C. S., Cervera Escario, J., Kassouma, J., Manolidis, S., Pringle, M. B., et al. (2013). Common cavity and custom-made electrodes: speech perception and audiological performance of children with common cavity implanted with a custom-made MED-EL electrode. *Int. J. Pediatr. Otorhinolaryngol.* 77, 1237–1243. doi: 10.1016/j.ijporl.2013.04.008

Beltrame, M. A., Bonfioli, F., and Frau, G. N. (2000). Cochlear implant in inner ear malformation: double posterior labyrinthotomy approach to common cavity. *Adv. Otorhinolaryngol.* 57, 113–119. doi: 10.1159/000059162

Beltrame, M. A., Frau, G. N., Shanks, M., Robinson, P., and Anderson, I. (2005). Double posterior labyrinthotomy technique: results in three Med-El patients with common cavity. *Otol. Neurotol.* 26, 177–182. doi: 10.1097/00129492-200503000-200503008

Bozorg Grayeli, A., Kalamarides, M., Bouccara, D., Ben Gamra, L., Ambert-Dahan, E., and Sterkers, O. (2007). Auditory brainstem implantation to rehabilitate profound hearing loss with totally ossified cochleae induced by pneumococcal meningitis. *Audiol. Neurotol.* 12, 27–30. doi: 10.1159/000096155

Brotto, D., Avato, I., Lovo, E., Muraro, E., Bovo, R., Trevisi, P., et al. (2019). Epidemiologic, imaging, audiologic, clinical, surgical, and prognostic issues in common cavity deformity: a narrative review. *JAMA Otolaryngol. Head Neck Surg.* 145, 72–78. doi: 10.1001/jamaoto.2018.2839

Buchman, C. A., Copeland, B. J., Yu, K. K., Brown, C. J., Carrasco, V. N., and Pillsbury, H. C. (2004). Cochlear implantation in children with congenital inner ear malformations. *Laryngoscope* 114, 309–316. doi: 10.1097/00005537-200402000-200402025

Casselman, J. W., Offeciers, E. F., De Foer, B., Govaerts, P., Kuhweide, R., and Somers, T. (2001). CT and MR imaging of congential abnormalities of the inner ear and internal auditory canal. *Eur. J. Radiol.* 40, 94–104. doi: 10.1016/S0720-048X(01)00377-371

Colletti, V., Carner, M., Miorelli, V., Guida, M., Colletti, L., and Fiorino, F. (2005). Auditory brainstem implant (ABI): new frontiers in adults and children. *Otolaryngol. Head Neck Surg.* 133, 126–138. doi: 10.1016/j.otohns.2005.03.022

Dettman, S., Sadeghi-Barzalighi, A., Ambett, R., Dowell, R., Trotter, M., and Briggs, R. (2011). Cochlear implants in forty-eight children with cochlear and/or vestibular abnormality. *Audiol. Neurotol.* 16, 222–232. doi: 10.1159/000320608

Dhanasingh, A., Dietz, A., Jolly, C., and Roland, P. (2019). Human inner-ear malformation types captured in 3d. *J. Int. Adv. Otol.* 15, 77–82. doi: 10.5152/iao.2019.6246

Fayad, J. N., and Linthicum, F. H. (2006). Multichannel cochlear implants: relation of histopathology to performance. *Laryngoscope* 116, 1310–1320. doi: 10.1097/01.mlg.0000227176.09500.28

Jackler, R. K., Luxford, W. M., and House, W. F. (1987). Congenital malformations of the inner ear: a classification based on embryo genesis. *Laryngoscope* 97, 2–14. doi: 10.1002/lary.5540971301

Liao, S., (2018). E-3D Medical Software V17.80 (Central South University, Changsha C. e3d-med n.d. Available online at: http://www.e3d-med.com/

Ma, H., Han, P., Liang, B., Tian, Z., Lei, Z., Kong, W., et al. (2008). Multislice spiral computed tomography imaging in congenital inner ear malformations. *J. Comput. Assist. Tomogr.* 32, 146–150. doi: 10.1097/rct.0b013e318063c64a

McElveen, J. T., Carrasco, V. N., Miyamoto, R. T., and Linthicum, F. H. (1997). Cochlear implantation in common cavity malformations using a transmastoid labyrinthotomy approach. *Laryngoscope* 107, 1032–1036. doi: 10.1097/00005537-199708000-00005

Molter, D. W., Pate, B. R., and Mcelveen, J. T. (1993). Cochlear implantation in the congenitally malformed Ear. *Otolaryngol. Neck Surg.* 108, 174–177. doi: 10.1177/019459989310800212

Papsin, B. C. (2005). Cochlear implantation in children with anomalous cochleovestibular anatomy. *Laryngoscope* 115, 1–26. doi: 10.1097/00005537-200501001-200501001

Robbins, A. M., Renshaw, J. J., and Berry, S. W. (1991). Evaluating meaningful auditory integration in profoundly hearing-impaired children. *Am. J. Otol.* 12, 144–150.

Sennaroğlu, L., and Bajin, M. D. (2017). Classification and current management of inner ear malformations. *Balkan Med. J.* 34, 397–411. doi: 10.4274/balkanmedj.2017.0367

Sennaroglu, L., Atay, G., and Bajin, M. D. (2014). A new cochlear implant electrode with a "cork"-type stopper for inner ear malformations. *Auris Nasus Larynx* 41, 331–336. doi: 10.1016/j.anl.2013.12.011

Sennaroğlu, L., Sennaroğlu, G., Yücel, E., Bilginer, B., Atay, G., Bajin, M. D., et al. (2016). Long-term results of abi in children with severe inner ear malformations. *Otol. Neurotol.* 37, 865–872. doi: 10.1097/MAO.0000000000001050

Shi, Y., Li, Y., Gong, Y., Chen, B., and Chen, J. (2019). Cochlear implants for patients with inner ear malformation: experience in a cohort of 877 surgeries. *Clin. Otolaryngol.* 44, 702–706. doi: 10.1111/coa.13360

Skinner, M. W., Ketten, D. R., Holden, L. K., Harding, G. W., Smith, P. G., Gates, G. A., et al. (2002). CT-derived estimation of cochlear morphology and electrode array position in relation to word recognition in nucleus-22 recipients. *JARO - J. Assoc. Res. Otolaryngol.* 3, 332–350. doi: 10.1007/s101620020013

Toh, E. H., and Luxford, W. M. (2008). Cochlear and Brainstem Implantation. *Neurosurg. Clin. N. Am.* 19, 317–329. doi: 10.1016/j.nec.2008.02.014

Tucci, D. L., Telian, S. A., Zimmerman-Phillips, S., Zwolan, T. A., and Kileny, P. R. (1995). Cochlear implantation in patients with cochlear malformations. *Arch. Otolaryngol. Neck Surg.* 121, 833–838. doi: 10.1001/archotol.1995.01890080005001

Verbist, B. M., Skinner, M. W., Cohen, L. T., Leake, P. A., James, C., Boëx, C., et al. (2010). Consensus panel on a cochlear coordinate system applicable in histologic, physiologic, and radiologic studies of the human cochlea. *Otol. Neurotol.* 31, 722–730. doi: 10.1097/MAO.0b013e3181d279e0

Wei, X., Li, Y., Chen, B., Gong, Y., Fu, Q. J., Liu, T., et al. (2017). Predicting auditory outcomes from radiological imaging in cochlear implant patients with cochlear nerve deficiency. *Otol. Neurotol.* 38, 685–693. doi: 10.1097/MAO.0000000000001382

Wei, X., Li, Y., Fu, Q., Gong, Y., Chen, B., Chen, J., et al. (2018). Slotted labyrinthotomy approach with customized electrode for patients with common cavity deformity. *Laryngoscope* 128, 468–472. doi: 10.1002/lary.26627

Xia, J., Wang, W., and Zhang, D. (2015). Cochlear implantation in 21 patients with common cavity malformation. *Acta Otolaryngol.* 135, 459–465. doi: 10.3109/00016489.2014.990054

Conflict of Interest: AD was employed by MED-EL Medical Electronics GmbH.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Neuroscience

ORIGINAL RESEARCH
published: 08 June 2022
doi: 10.3389/fnins.2022.904724

Relationship Between the Ability to Detect Frequency Changes or Temporal Gaps and Speech Perception Performance in Post-lingual Cochlear Implant Users

Dianzhao Xie, Jianfen Luo, Xiuhua Chao, Jinming Li, Xianqi Liu, Zhaomin Fan, Haibo Wang and Lei Xu*

Department of Otolaryngology-Head and Neck Surgery, Shandong Provincial ENT Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China

OPEN ACCESS

Edited by:
Fei Chen,
Southern University of Science and Technology, China

Reviewed by:
Jorge Humberto Ferreira Martins,
Coimbra Hospital and University Center, Portugal
Mustafa Yüksel,
Ankara Medipol University, Turkey

****Correspondence:***
Lei Xu
sdphxl@126.com

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *25 March 2022*
Accepted: *17 May 2022*
Published: *08 June 2022*

Citation:
Xie D, Luo J, Chao X, Li J, Liu X, Fan Z, Wang H and Xu L (2022) Relationship Between the Ability to Detect Frequency Changes or Temporal Gaps and Speech Perception Performance in Post-lingual Cochlear Implant Users. Front. Neurosci. 16:904724. doi: 10.3389/fnins.2022.904724

Previous studies, using modulation stimuli, on the relative effects of frequency resolution and time resolution on CI users' speech perception failed to reach a consistent conclusion. In this study, frequency change detection and temporal gap detection were used to investigate the frequency resolution and time resolution of CI users, respectively. Psychophysical and neurophysiological methods were used to simultaneously investigate the effects of frequency and time resolution on speech perception in post-lingual cochlear implant (CI) users. We investigated the effects of psychophysical results [frequency change detection threshold (FCDT), gap detection threshold (GDT)], and acoustic change complex (ACC) responses (evoked threshold, latency, or amplitude of ACC induced by frequency change or temporal gap) on speech perception [recognition rate of monosyllabic words, disyllabic words, sentences in quiet, and sentence recognition threshold (SRT) in noise]. Thirty-one adult post-lingual CI users of Mandarin Chinese were enrolled in the study. The stimuli used to induce ACCs to frequency changes were 800-ms pure tones (fundamental frequency was 1,000 Hz); the frequency change occurred at the midpoint of the tones, with six percentages of frequency changes (0, 2, 5, 10, 20, and 50%). Temporal silences with different durations (0, 5, 10, 20, 50, and 100 ms) were inserted in the middle of the 800-ms white noise to induce ACCs evoked by temporal gaps. The FCDT and GDT were obtained by two 2-alternative forced-choice procedures. The results showed no significant correlation between the CI hearing threshold and speech perception in the study participants. In the multiple regression analysis of the influence of simultaneous psychophysical measures and ACC responses on speech perception, GDT significantly predicted every speech perception index, and the ACC amplitude evoked by the temporal gap significantly predicted the recognition of disyllabic words in quiet and SRT in noise. We conclude that when the ability to detect frequency changes and the temporal gap is considered simultaneously, the ability to detect frequency changes may have no significant effect on speech perception, but the ability to detect temporal gaps could significantly predict speech perception.

Keywords: cochlear implant, frequency change detection, temporal gap detection, speech perception, psychophysical test, acoustic change complex

INTRODUCTION

For patients with severe-to-profound hearing loss, cochlear implantation (CI) is the most effective method for reconstructing hearing. While overall speech signal understanding has improved, there remains variability in performance across recipients, and speech perception in noise remains challenging (Firszt et al., 2004; Wilson and Dorman, 2008; Holden et al., 2013).

Understanding daily conversation depends on the ability of the auditory system to detect ongoing changes in the spectral and temporal patterns of the incoming signals (He et al., 2012). Unlike individuals with normal hearing who have approximately 3,500 inner hair cells and 12,000 outer hair cells to provide fine-grained spectral resolution, CI users rely on sound information conveyed by electrical stimulation through up to 22 electrodes (Liang et al., 2018). The real number of spectral channels used by most CI users is likely to be less than eight because of factors such as channel interactions and frequency-to-electrode mismatches (Fu et al., 2004). In addition, owing to signal processing (e.g., signal compression, bandpass filtering, and temporal envelope extraction), CI greatly attenuates the time-frequency information of sound. Furthermore, neural degeneration related to long-term deafness in CI users exacerbates their compromised ability to detect frequency differences in sound (Sek and Moore, 1995; Moore, 1996).

Exploring the influencing factors of CI users' speech perception has always been the interest of researchers. In terms of acoustical frequency resolution or time resolution of CI users, significant correlations were reported between spectral modulation sensitivity and speech perception outcomes for CI users (Henry et al., 2005; Won et al., 2007; Anderson et al., 2012), and there were significant correlations between speech perception performance and temporal modulation detection performance measured either through sound processor (Won et al., 2011; Gnansia et al., 2014) or direct stimulation in CI users (Cazals et al., 1994; Fu, 2002). Previous study have confirmed that the ability to detect spectrotemporal modulation, covaried in both the temporal and spectral domains, was related to CI users' speech recognition performance (Lawler et al., 2017). On this basis, some researchers (Won et al., 2015; Zhou N. et al., 2020) evaluated the correlation of speech recognition with the spectrotemporal modulation (STM) thresholds while controlling for either temporal or spectral modulation sensitivity, but different conclusions were drawn. Won et al. (2015) suggested that that slow spectral modulation rather than slow temporal modulation may be important for determining speech perception capabilities for CI users. However, Zhou N. et al. (2020) suggested that temporal information processing may limit performance more than spectral information processing in CI users. Considering the fact that similar method was applied but reached different conclusions, this study intended to use different methods to simultaneously investigate the frequency resolution and time resolution of CI users, and analyze their relative roles in predicting speech perception.

There are different approaches to measuring frequency discrimination. One of the approaches is frequency change detection, used here, examining detection of minimal frequency change within stimuli that have embedded frequency changes. The advantage of this approach is that it allows for the examination of neural response evoked by the frequency change (e.g., acoustic change complex, ACC) within stimuli (Zhang et al., 2019). Measurement of gap detection thresholds (GDTs), used here, is one of the most widely used methods for assessing temporal resolution in humans (Garadat and Pfingst, 2011; Lister et al., 2011).

Some studies have investigated the relationships between speech perception and frequency change detection or temporal gap detection. In terms of frequency change detection, a study on adult CI patients confirmed that the frequency change detection threshold (FCDT) is related to speech perception ability (Zhang et al., 2019). According to the FCDT results, the adult CI was divided into two groups: good and poor. The speech test results of the good CI group were significantly better than those of the poor CI group (McGuire et al., 2021). In a study of changing stimulated electrodes, there was a robust correlation between electrode-discrimination capacities and speech-perception performance in CI children with auditory neuropathy spectrum disorder (ANSD) (He et al., 2014). In terms of temporal gap detection, one study in young people with normal hearing and old people with hearing loss showed that, after excluding the influence of age and hearing loss, the GDT contributed to variance in speech recognition in noise (Hoover et al., 2015). However, some studies have failed to confirm the correlation between GDT and speech perception ability of CI users (Mussoi and Brown, 2019; Cesur and Derinsu, 2020). Luo et al. (2020) drew different conclusions for various subjects. For older acoustic-hearing listeners, gap detection ability was significantly correlated with SRT in noise, but this correlation was not observed in older CI users and younger listeners with normal hearing. Therefore, the relationship between gap-detection ability and speech perception requires further study.

The above FCDT or GDT are obtained by psychophysical tests. However, clinically, some CI users would not be able to participate in complicated auditory tests; therefore, it is necessary to find simple test methods to quickly estimate or predict the effects of CI. Cortical auditory evoked potentials (CAEPs), which can be recorded in a passive listening condition that does not require the participant's attention or voluntary responses, can thus serve as a suitable tool for difficult-to-test participants (Sharma and Dorman, 2006; Small and Werker, 2012). The auditory event–related potentials (ERPs), including the onset response and the acoustic change complex (ACC), are cortically generated potentials that can be recorded from surface electrodes placed on the scalp. The onset response is typically evoked by a brief stimulus, and its presence indicates sound detection. The ACC is elicited by stimulus change(s) that occur within an ongoing, long-duration stimulation. The ACC provides evidence of discrimination capacity across various stimulus dimensions at the level of the auditory cortex (Martin et al., 2008). The ACC differs from and has advantages over the mismatch negativity (MMN), another type of auditory evoked response reflecting auditory discrimination. First, in the stimulus paradigm for the ACC, every trial of the stimuli contributes to the ACC response. In MMN recordings, a large number of standard stimuli is required to embed a sufficient number of deviant stimuli. Thus, the recording time for the ACC is much shorter than that for the MMN. Second, the ACC has a much larger amplitude (higher signal-to-noise ratio) compared to the MMN, which enables the accurate identification of ACC peaks for latency and amplitude calculation (Martin and Boothroyd, 1999). Third, the MMN is an outcome of waveform subtraction between the response to the deviant stimuli and the response to the standard stimuli, while the ACC is a response directly collected from the participant (Liang et al., 2018).

There were also some studies on the relationship between ACC induced by frequency change or temporal gap and speech perception. On ACC induced by frequency change, the N1 latency of the ACC induced by the 160-Hz tone containing a 50% frequency change was significantly correlated with the clinically collected phonetic perception score [consonant-nucleus-consonant (CNC) monosyllabic word], although a correlation between N1 latency and AzBio sentences could not be established (Liang et al., 2018). In a study of changing stimulated electrodes, compared with those with poor speech performance, the electrically evoked auditory change complex (eACC) amplitude of those with better speech performance was larger (He et al., 2014). On ACC induced by temporal gap, studies of CI or non-CI children with ANSD showed that the eACC or ACC induced by temporal gap was significantly correlated with the phonetically balanced kindergarten (PBK) word score (He et al., 2013, 2015). Unlike the above studies, a study on people with normal hearing and hearing loss showed that, after considering the influence of hearing loss, there was no significant correlation between the ACC threshold induced by frequency change and speech reception thresholds (SRTs) in noise (Vonck et al., 2021). Most of the aforementioned studies that reported a relationship between frequency change detection and speech perception did not consider the influence of hearing threshold. Therefore, the relationship between subjective or objective frequency change detection and speech recognition requires further study after excluding the influence of hearing threshold.

In addition to failing to reach a consistent conclusion, most of the aforementioned studies did not investigate frequency change detection and temporal gap detection simultaneously; therefore, it is impossible to analyze their relative effects on speech perception. Furthermore, few studies have simultaneously used psychophysical and neurophysiological methods to investigate the effects of frequency change detection and temporal gap detection on CI users' speech perception. These were exactly what this research wanted to do.

For tonal languages, such as Mandarin Chinese, lexical tones make an essential contribution to understanding the meaning of words and sentences. Mandarin Chinese includes four tones: the high-level tone (tone 1), the rising tone (tone 2), the falling-rising tone (tone 3), and the high-falling tone (tone 4). These tones play an important role in understanding the meaning of monosyllabic words in Chinese language (Zhou Q. et al., 2020). Some studies have examined the role of temporal and spectral cues in mandarin tone recognition (Kong and Zeng, 2006; Wei et al., 2007), but few studies have examined the influence of frequency resolution and time resolution on CI users' speech perception in Mandarin Chinese.

This study addresses the following questions: (1) whether the speech perception ability of post-lingual CI users is affected by their CI hearing threshold; (2) whether FCDT or GDT obtained using psychophysical tests could predict speech perception in post-lingual CI users of Mandarin Chinese; (3) whether ACC response induced by frequency change or temporal gap can predict speech perception in these CI users; and (4) when considering both psychophysical and neurophysiological results of frequency change detection or temporal gap detection, which factors could best predict speech perception in these CI users.

MATERIALS AND METHODS

Participants

There were 31 CI users (11 females and 20 males; 16.3–51.4 years old; 27 unilateral and four bilateral CI users) participated in this study. Only the more satisfied side was tested in bilateral CI users, whereas the other was picked off. Bimodal CI users, who wore a hearing aid in the non-implanted ear, took off their hearing aid and had an earplug inserted into their non-implanted ear. All participants were post-lingually deafened, with a speech intelligibility rating (SIR) score above 4 (connected speech is intelligible to a listener who has little experience of a deaf person's speech; the listener does not need to concentrate unduly). Considering that the purpose of this study was to find the relationship between CI users' speech perception with auditory discrimination factors, and the speech perception of prelingually deafed adult CI users was greatly constrained, the participant was limited to post-lingually deafed CI users. All patients used CI for at least 6 months. All participants were native Mandarin Chinese speakers with no history of neurological or psychological disorders. Demographic data of the participants are presented in **Table 1**. Twenty-one participants used Chinese CIs.

This study was approved by the Medical Ethics Committee of Shandong Provincial ENT Hospital, Shandong, China.

TABLE 1 | Cochlear implant (CI) users' demographics.

CI user	Gender	Type of CI user	Age	Ear tested	Device	Duration of severe-to-profound deafness (yr)	Age at implantation	Duration of CI use (m)
01	M	Unilateral	51.88	R	Nurotron/CS-10A*	4	51.09	9.4
02	M	Unilateral	38.89	R	Nurotron/CS-10A*	10	38.21	8.9
03	M	Unilateral	39.93	R	Nurotron/CS-10A*	3	39.44	6.02
04	F	Unilateral	34.02	L	Nurotron/CS-10A*	10	33.26	9.17
05	F	Unilateral	46.42	R	Nurotron/CS-10A*	8	45.86	6.87
06	M	Unilateral	43.24	L	Nurotron/CS-10A*	3	42.49	9.79
07	M	Unilateral	18.99	L	Nurotron/CS-10A*	10	18.4	7.79
08	F	Unilateral	25.05	L	Nurotron/CS-10A*	10	24.36	9.13
09	M	Unilateral	35.82	R	Nurotron/CS-10A*	24	35.36	6.27
10	M	Unilateral	42.37	L	Nurotron/CS-10A*	5	41.83	7.43
11	M	Unilateral	47.98	L	Listent/LCI-20PI*	1	47.47	6.23
12	M	Bilateral	29.25	L	Med El/Sonata	22	28.72	6.37
13	M	Unilateral	35.16	L	Nurotron/CS-10A*	7	34.67	6.01
14	F	Unilateral	29.29	R	Nurotron/CS-10A*	2	28.75	6.41
15	F	Unilateral	29.43	R	Nurotron/CS-10A*	1	28.93	6.02
16	M	Unilateral	36.09	L	Nurotron/CS-10A*	6	35.59	6.05
17	M	Bilateral	34.79	R	Nurotron/CS-10A*	1	34.29	6.01
18	F	Unilateral	24.02	L	Nurotron/CS-10A*	13	23.49	6.34
19	M	Bilateral	37.14	L	Nurotron/CS-10A*	26	36.62	6.28
20	F	Unilateral	50.35	R	Nurotron/CS-10A*	4	49.79	6.8
21	M	Unilateral	29.78	R	Nurotron/CS-10A*	9	29.24	6.51
22	F	Unilateral	18.62	R	Med El/Sonata	11	17.69	11.24
23	M	Bilateral	19.13	L	Med El/Sonata	1	9.27	118.31
24	M	Unilateral	40.33	L	Nucleus/CI522	2	37.32	36.14
25	F	Unilateral	17.39	L	Nucleus/CI422	13	15.3	25.07
26	M	Unilateral	20.68	L	Nucleus/CI24RE(CA)	1	19.88	9.56
27	M	Unilateral	23.85	R	Listent/LCI-20PI*	1	23.04	9.69
28	F	Unilateral	25.92	L	Nucleus/CI24RE(CA)	1	24.88	12.42
29	F	Unilateral	16.25	L	Med El/Sonata	13	15.43	9.86
30	M	Unilateral	25.9	L	Nucleus/CI24RE(CA)	1	24.88	12.19
31	M	Unilateral	45.6	R	Med El/Sonata	29	45.08	6.18

**Nurotron and Listent were two Chinese domestic cochlear implant brands.*

All participants provided written informed consent before participating in the study.

Stimuli

Stimuli in psychophysical tests and CAEPs tests were generated using Audacity software at a sample rate of 44.1 kHz and presented by the E-Prime program (Psychology Software Tools, Pittsburgh, PA, United States).

A series of tones of 800 ms duration (including 10-ms raised-cosine onset and offset ramps) at f_{base} of 1,000 kHz that contained different magnitudes of upward F-changes at 400 ms after the tone onset were used in frequency discrimination tests. The F-change occurred at 0 phase (zero crossing); there was no audible transient when the F-change occurred (Dimitrijevic et al., 2008; Pratt et al., 2009). The amplitudes of all the stimuli were normalized. Similar stimuli were used in some other studies (Liang et al., 2018, 2020; Zhang et al., 2019; McGuire et al., 2021).

Compared with pure tone or narrow-band noise, broadband noise can activate more electrodes in CI electrode array. Therefore, white noise was used in this study, so as to investigate the overall gap detection ability of CI users. White noise with different durations of silent gaps added in the middle position was used in the gap detection tests. There were 10 ms rising and falling periods when white noise appeared and ended. We used the 4-ms fall/rise surrounding the gap to reduce the spectral splatter, which is usually introduced by rapid onsets and offsets. To minimize the availability of intensity cues resulting from the 4-ms fall/rise, both the gap and no-gap stimuli contained a 4-ms fall/rise. GDT was affected by the duration of stimulation before and after the gap, and gap detection thresholds of older adults were markedly higher than those of younger adults for marker durations of less than 250 ms (Schneider and Hamstra, 1999). In this study, the influence of age should be excluded, so the duration of noise before and after the gap should be longer than 250 ms. In this study, the stimulation duration before and after gap was 400 ms, which was also applied in some other studies (He et al., 2013, 2015, 2018; Mussoi and Brown, 2019). In GDT testing, the stimulation duration was fixed at 800 ms, and the gap was in the middle position. In ACC testing, the duration of stimulation ranged from 800 to 900 ms, in which the gap occurred at 400 ms after the noise onset, and gap duration ranged from 0 to 100 ms.

Studies have shown that GDT was affected by the intensity of stimulation, and gap detection was known to improve with increases in stimulus intensity level until asymptotic performance was achieved (Florentine and Buus, 1984; Horwitz et al., 2011). For people with normal hearing, to achieve a stable and high level of GDT, the noise intensity should be above 50 dB SPL (Florentine and Buus, 1984) or above 20 dB SL (Horwitz et al., 2011). In this study, in order to ensure that the subjects' GDT reaches their own high level, the noise intensity was set to not less than 70 dB SPL. Patients who were unable to tolerate sounds of 70 dB SPL were excluded from these tests.

Procedures

The participants were tested for pulse tone hearing thresholds to ensure audibility of the stimuli presented through their clinical processors. Their speech performance was tested with 35 dB HL sound intensity above their CI hearing thresholds (average of 0.25, 0.5, 1, 2, 4 kHz pulse tone hearing threshold). They were seated on a comfortable chair in a sound-treated booth for the psychophysical and CAEP tests. Stimuli were presented in the sound field *via* a single loudspeaker placed at ear level, 1.5 m in front of the participant. The stimuli were presented at an intensity corresponding to loudness level 7 (most comfortable level) on a 0–10-point (inaudible to too loud) numerical scale to the tested CI ear (Liang et al., 2018). The intensity level of stimuli was determined separately in frequency change detection and gap detection, so the stimuli intensity level of frequency change detection and gap detection may be different for one subject. The same stimulus intensity was used for both the psychophysical and CAEP tests in frequency change detection or temporal gap detection, making the results of psychophysical test and CAEP test be comparable.

Behavioral Tasks

Psychophysical Tests

An adaptive, 2-alternative forced-choice procedure was used to determine the FCDT and GDT. In each trial, a standard stimulus and target stimulus were included, and the participant was instructed to choose the target stimulus by pressing the corresponding button. The order of the standard and target stimuli was randomized and the interval between the stimuli in each trial was 0.5 s. A 2-down, 1-up staircase technique was used to track the 70.7% correct point on the psychometric function. Each response alteration was counted as response reversal. Each run generated 10 reversals. FCDT or GDT was calculated as the average of the last six reversals. The test was repeated thrice, and the average of the FCDTs or GDTs was recorded. The order of FCDT and GDT was random among different subjects.

The standard stimulus in the FCDT test was a 1,000 Hz tone, with no frequency change; the target stimulus was a 1,000 Hz tone containing a frequency change with a magnitude of up to 100%; the step size was 5% from 10 to 100% range, 0.5% from 0.5 to 10% range, and 0.05% from 0.05 to 0.5% range. The change of frequency began at 20%. The standard stimulus in the GDT test was white noise with no gap inserted, and the target stimulus was white noise with a gap inserted, in which the maximum gap duration was 100 ms; the step size was 5 ms from the 40 to 100 ms range, 2 ms from the 10 to 40 ms range, and 1 ms from the 1 to 10 ms range. The initial gap duration was 20 ms.

Speech Perception Tests

A computer-assisted Chinese speech audiometry platform was used to test speech perception (Xin et al., 2010). The recognition accuracy for monosyllabic words, disyllable words, and sentences in a quiet environment was tested. The SRT in noise (the SNR required for 50% correct word-in-sentence recognition in multi-talker, speech-babble noise) was only tested in participants whose recognition accuracy of a sentence in quiet exceeded 50%. In the monosyllabic words test, there were 25 syllables, and only when the consonants, vowels and tones were all correctly identified could the monosyllabic word be regarded as correctly identified. There were 40 disyllabic words in the disyllabic word test, and a single word was used as the scoring unit in the test. The sentence test in quiet consisted of 10 sentences with 50 key words, and the key word was used as the scoring unit in the test. In SRT test in noise, the initial SNR was set before test and the software could obtain the SNR corresponding to 50% correct recognition in noise by self-adapting SNR. In this study, the SRT in noisy was tested in 18 subjects. In this study, speech perception refers to the overall speech recognition ability, including the recognition accuracy of monosyllabic words, disyllabic words and sentences in quiet and the SRT in noise.

Electroencephalographic Recordings and Data Processing

The stimuli used to induce the F-change CAEPs were tones at f_{base} of 1 kHz containing six different percentages of F-changes (0, 2, 5, 10, 20, and 50%). The stimuli used to induce the temporal gap CAEPs were white noise with six different gap durations (0, 5, 10, 20, 50, and 100 ms). The six stimulus conditions in tone or white noise were randomized to prevent order effects. The inter-stimulus interval was 1,200 ms. The order of the F-change and temporal gap CAEP tests was random among the participants.

The participants were seated comfortably in chairs and invited to watch silent films with subtitles. They did not need to pay attention to the sound stimulation presented *via* the loudspeaker but needed to stay awake and quiet. Electroencephalogram (EEG) results were collected using a Brain Vision (1.22) system (London, United Kingdom) and a Brain Amp DC amplifier. According to the International Standard 10–20 system, FPz, Fz, Cz, C3, and C4 were used as recording electrodes, the electrode placed at the opposite mastoid of the implanted side was used as the reference, the electrode placed between the eyebrows was the grounding electrode, and the electrode placed under the opposite eyes of the implanted side was used to record eye blinks. The electrode impedances were maintained at below 5 kΩ. The EEG was sampled at 5,000 Hz and filtered online between 1 and 100 Hz. The artifact rejection threshold was $\pm$ 120 μV. The EEG was epoched and baseline-corrected online using a window of 1,100 ms, including a 100-ms pre-stimulus baseline and a 1,000-ms post-stimulus time. For each subject, at least 200 artifact-free sweeps were recorded for each stimulation condition. These recordings were digitally filtered offline between 1 and 35 Hz before response identification and amplitude measurements.

Time windows delimiting the possible occurrence of ACC responses were determined based on the grand mean average of all recorded responses of F-change or temporal gap CAEPs. The windows for the ACC response extended from 450 to 650 ms for F-change CAEPs and from 450 to 750 ms for temporal gap CAEPs, relative to the stimulus onset. The presence of ACC was determined on ERPs based on the following criteria: (i) an expected ACC wave morphology (N1-P2 complex) within the expected time window, and (ii) a visual difference in the waveforms between the F-change conditions vs. the no-change condition or gap-inserted conditions vs. no gap-inserted conditions. Finally, the peak components of ACC (N1 and P2) were labeled. In the ACC response, the first peak, P1, is considerably smaller than the N1 and P2 peaks. The low signal-to-noise ratio of this peak makes it difficult to reliably determine the amplitude and latency of P1 (Vonck et al., 2021). The N1 potential was used to represent the ACC potential. As in previous studies (Tremblay et al., 2001; Kim et al., 2009), the N1-P2 peak-to-peak amplitude was used to represent the amplitude of the ACC.

Statistical Analysis

The frequency change and temporal gap ACC thresholds were separately defined as the smallest frequency change and the shortest temporal gap in six change conditions that could be reliably used to evoke ACC responses. Among the six stimulus conditions, the latency and amplitude of ACC evoked by 50% frequency change and 100 ms temporal gap at Fz were used for statistical analysis. Statistical analyses were performed using the SPSS version 20.0 software (IBM, Armonk, NY, United States).

Linear regression was performed to determine whether speech perception could be predicted by FCDT/GDT, frequency change ACC thresholds/temporal gap ACC threshold, ACC latency, or amplitude induced by frequency change or temporal gap. In the cases of several significant correlations between speech perception and other measures, additional multiple stepwise regressions were conducted to assess their relative contributions to speech perception. A p-value less than 0.05 was used to determine factor entry into the regression model, while a p-value greater than 0.10 was used to determine factor removal from the model. The final model of the stepwise regression excluded any factors whose removal did not significantly impact the model fit while including any factors whose addition significantly improved the model fit. We checked for collinearity between independent variables [defined as tolerance < 0.1 and variance inflation factor (VIF) > 10] (Vonck et al., 2021).

RESULTS

Correlation of Speech Perception to Cochlear Implant Hearing Thresholds

Figure 1 shows the means and standard deviations of the hearing thresholds using pulse tones at frequencies of 0.25, 0.5, 1, 2, and 4 kHz. The thresholds at the different frequencies were significantly different [$F_{(4,\ 120)} = 24.09$; $p < 0.01$]. A one-way repeated analysis of variance showed that the hearing threshold of 250 Hz was significantly lower than that of the remaining four frequencies ($p < 0.01$), that of 500 Hz was significantly lower than that of 1k/2k/4k ($p < 0.01$), that of 1 kHz was significantly lower than that of 2k/4k ($p < 0.01$), and that there was no significant difference between 2k and 4k ($p > 0.05$). The correlations between speech perception and hearing thresholds of 0.25, 0.5, 1, 2, and 4 kHz or the average hearing threshold of the five frequencies were analyzed. Our results suggest no significant correlation between speech perception and hearing threshold ($p > 0.05$).

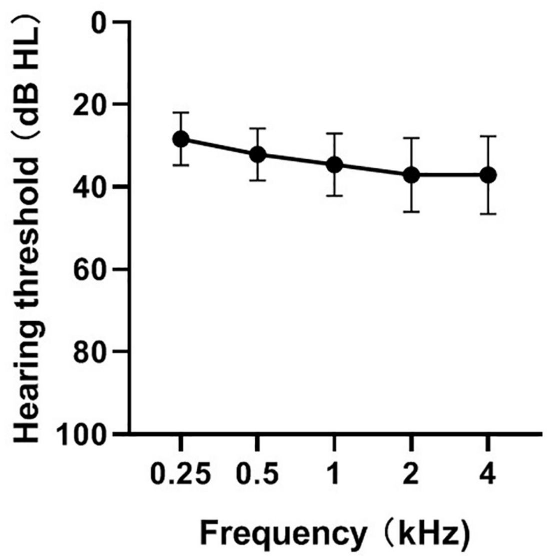

FIGURE 1 | Mean hearing thresholds at frequencies of 0.25, 0.5, 1, 2, 4 kHz in CI ears ($n = 31$). The mean (circle) and the standard deviations (error bars) of the mean are plotted.

Correlation of Speech Perception to Psychophysical Measures

In the simple linear regression analysis with FCDT or GDT as the independent variable and speech perception indicators as dependent variables, GDT was significantly correlated with all speech perception indicators ($p < 0.01$). FCDT was significantly correlated with SRT in noise ($p < 0.01$), but not with other speech recognition indicators ($p > 0.05$). The results are shown in **Figure 2**.

Correlation of Speech Perception to Acoustic Change Complex Thresholds

In the simple linear regression analysis with the frequency change ACC threshold or the temporal gap ACC threshold as the independent variable and speech perception indicators as dependent variables, the frequency change ACC threshold was significantly correlated with the recognition rates of disyllable

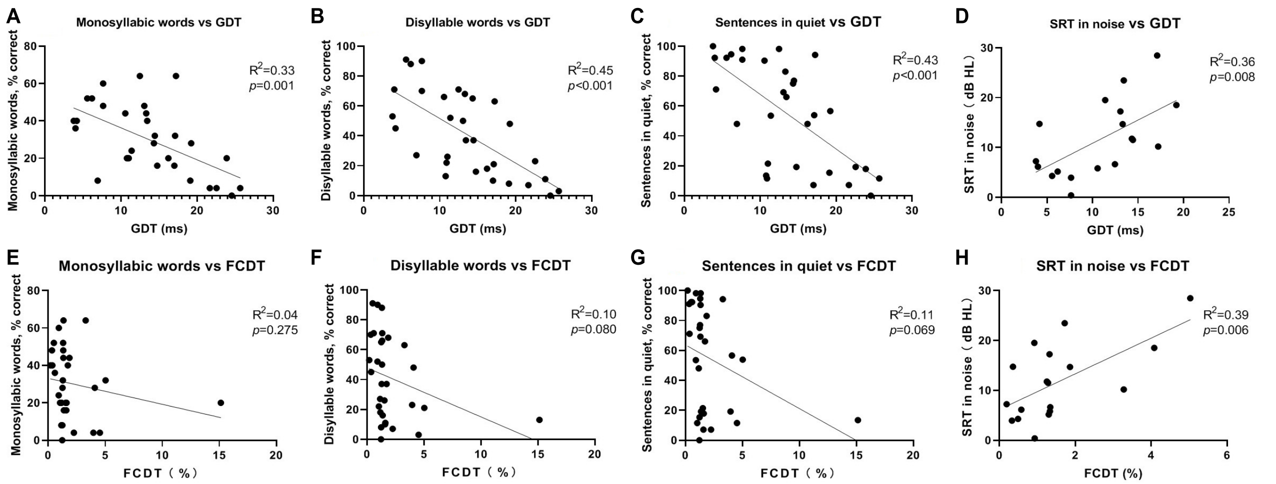

FIGURE 2 | The accuracy rate of monosyllabic words **(A)**, disyllable words **(B)**, sentences in quiet **(C)** as function of GDT for all subjects. The SRT in noise **(D)** as function of GDT for partial subjects. The accuracy rate of monosyllabic words **(E)**, disyllable words **(F)**, sentences in quiet **(G)** as function of FCDT for all subjects. The SRT in noise **(H)** as function of FCDT for partial subjects.

words ($r = -0.41$, $p < 0.05$) and sentences ($r = -0.42$, $p < 0.05$) in quiet, and with SRT in noise ($r = 0.57$, $p < 0.05$), but not with monosyllabic words ($r = -0.27$, $p > 0.05$) in quiet (**Table 2A**). The temporal gap ACC threshold was significantly correlated with the recognition rates of monosyllabic words ($r = -0.43$, $p < 0.05$), disyllable words ($r = -0.44$, $p < 0.05$), and sentences ($r = -0.50$, $p < 0.01$) in quiet, but not with SRT in noise ($r = 0.09$, $p > 0.05$) (**Table 2B**).

Correlation of Speech Perception to Potentials and Amplitudes of Acoustic Change Complex

In simple linear regression analysis with the potential or amplitude of ACC evoked by 50% F-change or 100 ms temporal gap as the independent variable and speech perception indicators as the dependent variables, the 50% F-change ACC potential had no significant correlation with any speech perception indicators ($p > 0.05$) (**Table 2C**); 50% F-change ACC amplitude was significantly correlated with recognition rates of disyllable words ($r = 0.38$, $p < 0.05$) and sentences ($r = 0.39$, $p < 0.05$) in quiet (**Table 2D**); the 100 ms gap ACC potential was significantly correlated with recognition rates of disyllable words ($r = -0.40$, $p < 0.05$) and sentences ($r = -0.40$, $p < 0.05$) in quiet (**Table 2E**); and 100 ms gap ACC amplitude was significantly correlated with recognition rates of disyllable words ($r = 0.45$, $p < 0.05$) in quiet and SRT in noise ($r = -0.62$, $p < 0.01$) (**Table 2F**).

Multiple Regression Analyses of Speech Perception to Psychophysical Measures and Acoustic Change Complex Responses

To simultaneously consider the influence of psychophysical measures, ACC thresholds, potentials, and amplitudes of ACCs on speech perception, multiple stepwise regression analyses were conducted (see **Table 3** for the results). In multiple stepwise regression, all factors related to speech perception in the simple linear regression analysis were treated as independent in the regression models.

In the multiple regression analysis with monosyllabic word recognition as the dependent variable, only GDT was included in the model, whereas the temporal gap ACC threshold was excluded (**Table 3A**). Therefore, the temporal gap ACC threshold no longer had a significant effect on monosyllabic word recognition once the effect of the GDT was considered.

In the multiple regression analysis with disyllable word recognition as the dependent variable, only GDT and 100 ms gap ACC amplitude were included in the model, whereas other factors were excluded (**Table 3B**). Therefore, other factors no longer had a significant effect on disyllable word recognition once the effects of the GDT and 100 ms gap ACC amplitude were considered.

In the multiple regression analysis with sentence recognition in quiet as the dependent variable, only GDT was included in the model, whereas other factors were excluded (**Table 3C**). Therefore, other factors no longer had a significant effect on sentence recognition in quiet once the effect of the GDT was taken into account.

In the multiple regression analysis with SRT in noise as the dependent variable, only GDT and 100 ms gap ACC amplitude were included in the model, whereas other factors were excluded (**Table 3D**). Therefore, other factors no longer had a significant effect on SRT in noise once the effects of the GDT and 100 ms gap ACC amplitude were taken into account.

No collinearity was found between the factors with multiple regressions for the abovementioned multiple regressions (tolerance > 0.9, VIF < 1.1).

TABLE 2 | Simple linear regression analysis.

	R	R^2	*p*
A. Frequency change ACC threshold vs. speech performance			
Monosyllabic words	−0.270	0.073	0.142
Disyllable words	−0.412	0.170	**0.021**
Sentences in quiet	−0.423	0.179	**0.018**
SRT in noise	0.571	0.326	**0.013**
B. Temporal gap ACC threshold vs. speech performance			
Monosyllabic words	−0.432	0.187	**0.015**
Disyllable words	−0.436	0.190	**0.014**
Sentences in quiet	−0.496	0.246	**0.005**
SRT in noise	0.086	0.007	0.735
C. 50% F_change ACC potential vs. speech performance			
Monosyllabic words	−0.008	0.00006	0.965
Disyllable words	−0.168	0.028	0.368
Sentences in quiet	−0.137	0.019	0.461
SRT in noise	−0.235	0.055	0.348
D. 50% F_change ACC amplitude vs. speech performance			
Monosyllabic words	0.288	0.083	0.116
Disyllable words	0.378	0.143	**0.036**
Sentences in quiet	0.385	0.148	**0.032**
SRT in noise	−0.155	0.024	0.538
E. 100ms gap ACC potential vs. speech performance			
Monosyllabic words	−0.343	0.118	0.074
Disyllable words	−0.400	0.160	**0.035**
Sentences in quiet	−0.403	0.162	**0.033**
SRT in noise	0.334	0.112	0.191
F. 100 ms gap ACC amplitude vs. speech performance			
Monosyllabic words	0.256	0.066	0.188
Disyllable words	0.445	0.198	**0.018**
Sentences in quiet	0.295	0.087	0.128
SRT in noise	−0.617	0.381	**0.008**

The bold numbers indicate $p < 0.05$.

DISCUSSION

The Speech Perception of Post-lingual Cochlear Implant Users Was Not Affected by Their Cochlear Implant Hearing Thresholds

In this study, there were significant differences in CI hearing thresholds at different frequencies, and the overall trend was that the hearing thresholds gradually increased from low to high. This may be related to the individual characteristics of the study participants. Some of the study participants had long-term hearing loss and they did not wear hearing aids or the hearing aids did not adequately compensate for high-frequency sounds; therefore, they were unable to hear high-frequency sounds when they suffered from hearing loss, which made them intolerant of high-frequency sounds. Given the patient's level of tolerance, the sensitivity to high-frequency sounds did not reach that to low frequency when mapped.

Research on people with normal hearing and hearing loss has shown that the degree of hearing loss has an important influence on speech perception (Vonck et al., 2021). In this study, correlation analyses of speech perception and different hearing frequency thresholds and mean CI hearing thresholds were conducted, but the effect of hearing thresholds on speech perception has not been confirmed. Therefore, the CI hearing threshold was excluded from the factors influencing speech perception. Nevertheless, the relationship between speech perception and CI hearing thresholds cannot be completely negated as there were no CI users with excessively poor CI hearing thresholds among the study participants. Therefore, we are only able to affirm that in the case of acceptable CI hearing thresholds, speech perception was not affected by the CI hearing threshold in post-lingual CI users.

Speech Perception of Post-lingual Cochlear Implant Users May Not Be Affected by the Ability to Detect Frequency Change

In simple linear regression analyses, FCDT was correlated with SRT in noise, and the ACC threshold or ACC amplitude evolved by frequency change was correlated with some speech perception indicators. However, FCDT, ACC threshold, and ACC amplitude were all excluded from the subsequent multiple regression models in which, besides FCDT and frequency change ACC, GDT, and temporal gap ACC were also treated as independent variables. This indicated that FCDT or frequency change ACC did not play a significant role in explaining differences in speech perception once the effect of the GDT or temporal gap ACC was considered. This result is not in line with previous research results. Previous studies have confirmed that speech perception in CI users is related to the FCDT (Zhang et al., 2019), spectral ripple discrimination (Luo et al., 2020), and electrode discrimination ability (He et al., 2014), and that speech perception is related to the ACC response induced by frequency changes or stimulation electrode changes (He et al., 2014; Liang et al., 2018). Except for the study on spectral ripple discrimination (Luo et al., 2020), none of the aforementioned studies (He et al., 2014; Liang et al., 2018; Zhang et al., 2019) examined the temporal resolution of the subjects simultaneously. Therefore, it is uncertain whether the relationship between frequency resolution and speech perception still exists in these studies when temporal and frequency resolution were taken into account at the same time. In addition, the aforementioned studies did not rule out the possible influence of CI hearing threshold on speech perception. A study reported that the degree of hearing loss had an important influence on speech perception, and that the correlation between frequency change ACC threshold and speech perception could mainly be explained by the degree of hearing loss (Vonck et al., 2021). Therefore, the correlation between subjective or objective frequency resolution and speech perception of CI users needs to be further studied, considering both temporal resolution and CI hearing threshold.

TABLE 3 | Multiple stepwise regression analysis.

Final mode			Included variables	β	*t*	*p*	Excluded variables	β	*t*	*p*
R2	*F*	*p*								
A. Multiple regression analyses with monosyllabic words as dependent variable										
0.331	14.336	0.001	GDT	−1.71	−3.786	0.001	Temporal gap ACC threshold	−0.212	−1.253	0.221
B. Multiple regression analyses with disyllable words as dependent variable										
0.686	27.331	<0.001	GDT	−3.031	−6.235	<0.001	Frequency change ACC threshold	−0.084	−0.706	0.487
							Temporal gap ACC threshold	−0.132	−1.015	0.32
			100 ms gap ACC amplitude	8.741	3.699	0.001	50% F_change ACC amplitude	0.013	0.104	0.918
							100 ms gap ACC potential	0.126	0.911	0.371
C. Multiple regression analyses with sentences in quiet as dependent variable										
0.55	31.841	<0.001	GDT	−3.945	−5.643	<0.001	Frequency change ACC threshold	−0.181	−1.361	0.186
							Temporal gap ACC threshold	−0.172	−1.134	0.268
							50% F_change ACC amplitude	0.153	1.147	0.262
							100 ms gap ACC potential	−0.05	−0.322	0.75
D. Multiple regression analyses with STR in noise as dependent variable										
0.577	9.562	0.002	100 ms gap ACC amplitude	−2.337	−3.073	0.008	FCDT	−0.068	−0.203	0.842
			GDT	0.61	2.551	0.023	Frequency change ACC threshold	0.029	0.148	0.885

The Speech Perception of Post-lingual Cochlear Implant Users Was Affected by the Ability to Detect Temporal Gap

Multiple regression analysis showed that, as anticipated, speech perception could be partly predicted by the GDT and ACC amplitudes induced by the temporal gap. The auditory system uses temporal cues, such as the duration of speech segments and silent intervals between speech segments, to differentiate various speech sounds (Dorman et al., 1985). Current CIs mainly use an envelope-based speech-processing strategy to encode time-varying amplitudes in several frequency bands (Wilson et al., 1991; Shannon et al., 1995). Moreover, the spectral information provided by the CI is degraded and, therefore, significantly poorer than that heard by listeners with normal hearing (Xu et al., 2005; Sagi et al., 2009). The temporal resolution (i.e., the ability to follow rapid changes in the time waveform) is critical for speech recognition in CI users (Luo et al., 2020).

However, some studies (Mussoi and Brown, 2019; Cesur and Derinsu, 2020) have failed to confirm the correlation between GDT and speech perception. Several stimulus parameters, such as intensity, duration, temporal envelopes, and similarity of pre- and post-gate frequencies (Phillips et al., 1997; Schneider and Hamstra, 1999; Pichora-Fuller et al., 2006; Garadat and Pfingst, 2011; Horwitz et al., 2011), have also been shown to affect GDTs in CI users. The discrepancies in the literature regarding GDTs in CI users may be partially caused by the variety of stimulus parameters used in these studies (Blankenship et al., 2016). In addition to the stimulus parameters, the relationship between GDT and speech perception may also be modified by other factors affecting speech perception. Speech recognition requires cognitive processes such as attention, memory, and intelligence, as well as intact auditory pathways. The existence of various factors, such as peripheral, central, and cognitive processes, makes it difficult to evaluate the effects of temporal resolution on speech comprehension problems alone (Cesur and Derinsu, 2020). Therefore, in studies that fail to confirm the correlation between GDT and speech perception, the results may be influenced by the aforementioned stimulus parameters or factors that affect speech perception.

Clinical Implications

This study found that GDT could significantly predict speech perception in quiet or noisy environments of post-lingual CI users. This indicated that the GDT may provide an easy, quick, and non-linguistic tool to "screen out" poor CI ears for target intervention. Interventions may include doing auditory discrimination training, designing language rehabilitation courses for specific CI users. This tool is useful when patients cannot be reliably expected to perform well on clinical speech tests (e.g., young children) or when they have language barriers (e.g., non-native speakers).

This study also found that the ACC amplitude induced by the temporal gap can significantly predict the perception of disyllabic words in quiet and speech perception in noisy. This suggests, to some extent, that ACC amplitude induced by the temporal gap can be used as an objective tool to predict speech outcomes. This tool is useful when patients cannot be reliably expected to perform well on psychophysical tests or speech tests.

Limitations and Future Studies

Among the five conditions with frequency change or gap duration change, the minimum change condition that could induce an ACC response was defined as the frequency change ACC threshold or temporal gap ACC threshold in this study.

However, this was only an approximate estimate of the actual ACC threshold. Future research can use an adaptive program and real-time data analysis to calculate the exact threshold of ACC response, as has been done by Vonck et al. (2021).

In multiple regression analysis, the ACC amplitude induced by temporal gap could significantly predict disyllabic words in quiet and SRT in noise, but not monosyllabic words or sentences in quiet. We can't explain this result well, but it may be related to the differences among materials of speech tests. Compared with identifying monosyllabic words and sentences, identifying monosyllabic words may be a moderately difficult task. There is no hint of other syllables, so it is relatively difficult to recognize monosyllabic words. Meanwhile, it may be relatively simple to recognize the sentences in quiet because of the hints of the preceding and following words in the sentences. Therefore, the recognition of disyllabic words may better represent the recognition ability of speech sounds. But more research is needed to verify this argument.

Although GDT and ACC amplitude evoked by temporal gaps were associated with speech perception in this study, this could explain only approximately half of the variability in speech perception scores. Other factors that were not considered in the present study might help explain the remaining variability in speech perception. Cognitive abilities and listening efforts are likely to be important for speech perception. For example, according to Mussoi and Brown (2019), digit span and cognitive ability are correlated with speech perception performance.

CONCLUSION

In this study, no influence of CI hearing threshold on speech perception was found in post-lingual CI users. Psychophysical and neurophysiological methods were used to investigate the influence of the ability to detect frequency changes or temporal gaps on the speech perception of post-lingual CI users. When the ability to detect frequency changes and the temporal gap was considered simultaneously, the ability to detect frequency changes had no significant effect on speech perception, but the ability to detect temporal gaps could significantly predict speech perception. The GDT obtained using the psychophysical method is a good predictor of speech perception, and the ACC amplitude induced by the temporal gap can also predict speech perception to a certain extent.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Medical Ethics Committee of Shandong Provincial ENT Hospital. Written informed consent to participate in this study was provided by the participants.

AUTHOR CONTRIBUTIONS

LX, HW, and ZF contributed to the conception of the study. DX, JLu, and XC contributed to the experimental design. JLi and XL selected the data and performed the analysis. All authors contributed to manuscript revision and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (grant nos. 81800905 and 82071053) and the Shandong Provincial ENT Hospital (1237000049557032XQ).

ACKNOWLEDGMENTS

We thank all participants for their participation in this study.

REFERENCES

Anderson, E. S., Oxenham, A. J., Nelson, P. B., and Nelson, D. A. (2012). Assessing the role of spectral and intensity cues in spectral ripple detection and discrimination in cochlear-implant users. *J. Acoust. Soc. Am.* 132, 3925–3934. doi: 10.1121/1.4763999

Blankenship, C., Zhang, F., and Keith, R. (2016). Behavioral measures of temporal processing and speech perception in cochlear implant users. *J. Am. Acad. Audiol.* 27, 701–713. doi: 10.3766/jaaa.15026

Cazals, Y., Pelizzone, M., Saudan, O., and Boex, C. (1994). Low-pass filtering in amplitude modulation detection associated with vowel and consonant identification in subjects with cochlear implants. *J. Acoust. Soc. Am.* 96, 2048–2054. doi: 10.1121/1.410146

Cesur, S., and Derinsu, U. (2020). Temporal processing and speech perception performance in postlingual adult users of cochlear implants. *J. Am. Acad. Audiol.* 31, 129–136. doi: 10.3766/jaaa.19002

Dimitrijevic, A., Michalewski, H. J., Zeng, F. G., Pratt, H., and Starr, A. (2008). Frequency changes in a continuous tone: auditory cortical potentials. *Clin. Neurophysiol.* 119, 2111–2124. doi: 10.1016/j.clinph.2008.06.002

Dorman, M. F., Lindholm, J. M., and Hannley, M. T. (1985). Influence of the first formant on the recognition of voiced stop consonants by hearing-impaired listeners. *J. Speech Hear. Res.* 28, 377–380. doi: 10.1044/jshr.2803.377

Firszt, J. B., Holden, L. K., Skinner, M. W., Tobey, E. A., Peterson, A., Gaggl, W., et al. (2004). Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems. *Ear Hear.* 25, 375–387. doi: 10.1097/01.aud.0000134552.22205.ee

Florentine, M., and Buus, S. (1984). Temporal gap detection in sensorineural and simulated hearing impairments. *J. Speech Hear. Res.* 27, 449–455. doi: 10.1044/jshr.2703.449

Fu, Q. J. (2002). Temporal processing and speech recognition in cochlear implant users. *Neuroreport* 13, 1635–1639. doi: 10.1097/00001756-200209160-00013

Fu, Q. J., Chinchilla, S., and Galvin, J. J. (2004). The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users. *J. Assoc. Res. Otolaryngol.* 5, 253–260. doi: 10.1007/s10162-004-4046-1

Garadat, S. N., and Pfingst, B. E. (2011). Relationship between gap detection thresholds and loudness in cochlear-implant users. *Hear. Res.* 275, 130–138. doi: 10.1016/j.heares.2010.12.011

Gnansia, D., Lazard, D. S., Leger, A. C., Fugain, C., Lancelin, D., Meyer, B., et al. (2014). Role of slow temporal modulations in speech identification for cochlear implant users. *Int J Audiol.* 53, 48–54. doi: 10.3109/14992027.2013.844367

He, S., Grose, J. H., and Buchman, C. A. (2012). Auditory discrimination: the relationship between psychophysical and electrophysiological measures. *Int. J. Audiol.* 51, 771–782. doi: 10.3109/14992027.2012.699198

He, S., Grose, J. H., Teagle, H. F., Woodard, J., Park, L. R., Hatch, D. R., et al. (2015). Acoustically evoked auditory change complex in children with auditory neuropathy spectrum disorder: a potential objective tool for identifying cochlear implant candidates. *Ear Hear.* 36, 289–301. doi: 10.1097/AUD.0000000000000119

He, S., Grose, J. H., Teagle, H. F. B., and Buchman, C. A. (2014). Objective measures of electrode discrimination with electrically evoked auditory change complex and speech-perception abilities in children with auditory neuropathy spectrum disorder. *Ear Hear.* 35, e63–e74. doi: 10.1097/01.aud.0000436605.92129.1b

He, S., Grose, J. H., Teagle, H. F. B., Woodard, J., Park, L. R., Hatch, D. R., et al. (2013). Gap detection measured with electrically evoked auditory event-related potentials and speech-perception abilities in children with auditory neuropathy spectrum disorder. *Ear Hear.* 34, 733–744. doi: 10.1097/AUD.0b013e3182944bb5

He, S., McFayden, T. C., Shahsavarani, B. S., Teagle, H. F. B., Ewend, M., Henderson, L., et al. (2018). The electrically evoked auditory change complex evoked by temporal gaps using cochlear implants or auditory brainstem implants in children with cochlear nerve deficiency. *Ear Hear.* 39, 482–494. doi: 10.1097/aud.0000000000000498

Henry, B. A., Turner, C. W., and Behrens, A. (2005). Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. *J. Acoust. Soc. Am.* 118, 1111–1121. doi: 10.1121/1.1944567

Holden, L. K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., Potts, L. G., et al. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear Hear.* 34, 342–360. doi: 10.1097/AUD.0b013e3182741aa7

Hoover, E., Pasquesi, L., and Souza, P. (2015). Comparison of clinical and traditional gap detection tests. *J. Am. Acad. Audiol.* 26, 540–546. doi: 10.3766/jaaa.14088

Horwitz, A. R., Ahlstrom, J. B., and Dubno, J. R. (2011). Level-dependent changes in detection of temporal gaps in noise markers by adults with normal and impaired hearing. *J. Acoust. Soc. Am.* 130, 2928–2938. doi: 10.1121/1.3643829

Kim, J. R., Brown, C. J., Abbas, P. J., Etler, C. P., and O'Brien, S. (2009). The effect of changes in stimulus level on electrically evoked cortical auditory potentials. *Ear Hear.* 30, 320–329. doi: 10.1097/AUD.0b013e31819c42b7

Kong, Y.-Y., and Zeng, F.-G. (2006). Temporal and spectral cues in Mandarin tone recognition. *J. Acoust. Soc. Am.* 120, 2830–2840. doi: 10.1121/1.2346009

Lawler, M., Yu, J., and Aronoff, J. M. (2017). Comparison of the spectral-temporally modulated ripple test with the Arizona Biomedical Institute Sentence Test in cochlear implant users. *Ear Hear.* 38, 760–766. doi: 10.1097/AUD.0000000000000496

Liang, C., Houston, L. M., Samy, R. N., Abedelrehim, L. M. I., and Zhang, F. (2018). Cortical processing of frequency changes reflected by the acoustic change complex in adult cochlear implant users. *Audiol. Neurootol.* 23, 152–164. doi: 10.1159/000492170

Liang, C., Wenstrup, L. H., Samy, R. N., Xiang, J., and Zhang, F. (2020). The effect of side of implantation on the cortical processing of frequency changes in adult cochlear implant users. *Front. Neurosci.* 14:368. doi: 10.3389/fnins.2020.00368

Lister, J. J., Roberts, R. A., Krause, J. C., Debiase, D., and Carlson, H. (2011). An adaptive clinical test of temporal resolution: within-channel and across-channel gap detection. *Int. J. Audiol.* 50, 375–384. doi: 10.3109/14992027.2010.551217

Luo, X., Kolberg, C., Pulling, K. R., and Azuma, T. (2020). Psychoacoustic and demographic factors for speech recognition of older adult cochlear implant users. *J. Speech Lang. Hear. Res.* 63, 1712–1725. doi: 10.1044/2020_JSLHR-19-00225

Martin, B. A., and Boothroyd, A. (1999). Cortical, auditory, event-related potentials in response to periodic and aperiodic stimuli with the same spectral envelope. *Ear Hear.* 20, 33–44. doi: 10.1097/00003446-199902000-00004

Martin, B. A., Tremblay, K. L., and Korczak, P. (2008). Speech evoked potentials: from the laboratory to the clinic. *Ear Hear.* 29, 285–313. doi: 10.1097/AUD.0b013e3181662c0e

McGuire, K., Firestone, G. M., Zhang, N., and Zhang, F. (2021). The acoustic change complex in response to frequency changes and its correlation to cochlear implant speech outcomes. *Front. Hum. Neurosci.* 15:757254. doi: 10.3389/fnhum.2021.757254

Moore, B. C. (1996). Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. *Ear Hear.* 17, 133–161. doi: 10.1097/00003446-199604000-00007

Mussoi, B. S. S., and Brown, C. J. (2019). Age-related changes in temporal resolution revisited: electrophysiological and behavioral findings from cochlear implant users. *Ear Hear.* 40, 1328–1344. doi: 10.1097/AUD.0000000000000732

Phillips, D. P., Taylor, T. L., Hall, S. E., Carr, M. M., and Mossop, J. E. (1997). Detection of silent intervals between noises activating different perceptual channels: some properties of "central" auditory gap detection. *J. Acoust. Soc. Am.* 101, 3694–3705. doi: 10.1121/1.419376

Pichora-Fuller, M. K., Schneider, B. A., Benson, N. J., Hamstra, S. J., and Storzer, E. (2006). Effect of age on detection of gaps in speech and nonspeech markers varying in duration and spectral symmetry. *J. Acoust. Soc. Am.* 119, 1143–1155. doi: 10.1121/1.2149837

Pratt, H., Starr, A., Michalewski, H. J., Dimitrijevic, A., Bleich, N., and Mittelman, N. (2009). Auditory-evoked potentials to frequency increase and decrease of high- and low-frequency tones. *Clin. Neurophysiol.* 120, 360–373. doi: 10.1016/j.clinph.2008.10.158

Sagi, E., Kaiser, A. R., Meyer, T. A., and Svirsky, M. A. (2009). The effect of temporal gap identification on speech perception by users of cochlear implants. *J. Speech Lang. Hear. Res.* 52, 385–395. doi: 10.1044/1092-4388(2008/07-0219)

Schneider, B. A., and Hamstra, S. J. (1999). Gap detection thresholds as a function of tonal duration for younger and older listeners. *J. Acoust. Soc. Am.* 106, 371–380. doi: 10.1121/1.427062

Sek, A., and Moore, B. C. (1995). Frequency discrimination as a function of frequency, measured in several ways. *J. Acoust. Soc. Am.* 97, 2479–2486. doi: 10.1121/1.411968

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science.* 270, 303–304. doi: 10.1126/science.270.5234.303

Sharma, A., and Dorman, M. F. (2006). Central auditory development in children with cochlear implants: clinical implications. *Adv. Oto Rhino Laryngol.* 64, 66–88. doi: 10.1159/000094646

Small, S. A., and Werker, J. F. (2012). Does the ACC have potential as an index of early speech discrimination ability? A preliminary study in 4-month-old infants with normal hearing. *Ear Hear.* 33, e59–e69. doi: 10.1097/AUD.0b013e31825f29be

Tremblay, K., Kraus, N., McGee, T., Ponton, C., and Otis, B. (2001). Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear Hear.* 22, 79–90. doi: 10.1097/00003446-200104000-00001

Vonck, B. M. D., Lammers, M. J. W., Schaake, W. A. A., van Zanten, G. A., Stokroos, R. J., and Versnel, H. (2021). Cortical potentials evoked by tone frequency changes compared to frequency discrimination and speech perception: thresholds in normal-hearing and hearing-impaired subjects. *Hear. Res.* 401, 108154. doi: 10.1016/j.heares.2020.108154

Wei, C., Cao, K., Jin, X., Chen, X., and Zeng, F.-G. (2007). Psychophysical performance and Mandarin tone recognition in noise by cochlear implant users. *Ear Hear.* 28, 62S–65S. doi: 10.1097/AUD.0b013e318031512c

Wilson, B. S., and Dorman, M. F. (2008). Cochlear implants: a remarkable past and a brilliant future. *Hear Res.* 242, 3–21. doi: 10.1016/j.heares.2008.06.005

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature.* 352, 236–238. doi: 10.1038/352236a0

Won, J. H., Drennan, W. R., Nie, K., Jameyson, E. M., and Rubinstein, J. T. (2011). Acoustic temporal modulation detection and speech perception in cochlear implant listeners. *J Acoust Soc Am.* 130, 376–388. doi: 10.1121/1.3592521

Won, J. H., Drennan, W. R., and Rubinstein, J. T. (2007). Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *J. Assoc. Res. Otolaryngol.* 8, 384–392. doi: 10.1007/s10162-007-0085-8

Won, J. H., Moon, I. J., Jin, S., Park, H., Woo, J., Cho, Y. S., et al. (2015). Spectrotemporal Modulation Detection and Speech Perception by Cochlear Implant Users. *PLoS One.* 10:e0140920. doi: 10.1371/journal.pone.0140920

Xin, X., Gao-yang, H., Fei, J., Jia, J., Ai-ting, C., Yang, Z., et al. (2010). The establishment of computer-assisted Chinese speech audiometry platform. *Chinese Sci. J. Hear. Speech Rehabil.* 41, 31–34. doi: 10.3969/j.issn.1672-4933.2010.04.005

Xu, L., Thompson, C. S., and Pfingst, B. E. (2005). Relative contributions of spectral and temporal cues for phoneme recognition. *J. Acoust. Soc. Am.* 117, 3255–3267. doi: 10.1121/1.1886405

Zhang, F., Underwood, G., McGuire, K., Liang, C., Moore, D. R., and Fu, Q. J. (2019). Frequency change detection and speech perception in cochlear implant users. *Hear. Res.* 379, 12–20. doi: 10.1016/j.heares.2019.04.007

Zhou, N., Dixon, S., Zhu, Z., Dong, L., and Weiner, M. (2020). Spectrotemporal modulation sensitivity in cochlear-implant and normal-hearing listeners: is the performance driven by temporal or spectral modulation sensitivity? *Trends Hear.* 24, 1–11. doi: 10.1177/2331216520948385

Zhou, Q., Bi, J., Song, H., Gu, X., and Liu, B. (2020). Mandarin lexical tone recognition in bimodal cochlear implant users. *Int. J. Audiol.* 59, 548–555. doi: 10.1080/14992027.2020.1719437

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Neuroscience

BRIEF RESEARCH REPORT
published: 10 June 2022
doi: 10.3389/fnins.2022.888596

Check for updates

The P300 Auditory Event-Related Potential May Predict Segregation of Competing Speech by Bimodal Cochlear Implant Listeners

Duo-Duo Tao[1†], Yun-Mei Zhang[2†], Hui Liu[1], Wen Zhang[1], Min Xu[1], John J. Galvin III[3], Dan Zhang[2] and Ji-Sheng Liu[2*]*

[1] Department of Ear, Nose, and Throat, Shaanxi Provincial People's Hospital, Xi'An, China, [2] Department of Ear, Nose, and Throat, The First Affiliated Hospital of Soochow University, Suzhou, China, [3] House Institute Foundation, Los Angeles, CA, United States

OPEN ACCESS

Edited by:
Xin Luo,
Arizona State University, United States

Reviewed by:
Yingjiu Nie,
James Madison University,
United States
Zilong Xie,
University of Kansas Medical Center,
United States

***Correspondence:**
Duo-Duo Tao
entdtao@163.com
Ji-Sheng Liu
ljswwq@sina.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: 03 March 2022
Accepted: 16 May 2022
Published: 10 June 2022

Citation:
Tao D-D, Zhang Y-M, Liu H, Zhang W, Xu M, Galvin JJ III, Zhang D and Liu J-S (2022) The P300 Auditory Event-Related Potential May Predict Segregation of Competing Speech by Bimodal Cochlear Implant Listeners. Front. Neurosci. 16:888596. doi: 10.3389/fnins.2022.888596

Compared to normal-hearing (NH) listeners, cochlear implant (CI) listeners have greater difficulty segregating competing speech. Neurophysiological studies have largely investigated the neural foundations for CI listeners' speech recognition in quiet, mainly using the P300 component of event-related potentials (ERPs). P300 is closely related to cognitive processes involving auditory discrimination, selective attention, and working memory. Different from speech perception in quiet, little is known about the neurophysiological foundations for segregation of competing speech by CI listeners. In this study, ERPs were measured for a 1 vs. 2 kHz contrast in 11 Mandarin-speaking bimodal CI listeners and 11 NH listeners. Speech reception thresholds (SRTs) for a male target talker were measured in steady noise or with a male or female masker. Results showed that P300 amplitudes were significantly larger and latencies were significantly shorter for the NH than for the CI group. Similarly, SRTs were significantly better for the NH than for the CI group. Across all participants, P300 amplitude was significantly correlated with SRTs in steady noise ($r = -0.65$, $p = 0.001$) and with the competing male ($r = -0.62$, $p = 0.002$) and female maskers ($r = -0.60$, $p = 0.003$). Within the CI group, there was a significant correlation between P300 amplitude and SRTs with the male masker ($r = -0.78$, $p = 0.005$), which produced the most informational masking. The results suggest that P300 amplitude may be a clinically useful neural correlate of central auditory processing capabilities (e.g., susceptibility to informational masking) in bimodal CI patients.

Keywords: cochlear implant, competing speech, informational masking, event-related potentials, P300

INTRODUCTION

While cochlear implants (CIs) provide sufficient spectro-temporal resolution for speech recognition in quiet by deaf individuals, masked speech recognition is often difficult for CI users. Steady noise is thought to largely produce "energetic" masking; the spectro-temporal overlap between the target and masker occurs at the periphery (e.g., Brungart, 2001; Kidd et al., 2002). Competing speech is thought to produce some combination of energetic masking, "envelope" masking (target and masker envelope interference even when there is no spectral overlap; e.g., Stone and Canavan, 2016), and "informational" masking (e.g., lexical interference, target/masker similarities, etc.; Brungart, 2001; Kidd et al., 2002, 2016). Different from normal-hearing (NH) listeners, who have greater difficulty with competing noise than with competing speech, CI listeners have greater difficulty with competing speech than with competing noise (e.g., Stickney et al., 2004; Cullington and Zeng, 2008; Tao et al., 2018). The coarse spectro-temporal resolution is thought to limit CI users' segregation of target and masker speech (e.g., Friesen et al., 2001; Shannon et al., 2004; Fu and Nogaki, 2005; Luo and Fu, 2009).

Cortical measures have been used to characterize NH and CI listeners' auditory processing. Auditory event-related potentials (ERPs) reflect the brain's response to changes in an ongoing stimulus (e.g., deviant stimuli in the context of frequent stimuli in an oddball paradigm). Exogenous, pre-attentive responses (e.g., P1, N1, P2, N2 peaks) typically occur within the first 250 ms and do not reflect cognitive processing (e.g., Martin et al., 2008; Lightfoot, 2016). The latency of the endogenous P3 (or P300) response is typically between 250 and 400 ms, and is thought to reflect attention and/or arousal (e.g., Polich and Kok, 1995; Kok, 2001). Recording of P300 responses requires some sort of behavioral response to the deviant stimulus (e.g., counting the number of deviant stimuli during a test run, indicating when a deviant stimulus was heard, etc.). P300 latency has been shown to be related to the speed of information processing (e.g., Ritter et al., 1972; Kutas et al., 1977; Parasuraman and Beatty, 1980; Donchin and Coles, 1988). P300 amplitude has been shown to decrease with increasing task difficulty (e.g., Parasuraman and Beatty, 1980). Uncertainty in discrimination of sounds may be reflected in reduced P300 amplitude (e.g., Sutton et al., 1965; Hillyard et al., 1971; Squires et al., 1973; Picton, 2011).

There is great variability in CI outcomes that is largely unexplained but may be related to individual central auditory processing capacities (e.g., Dunn et al., 2005). In CI users, ERPs may be used to observe detection (exogenous components) and discrimination (endogenous components) of stimulus contrasts. Assuming there are no cognitive deficits, device-related factors (e.g., the number of implanted electrodes, frequency allocation) and patient-related factors (e.g., the electrode-neural interface, patterns of neural survival, etc.) may affect ERP responses. As such, it is important to select stimuli that are sufficiently contrastive when measuring ERPs. Some studies have used pure-tone contrasts (e.g., Groenen et al., 2001; Beynon et al., 2002; Sasaki et al., 2009; Obuchi et al., 2012; Calderaro et al., 2020; Van Yper et al., 2020; Wedekind et al., 2021) while others have used phonemic contrasts (e.g., Groenen et al., 2001; Beynon et al., 2002, 2005; Beynon and Snik, 2004; Henkin et al., 2009; Micco et al., 1995). ERPs have also been used to observe the evolution of auditory processing after cochlear implantation in longitudinal studies (e.g., Kubo et al., 2001).

P300 is closely related to cognitive processes involving auditory discrimination, selective attention, and working memory (e.g., Polich, 2007). Segregation of competing speech has been shown to involve cognitive processes (e.g., Francis, 2010). Some CI studies have compared P300 responses to standard clinical measures such as word recognition in quiet (e.g., Kileny et al., 1997; Groenen et al., 2001; Grasel et al., 2018; Abrahamse et al., 2021; Amaral et al., 2021). Others have compared P300 responses to phoneme recognition in quiet (e.g., Groenen et al., 2001; Beynon et al., 2002) or to speech recognition in steady noise (e.g., Iwaki et al., 2004). Kileny et al. (1997) found a significant correlation between P300 amplitude and sentence recognition in pediatric CI users. Groenen et al. (2001) found a significant correlation between P300 amplitude and word/phoneme recognition in quiet in adult CI users.

Bimodal listening [CI in one ear, hearing aid (HA) in the other ear] provides important low-frequency temporal fine-structure cues that benefit pitch-mediated perception (e.g., music, talker identity, prosody) and segregation of target speech and maskers (e.g., Gifford et al., 2007; Cullington and Zeng, 2008; Dorman et al., 2008; Yoon et al., 2012; Crew et al., 2015; Liu et al., 2019). Previous studies have shown more robust P300 responses with bimodal than with CI-only listening. Iwaki et al. (2004) found that sentence recognition in noise was significantly better and P300 latency was significantly shorter with bimodal than with CI-only listening. Sasaki et al. (2009) also reported shorter P300 latency and better word recognition in quiet with bimodal than with CI-only listening. However, the relationship between P300 responses and segregation of competing speech with bimodal listening remains unclear.

In this study, P300 responses to pure-tone stimuli were recorded in NH listeners and bimodal CI users; speech recognition was measured in the presence of steady noise or competing speech. Given that the present participants used bimodal listening in daily life, only bimodal listening was tested. Also, previous studies have shown more robust P300 responses with bimodal than with CI-only listening (e.g., Iwaki et al., 2004; Sasaki et al., 2009). Consistent with previous studies (e.g., Beynon et al., 2005; Obuchi et al., 2012; Grasel et al., 2018), we expected greater P300 amplitudes and shorter P300 latencies in NH than in CI listeners. Given the great variability in speech performance among CI users (e.g., Stickney et al., 2004; Cullington and Zeng, 2008) and given that P300 is sensitive to auditory task difficulty (Parasuraman and Beatty, 1980; Polich, 1987; Causse et al., 2016), we expected that P300 responses would be related to masked speech recognition, especially for the more difficult segregation of competing speech by CI users.

METHODS

Participants

Eleven Mandarin-speaking CI listeners (six females, five males) participated in the study; the mean age at testing was 21.5 ± 9.2 years. All were users of Med-El devices. All except for CI-4 were implanted with the Sonata ti10 device with the Standard electrode array (31.5 mm); CI-4 was implanted with Concerto device and the Flex 28 electrode array (28 mm). All used the Opus 2 processor, and all used the FS4 strategy. All were bimodal listeners, using a CI in one ear and a hearing aid in the other ear in every day listening. The mean duration of deafness prior to implantation was 12.8 ± 6.5 years. The mean CI experience was 2.0 ± 1.9 years.

TABLE 1 | Demographic information of CI participants.

Participant	Sex	Age at test (yrs)	Dur deaf (yrs)	CI exp (yrs)	CI ear	Etiology	PTA (dB HL)	Mean HA gain (dB)
CI-C1	M	6.5	3.4	3.3	R	Congenital	94.2	33.3
CI-C2	F	10.6	8.0	0.8	L	Congenital	81.7	40.8
CI-C3	F	14.9	7.9	0.5	R	Unknown	82.5	29.2
CI-A1	F	20.1	15.0	0.8	R	Progressive	91.7	38.3
CI-A2	F	20.3	20.3	3.5	R	Congenital	90.8	29.2
CI-A3	F	20.7	20.0	0.6	L	Congenital	85.8	25.0
CI-A4	M	23.8	14.8	1.2	L	Unknown	85.0	23.3
CI-A5	M	23.9	12.9	1.0	L	Unknown	69.2	22.5
CI-A6	M	24.1	21.0	6.7	R	Congenital	81.7	20.0
CI-A7	M	31.3	15.0	0.6	R	Progressive	86.7	41.7
CI-A8	F	40.2	2.5	2.5	R	Sudden	83.3	17.5

All participants were users of Med-El devices, and all were everyday bimodal listeners (CI in one ear, hearing aid in the other ear).
Unaided PTA thresholds were calculated across 0.25, 0.5, 1.0, 2.0, 4.0, and 8.0 kHz. The mean HA gain was calculated as the mean difference between unaided and aided PTA thresholds. Dur deaf, duration of deafness before cochlear implantation; CI exp, experience with the CI device; CI-C, CI children; CI-A, CI adult.

CI participants C1, C2, A2, A3, A6 were prelingually deaf, and C3, A1, A4, A5, A7, A8 were postlingually deaf. **Table 1** shows demographic information for the CI participants. Eleven Mandarin-speaking NH listeners (seven females, four males) also participated in the study; the mean age at testing was 22.1 ± 9.5 years. A *t*-test showed no significant difference in age at testing between the CI and NH groups [$t(20) = 0.2$ $p = 0.882$]. All participants were recruited from Department of Ear, Nose, and Throat, The First Affiliated Hospital of Soochow University. The Ethical Committee from The First Affiliated Hospital of Soochow University specifically approved this study (Approval number 2021122). All participants provided written informed consent before participating in the study; parental approval was obtained for pediatric CI and NH listeners.

Speech Perception

The Closed-set Mandarin Speech (CMS; Tao et al., 2017) test materials were used to test speech recognition with the different maskers. The CMS test materials consist of familiar words selected to represent the natural distribution of vowels, consonants, and lexical tones found in Mandarin Chinese. Ten keywords in each of five categories (Name, Verb, Number, Color, and Fruit) were produced by native Mandarin talkers.

Speech reception thresholds (SRTs), defined as the target-to-masker ratio (TMR) that produced 50% correct keyword recognition, were adaptively measured using a modified coordinate response matrix test (Brungart, 2001). Two target keywords (randomly selected from the Number and Color categories) were embedded in a five-word carrier sentence uttered by a male target talker [mean fundamental frequency (F0) across all words = 136 Hz]. The first word in the target sentence was always the Name "Xiaowang," followed by randomly selected words from the remaining categories. Thus, the target sentence could be (translated from Mandarin) "**Xiaowang** sold ***Three Red*** strawberries" or "**Xiaowang** chose ***Four Brown*** bananas," etc. (Name to cue target talker in bold; target keywords in bold italic).

Recognition of the target keywords was measured in the presence of steady state noise (SSN) or competing speech; maskers were co-located with the target (0° azimuth). The spectrum of the SSN was matched to the long-term average spectrum of the target talker, averaged across all words. For competing speech, the masker was a female talker (mean F0 across all words = 248 Hz) or a different male talker (mean F0 = 178 Hz). Masker sentences were randomly generated for each test trial; words were randomly selected from each category, excluding the words used in the target sentence. Thus, the masker sentence could be "Xiaozhang saw *Two Blue* kumquats," "Xiaodeng took *Eight Green* papayas," etc. (competing keywords in italic).

All stimuli were presented in the sound field at 65 dBA *via* a single loudspeaker; subjects were seated in a sound-attenuated booth, directly facing the loudspeaker at a 1-m distance. For CI participants, SRTs were measured using the clinical settings for their devices, which were not changed throughout the study. During each test trial, a sentence was presented at the desired TMR; the initial TMR was 10 dB. Participants were instructed to listen to the target sentence (produced by the male target talker and beginning with the name "Xiaowang") and then click on one of the 10 response choices for each of the Number and Color categories; no selections could be made from the remaining categories, which were grayed out. If the subject correctly identified both keywords, the TMR was reduced by 4 dB (initial step size); if the subject did not correctly identify both keywords, the TMR was increased by 4 dB. After two reversals, the step size was reduced to 2 dB. The SRT was calculated by averaging the last six reversals in TMR. If there were fewer than six reversals within 20 trials, the test run was discarded and another run was measured. Two test runs were completed for each condition and the SRT was averaged across runs. The masker conditions were randomized within and across participants.

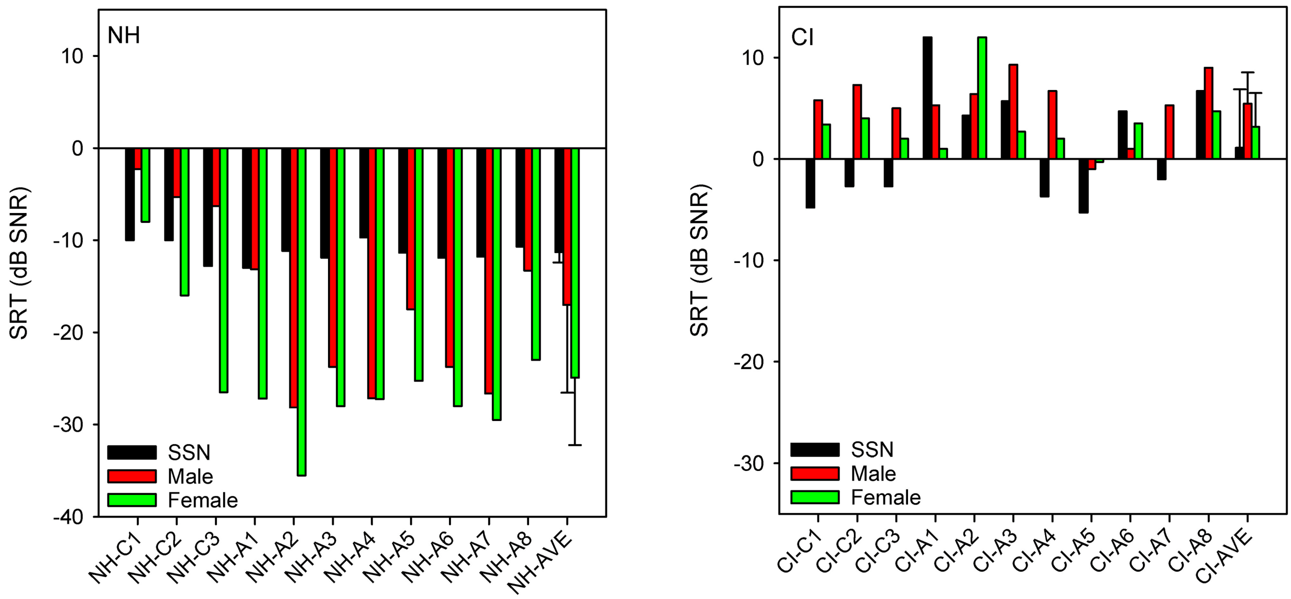

FIGURE 1 | (Left): SRTs with SSN, competing female, or competing masker for individual NH participants; mean SRTs across NH participants are shown at right. **(Right)**: Same as left panel, but for CI participants. In both panels, participants are ordered in terms of age at testing, with "C" indicating child listeners and "A" indicating adult listeners. The error bars show the standard deviation.

P300 Recordings

P300 ERPs were recorded using the Smart EP software (Intelligent Hearing System, Miami, FL, USA) and a multichannel recording paradigm. Disposable electrodes were placed at the high forehead (non-inverting electrode), both sides of the mastoid (inverting electrode), and low forehead (ground electrode). Absolute impedances and inter-electrode impedances were <5 and 3 kΩ, respectively. Responses were filtered online using a band-pass filter between 1 and 100 Hz. Pure-tone acoustic stimuli (1 or 2 kHz) with 50-ms duration and 5-ms rise and decay times were presented to the subjects every 1 s. Pure-tone stimuli were used instead of speech stimuli because pure-tone stimuli show better P300 reproducibility (e.g., Perez et al., 2017). The intensity of the stimuli was 20–30 dB above the aided PTA thresholds at 1 or 2 kHz to ensure that stimuli were clearly and comfortably audible for all participants.

Participants were seated in an electrically-shielded, sound-attenuated examination room. The stimuli were presented *via* two loudspeakers placed at ear level, 1 m away, ±45° relative to center. The probability was set at 80% for the frequent stimulus (1 kHz tone) and 20% for the rare stimulus (2 kHz tone). Participants were instructed to count the number of 2 kHz stimuli (oddball paradigm). All participants were able to discriminate between 1 and 2 kHz with 100% accuracy. In each run where all 20 oddball stimuli were identified, 20 ERPs for the rare stimuli were averaged. The recording window was comprised of a pre-stimulus baseline of 200 ms and a 500 ms post-stimulus epoch with a sampling rate of 1,000 Hz. Artifact rejection level was set at 100 mV. To avoid artifacts due to eye blinks, participants were instructed to close their eyes during the recording (Groenen et al., 2001). To reduce unwanted alpha rhythm, the inter-stimulus-interval was jittered by ±0.1 s (±10%), which made stimulus presentation less predictable and participants more attentive. Also, alpha rhythm was partially canceled out during the average processing because the onset of the P300 ERP is random relative to the phase of the alpha wave (Talsma and Woldorff, 2005).

P300 amplitude was calculated between the most positive point in the waveform between ≈250–400 ms and the following most negative point. This approach was chosen because the following most negative point was more distinct than the previous negative point. P300 latency was identified according to the P300 positive point. A minimum of three runs were tested, with more as needed if the participant did not identify all 20 oddball stimuli; only test runs where all 20 oddball stimuli were identified were included in analyses. Rest periods were taken between sessions to keep the participants alert.

RESULTS

Figure 1 shows SRTs with SSN or with a competing male or female talker for the NH and CI listeners. SRTs were much lower (better) for NH than for CI listeners. For NH listeners, mean SRTs progressively improved from SSN (−11.3 ± 1.1 dB) to the male masker (−17.0 ± 9.0 dB) and then to the female masker (−24.9 ± 7.3 dB). For CI listeners, mean SRTs were poorer with the competing male (5.5 ± 3.1 dB) or female talker (3.2 ± 3.3 dB) than with SSN (1.1 ± 5.9 dB). A mixed-design analysis of variance (ANOVA) was performed on the SRT data, with masker (SSN, male, female) as the within-subject factor and group (NH, CI) as the between-subject factor. Results showed significant effects of

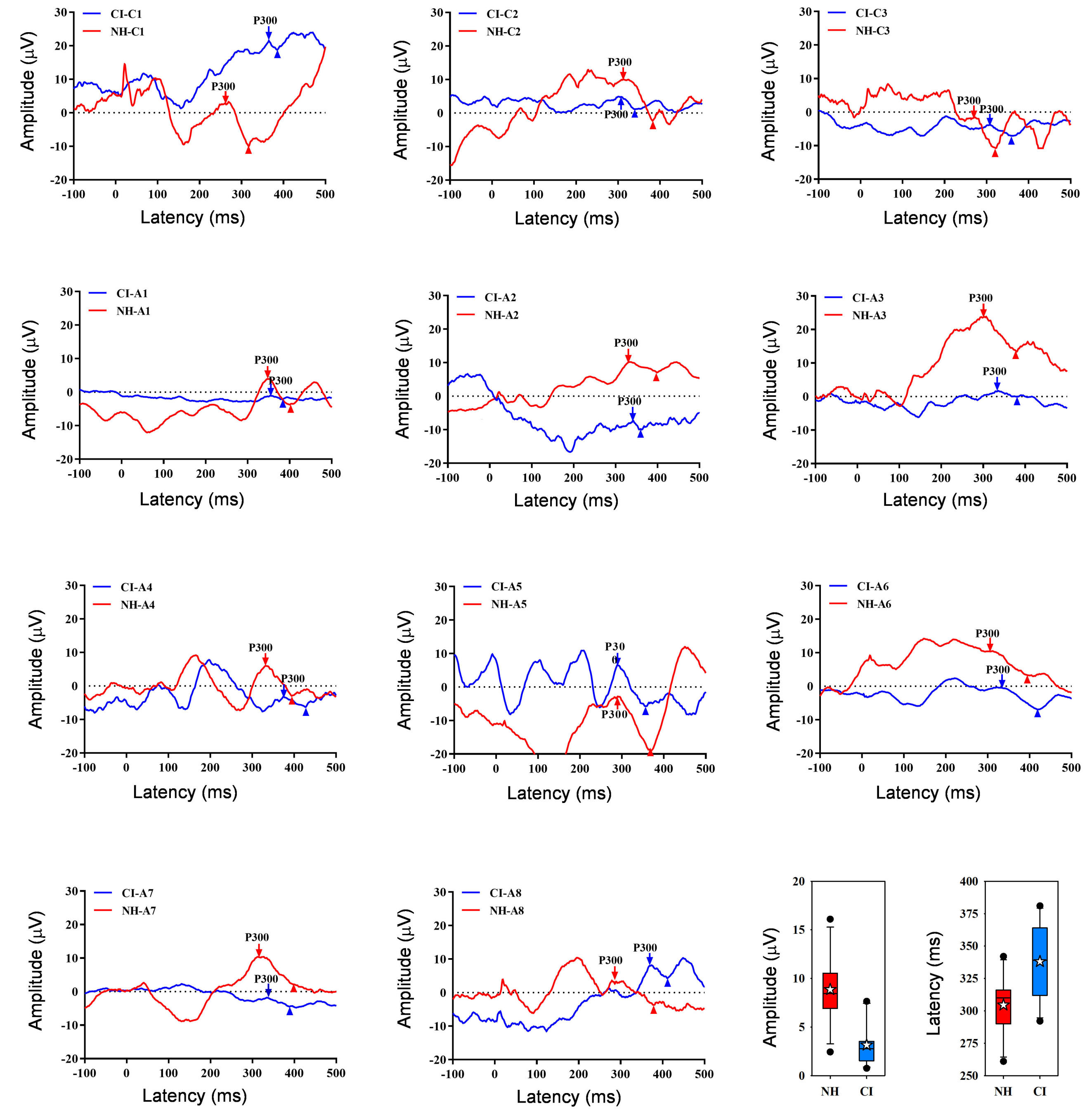

FIGURE 2 | Individual age-matched NH (red) and CI listener (blue) waveforms showing P300 responses averaged across the three test runs. The downward arrows show P300, and the upward triangles show the following negative point; P300 amplitude was calculated between P300 and the negative point. Panels are ordered in terms of age at testing; the top row shows data for child ("C") participants and the next two rows show data for adult ("A") participants. The panels at bottom right show boxplots of P300 amplitude and latency across all three runs for NH (red) and CI listeners (blue); the boxes show the 25th and 75th percentiles, the error bars show the 10th and 90th percentiles, the filled circles show outliers, the horizontal lines show the median, and the white stars show the mean.

group [$F_{(1,40)} = 125.1$, $p < 0.001$] and masker [$F_{(2,40)} = 10.7$, $p < 0.001$]; there was a significant interaction [$F_{(2,40)} = 16.8$, $p < 0.001$]. *Post-hoc* Bonferroni pairwise comparisons showed that for the NH group, SRTs were significantly higher (poorer) with SSN than with the male ($p = 0.016$) or female masker ($p < 0.001$), and significantly higher with the male than with the female masker ($p < 0.001$). There were no significant differences among the maskers for the CI group. SRTs were significantly lower (better) for the NH than for the CI group for all maskers ($p < 0.001$ for all comparisons).

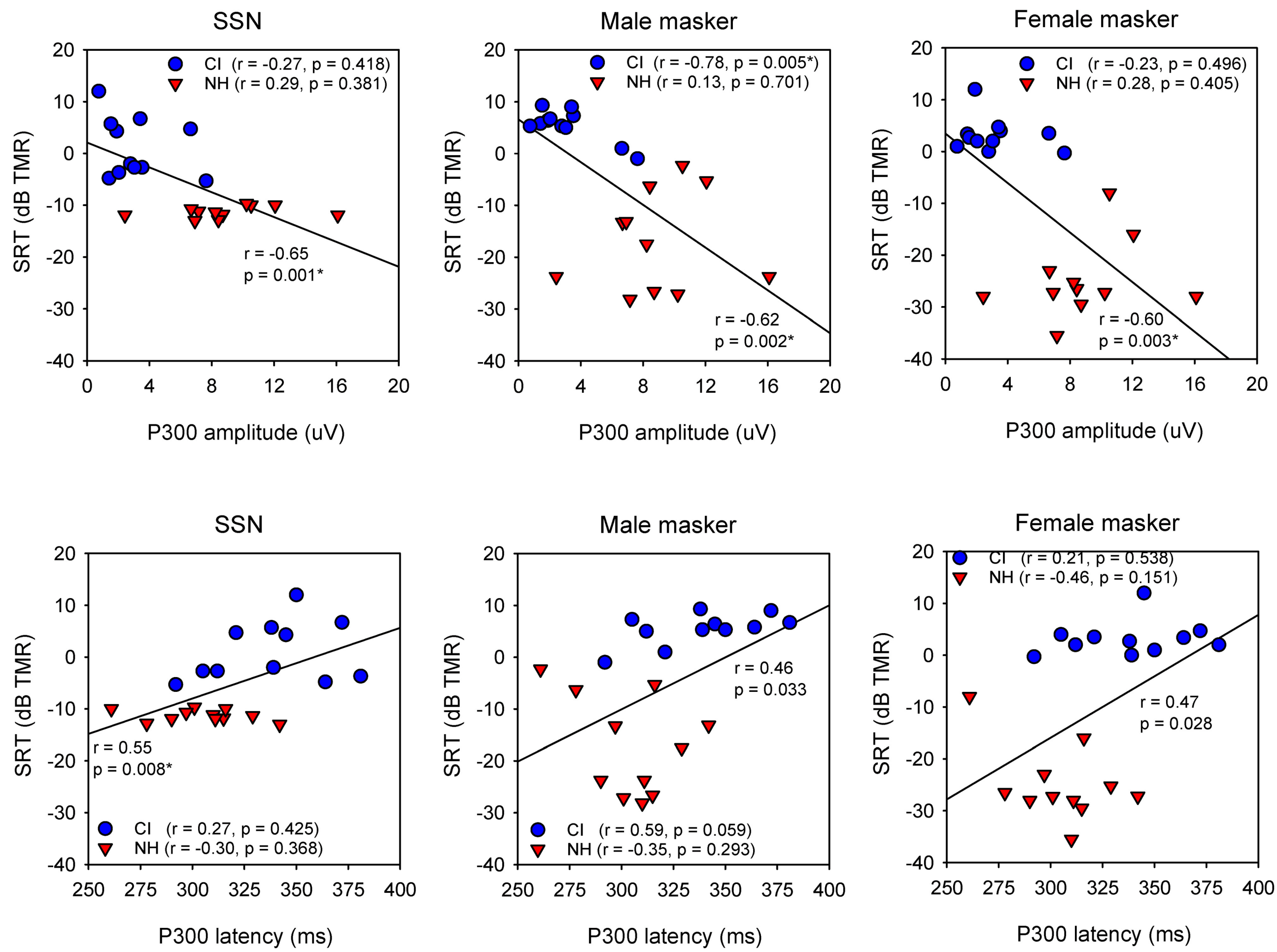

FIGURE 3 | (Top) Scatter plots of SRTs with SSN (left) or with a competing male (middle) or female talker (right) as a function of P300 amplitude, for the NH (red triangles) and CI listeners (blue circles). The diagonal line shows the linear regression across all data; the correlation coefficient and *p* value are shown near the line. Correlation coefficients and *p* values are shown for the CI data and NH data in the legend. Significant relationships after Bonferroni correction for multiple comparisons are indicated by asterisks. **(Bottom)** Same as top, but for SRTs as a function of P300 latency.

Figure 2 shows waveforms with the peak P300 response averaged across the three test runs for individual NH and CI listeners. Note that the intra-class correlation coefficient was 0.99 and 0.97 for P300 amplitude and latency, respectively, suggesting good test-retest reliability across the three runs. Because RM ANOVAs showed no significant effect of test run for NH or CI participants ($p > 0.05$ for all analyses), data were averaged across runs. Mean P300 amplitude was higher for the NH group (8.9 ± 3.5 μV) than for the CI group (3.2 ± 2.2 μV); mean P300 latency was shorter for the NH group (305 ± 23 ms) than for the CI group (338 ± 28 ms). *T*-tests showed that P300 amplitude was significantly higher for the NH than for the CI group [$t(20) = 4.6$, $p < 0.001$], and that P300 latency was significantly shorter for the NH than for the CI group [$t(20) = -3.1$, $p = 0.006$].

Figure 3 shows SRTs with SSN or with a competing male or female talker for the NH and CI groups as a function of P300 amplitude and latency; each data point shows the mean across three test runs. When all NH and CI data were combined, and after Bonferroni correction for multiple comparisons (adjusted $p = 0.016$), Pearson correlation analysis showed significant relationships between P300 amplitude and SRTs with SSN ($r = -0.65$, $p = 0.001$), and with the male ($r = -0.62$, $p = 0.002$) and female maskers ($r = -0.60$, $p = 0.003$). A significant relationship was observed between P300 latency and SRTs with SSN ($r = 0.60$, $p = 0.008$), but not for SRTs with the male or female masker. For the CI group, Pearson correlation analysis showed a significant relationship only between P300 amplitude and SRTs with the male masker ($r = -0.78$, $p = 0.005$); the correlation remained significant after controlling for age at testing, duration of deafness, and CI experience ($r = -0.81$, $p = 0.016$). No significant correlations were observed between P300 amplitude and SRTs with SSN or with the female masker, or between P300 latency and SRTs with any of the maskers. For the NH group, no significant relationships were observed between P300 amplitude or latency and SRTs with any of the maskers. For the CI group, a significant correlation was observed between P300 amplitude and unaided PTA thresholds (across all frequencies; $r = -0.87$, $p < 0.001$); there were no significant correlations between P300 amplitude and aided PTA thresholds. Significant correlations were observed between P300 latency and unaided PTA thresholds ($r = 0.67$, $p = 0.025$) and aided PTA thresholds ($r = 0.68$, $p = 0.021$). Note that statistical power was >0.80 for all of the above correlations, except for P300 latency vs. unaided PTA thresholds (power = 0.63) or aided PTA thresholds (power = 0.65).

DISCUSSION

Consistent with previous studies (e.g., Kubo et al., 2001; Beynon et al., 2005; Obuchi et al., 2012; Soshi et al., 2014; Grasel et al., 2018; Han et al., 2020), P300 amplitudes were significantly larger and latencies were significantly shorter for the NH group than for the CI group. For the present bimodal CI listeners, mean P300 amplitude and/or latency values were comparable to those observed in previous studies with CI listeners (e.g., Iwaki et al., 2004; Sasaki et al., 2009; Grasel et al., 2018; Abrahamse et al., 2021; Calderaro et al., 2020; Van Yper et al., 2020). P300 responses were elicited in all CI participants, consistent with Obuchi et al. (2012).

Mean SRTs for all maskers were lower (better) for the NH group than for the CI group, and values were comparable to those in previous studies using similar methods and stimuli (Tao et al., 2018; Zhang et al., 2020). Different from previous CI studies that showed lower SRTs in SSN than in competing speech (e.g., Cullington and Zeng, 2008; Croghan and Smith, 2018; Tao et al., 2018; Liu et al., 2019), there was no significant difference in SRTs between the SSN and competing speech maskers within the CI group. Note that CI listeners were tested while wearing contralateral hearing aids, which likely aided in segregation of competing speech, thereby reducing the deficit relative to SSN.

Across all NH and CI listeners, significant correlations were observed between P300 amplitude and SRTs with the SSN, male, and female maskers; a significant correlation was also observed between P300 latency and SRTs with SSN. These correlations were largely driven by across-group differences in speech performance and P300 responses. In general, higher P300 amplitude and shorter P300 latency were associated with better masked speech recognition.

In the NH group, there were no significant correlations between P300 responses and SRTs with any of the maskers. In the CI group, a significant correlation was observed only between P300 amplitude and SRTs with the male masker, the most challenging listening condition with the greatest informational masking. The correlation between P300 amplitude and SRTs with the male masker suggests some common relation to informational masking, a central auditory process. With the female masker, informational masking was reduced, and SSN produced largely energetic masking. Given the correlations between unaided PTA thresholds and P300 amplitude and latency and between aided PTA thresholds and P300 latency, differences in P300 response across CI listeners may have represented differences in segregation of the competing male talkers with residual acoustic hearing that provided low-frequency pitch cues.

Different from Soshi et al. (2014), we observed a significant correlation between P300 amplitude and SRTs with the male masker, but not between P300 amplitude and SRTs with SSN. Differences in cortical measure stimuli (1 vs. 2 kHz contrasts; consonant contrast), speech tests, methods, and CI patients (bimodal vs. CI-only listening) may have contributed to differences in results across studies. The 1 and 2 kHz stimuli used for ERP recording were presented at 20–30 dB above the aided thresholds, meaning that the aided acoustic hearing should have contributed to the response.

In the present study, ERPs and speech performance were measured only with bimodal listening. Some studies have shown greater P300 response and speech performance with bimodal than with CI-only listening (e.g., Iwaki et al., 2004; Sasaki et al., 2009). Interestingly, Wedekind et al. (2021) found no significant difference in P300 response between the NH ear and the CI ear in unilaterally deaf CI recipients; speech recognition in noise was better with the CI on than off. While it was not directly measured in Wedekind et al. (2021), speech performance would be expected to be much poorer with the CI ear alone than with the NH ear alone (e.g., Galvin et al., 2019). It is unclear why the P300 response would be similar across ears when speech performance would be different. As shown in **Figure 3**, significant relationships were observed between P300 amplitude and masked SRTs, presumably due to the underlying spectro-temporal resolution that was much better for NH than for CI listeners. However, some caution is warranted regarding the correlational analyses, given the limited number of participants and test runs. ERPs and speech performance were not measured with the acoustic-hearing ear alone or the CI ear alone in this study. It is possible that strong P300 responses may have been elicited within the acoustic-hearing ear alone, despite the expectedly poor speech performance. In future studies, it would be worthwhile to collect ERPs and speech performance with each ear alone and both ears together to better understand how the peripheral representations might affect the relationship between ERPs and speech performance.

The present results show some evidence that ERPs may be a useful objective measure to predict complex perception such as segregation of competing speech. However, eliciting P300 also requires a behavioral component in the oddball presentation, and the magnitude of the response may depend on the strength of the stimulus contrast. Obuchi et al. (2012) showed increasing P300 amplitude in CI listeners as the stimulus frequency contrast was increased from 1.5 to 4 kHz. Depending on the acoustic-to-electric frequency allocation and the electrode-neural interface (electrode position relative to healthy neurons), small contrasts (e.g., 1 vs. 1.5 kHz) may be perceived differently among CI listeners. The 1 vs. 2 kHz contrast in this study appeared to be sufficiently large to be discriminated by the present MED-EL CI users, most likely resulting in stimulation of electrodes 6 and 8, given the default frequency allocation. Note that there may have been some contribution from residual acoustic hearing for discrimination of the stimuli contrast.

CONCLUSIONS

Auditory ERPs and speech recognition in steady noise or competing speech were measured in NH and bimodal CI listeners. P300 amplitude was larger and latency was shorter in the NH group than in the CI group. Similarly, speech performance was better for the NH group than for the CI group. Significant correlations were observed across all participants between P300 amplitude and SRTs with steady noise and the male and female maskers. Within the CI group, P300 amplitude was significantly correlated with SRTs with the male masker, suggesting some relation between cortical response and informational masking.

DATA AVAILABILITY STATEMENT

The raw data from the study are included in the supplementary material; further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee from the First Affiliated Hospital of Soochow University (Approval number 2021122). All participants provided written informed consent before participating in the study; approval was obtained for pediatric CI and NH participants from their parents or adult next of kin.

AUTHOR CONTRIBUTIONS

D-DT: study design, data analysis, and writing of manuscript. Y-MZ and DZ: data analysis and writing of manuscript. JG: data analysis, data visualization, and writing of manuscript. HL, WZ, and MX: study supervision. J-SL: study design. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (81970877), the Jiangsu Provincial Natural Science Foundation—Outstanding Youth Foundation (BK20200054), the Jiangsu Provincial Key Research and Development Program Special Funds (BE2019670), and the Suzhou Science and Technology Project (SYS2019049).

ACKNOWLEDGMENTS

We thank the participants and their families for their time and contribution to this study.

REFERENCES

Abrahamse, R., Beynon, A., and Piai, V. (2021). Long-term auditory processing outcomes in early implanted young adults with cochlear implants: the mismatch negativity vs. P300 response. *Clin. Neurophysiol.* 132, 258–268. doi: 10.1016/j.clinph.2020.09.022

Amaral, M. S. A. D., Calderaro, V. G., Pauna, H. F., Massuda, E. T., Reis, A. C. M. B., and Hyppolito, M. Â. (2021). Is there a change in P300 evoked potential after 6 months in cochlear implant users? *Braz. J. Otorhinolaryngol.* 10, S1808-8694(21)00182-8. doi: 10.1016/j.bjorl.2021.10.002

Beynon, A. J., and Snik, A. F. (2004). Use of the event-related P300 potential in cochlear implant subjects for the study of strategy-dependent speech processing. *Int. J. Audiol.* 43(Suppl. 1), S44–S47.

Beynon, A. J., Snik, A. F., Stegeman, D. F., and van den Broek, P. (2005). Discrimination of speech sound contrasts determined with behavioral tests and event-related potentials in cochlear implant recipients. *J. Am. Acad. Audiol.* 16, 42–53. doi: 10.3766/jaaa.16.1.5

Beynon, A. J., and Snik, A. F., and van den Broek, P. (2002). Evaluation of cochlear implant benefit with auditory cortical evoked potentials. *Int. J. Audiol.* 41, 429–435. doi: 10.3109/14992020209090420

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *J. Acoust. Soc. Am.* 109, 1101–1109. doi: 10.1121/1.1345696

Calderaro, V. G., Amaral, M. S. A. D., Luz, B. A. B. D., Bernal, S. C., Hyppolito, M. Â., and Reis, A. C. M. B. (2020). Behavioral and electrophysiological assessment of adults who underwent cochlear implantation after hearing aid experience. *Int. Arch. Otorhinolaryngol.* 24, e132–e139. doi: 10.1055/s-0039-1695022

Causse, M., Peysakhovich, V., and Fabre, E. F. (2016). High working memory load impairs language processing during a simulated piloting task: an ERP and pupillometry study. *Front. Hum. Neurosci.* 10:240. doi: 10.3389/fnhum.2016.00240

Crew, J. D., Galvin, J. J. 3rd, Landsberger, D. M., and Fu, Q. J. (2015). Contributions of electric and acoustic hearing to bimodal speech and music perception. *PLoS ONE*. 10, e0120279. doi: 10.1371/journal.pone.0120279

Croghan, N. B. H., and Smith, Z. M. (2018). Speech understanding with various maskers in cochlear-implant and simulated cochlear-implant hearing: effects of spectral resolution and implications for masking release. *Trends Hear.* 22, 2331216518787276. doi: 10.1177/2331216518787276

Cullington, H. E., and Zeng, F. G. (2008). Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects. *J. Acoust. Soc. Am.* 123, 450–461. doi: 10.1121/1.2805617

Donchin, E., and Coles, M. G. H. (1988). On the conceptual foundations of cognitive psychology. *Behav. Brain Sci.* 1, 406–417. doi: 10.1017/S0140525X00058246

Dorman, M. F., Gifford, R. H., Spahr, A. J., and McKarns, S. A. (2008). The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. *Audiol. Neurootol.* 13, 105–112. doi: 10.1159/000111782

Dunn, C. C., R. S., and Tyler, and, S. A., Witt (2005). Benefit of wearing a hearing aid on the unimplanted ear in adult users of a cochlear implant. *J. Speech Lang. Hear. Res.* 48, 668–680. doi: 10.1044/1092-4388(2005/046)

Francis, A. L. (2010). Improved segregation of simultaneous talkers differentially affects perceptual and cognitive capacity demands for recognizing speech in competing speech. *Atten. Percept. Psychophys.* 72, 501–516. doi: 10.3758/APP.72.2.501

Friesen, L. M., Shannon, R. V., Baskent, D., and Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J. Acoust. Soc. Am.* 110, 1150–1163. doi: 10.1121/1.1381538

Fu, Q. J., and Nogaki, G. (2005). Noise susceptibility of cochlear implant users: the role of spectral resolution and smearing. *J. Assoc. Res. Otolaryngol.* 6, 19–27. doi: 10.1007/s10162-004-5024-3

Galvin, J. J., 3rd, Fu, Q. J., Wilkinson, E. P., Mills, D., Hagan, S. C., Lupo, J. E., et al. (2019). Benefits of cochlear implantation for single-sided deafness: data from the House Clinic-University of Southern California-University of California, Los Angeles clinical trial. *Ear Hear.* 40, 766–781. doi: 10.1097/AUD.0000000000000671

Gifford, R. H., Dorman, M. F., McKarns, S. A., and Spahr, A. J. (2007). Combined electric and contralateral acoustic hearing: word and sentence recognition with bimodal hearing. *J. Speech Lang. Hear. Res.* 50, 835–843. doi: 10.1044/1092-4388(2007/058)

Grasel, S., Greters, M., Goffi-Gomez, M. V. S., Bittar, R., Weber, R., Oiticica, J., et al. (2018). P3 cognitive potential in cochlear implant users. *Int. Arch. Otorhinolaryngol.* 22, 408–414. doi: 10.1055/s-0037-1613687

Groenen, P. A., Beynon, A. J., Snik, A. F., and van den Broek, P. (2001). Speech-evoked cortical potentials and speech recognition in cochlear implant users. *Scan Audiol.* 30, 31–40. doi: 10.1080/010503901750069554

Han, J. H., Lee, J., and Lee, H., J. (2020). Noise-induced change of cortical temporal processing in cochlear implant users. *Clin. Exp. Otorhinolaryngol.* 13, 241–248. doi: 10.21053/ceo.2019.01081

Henkin, Y., Tetin-Schneider, S., Hildesheimer, M., and Kishon-Rabin, L. (2009). Cortical neural activity underlying speech perception in postlingual adult cochlear implant recipients. *Audiol. Neurootol.* 14, 39–53. doi: 10.1159/000153434

Hillyard, S. A., Squires, K. C., Bauer, J. W., and Lindsay, P. H. (1971). Evoked potential correlates of auditory signal detection. *Science.* 172, 1357–1360. doi: 10.1126/science.172.3990.1357

Iwaki, T., Matsushiro, N., Mah, S. R., Sato, T., Yasuoka, E., Yamamoto, K., et al. (2004). Comparison of speech perception between monaural and binaural hearing in cochlear implant patients. *Acta Otolaryngol.* 124, 358–362. doi: 10.1080/00016480310000548a

Kidd, G. Jr, Mason, C. R., Swaminathan, J., Roverud, E., Clayton, K. K., and Best, V. (2016). Determining the energetic and informational components of speech-on-speech masking. *J. Acoust. Soc. Am.* 140, 132. doi: 10.1121/1.4954748

Kidd, G. Jr., Arbogast, T. L., Mason, C. R., and Walsh, M. (2002). Informational masking in listeners with sensorineural hearing loss. *J. Assoc. Res. Otolaryngol.* 3, 107–119. doi: 10.1007/s101620010095

Kileny, P. R., Boerst, A., and Zwolan, T. (1997). Cognitive evoked potentials to speech and tonal stimuli in children with implants. *Otolaryngol. Head Neck Surg.* 117, 161–169. doi: 10.1016/s0194-5998(97)70169-4

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiol.* 38, 557–577. doi: 10.1017/s0048577201990559

Kubo, T., Yamamoto, K., Iwaki, T., Matsukawa, M., Doi, K., and Tamura, M. (2001). Significance of auditory evoked responses (EABR and P300) in cochlear implant subjects. *Acta Otolaryngol.* 121, 257–261. doi: 10.1080/000164801300043749

Kutas, M., McCarthy, G., and Donchin, E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science.* 197, 792–795. doi: 10.1126/science.887923

Lightfoot, G. (2016). Summary of the N1-P2 cortical auditory evoked potential to estimate the auditory threshold in adults. *Semin. Hear.* 37, 1–8. doi: 10.1055/s-0035-1570334

Liu, Y. W., Tao, D. D., Chen, B., Cheng, X., Shu, Y., Galvin, J. J., et al. (2019). Factors affecting bimodal benefit in pediatric Mandarin-speaking Chinese cochlear implant users. *Ear Hear.* 40, 1316–1327. doi: 10.1097/AUD.0000000000000712

Luo, X., and Fu, Q. J. (2009). Concurrent-vowel and tone recognitions in acoustic and simulated electric hearing. *J. Acoust. Soc. Am.* 125, 3223–3233. doi: 10.1121/1.3106534

Martin, B. A., K. L., and Tremblay, and, P., Korczak (2008). Speech evoked potentials: from the laboratory to the clinic. *Ear Hear.* 29, 285–313. doi: 10.1097/AUD.0b013e3181662c0e

Micco, A. G., Kraus, N., Koch, D. B., McGee, T. J., Carrell, T. D., Sharma, A., et al. (1995). Speech-evoked cognitive P300 potentials in cochlear implant recipients. *Am. J. Otol.* 16, 514–520.

Obuchi, C., Harashima, T., and Shiroma, M. (2012). Auditory evoked potentials under active and passive hearing conditions in adult cochlear implant users. *Clin. Exp. Otorhinolaryngol.* 5(Suppl. 1), S6–S9. doi: 10.3342/ceo.2012.5.S1.S6

Parasuraman, R., and Beatty, J. (1980). Brain events underlying detection and recognition of weak sensory signals. *Science.* 210, 80–83. doi: 10.1126/science.7414324

Perez, A. P., Ziliotto, K., and Pereira, L. D. (2017). Test-retest of long latency auditory evoked potentials (P300) with pure tone and speech stimuli. *Int. Arch. Otorhinolaryngol.* 21, 134–139. doi: 10.1055/s-0036-1583527

Picton, T. W. (2011). "Human auditory evoked potentials," in *Endogenous Auditory Evoked Potentials: Attention Must be Paid*, ed. T. W. Picton (San Diego, CA: Plural Publishing), 399–448.

Polich, J. (1987). Task difficulty, probability, and inter-stimulus interval as determinants of P300 from auditory stimuli. *Electroencephalogr. Clin. Neurophysiol.* 68, 311–320. doi: 10.1016/0168-5597(87)90052-9

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Polich, J., and Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biol. Psych.* 41, 103–146. doi: 10.1016/0301-0511(95)05130-9

Ritter, W., Simson, R., and Vaughan H. G. Jr. (1972). Association cortex potentials and reaction time in auditory discrimination. *Electroenceph. Clin. Neurophysiol.* 33, 547–555. doi: 10.1016/0013-4694(72)90245-3

Sasaki, T., Yamamoto, K., Iwaki, T., and Kubo, T. (2009). Assessing binaural/bimodal advantages using auditory event-related potentials in subjects with cochlear implants. *Auris Nasus Larynx.* 36, 541–546. doi: 10.1016/j.anl.2008.12.001

Shannon, R. V., Fu, Q. J., and Galvin, J. 3rd (2004). The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Otolaryngol. Suppl.* 552, 50–54. doi: 10.1080/03655230410017562

Soshi, T., Hisanaga, S., Kodama, N., Kanekama, Y., Samejima, Y., Yumoto, E., et al. (2014). Event-related potentials for better speech perception in noise by cochlear implant users. *Hear Res.* 316, 110–121. doi: 10.1016/j.heares.2014.08.001

Squires, K. C., Hillyard, S. A., and Lindsay, P. H. (1973). Vertex potentials evoked during auditory signal detection: relation to decision criteria. *Percept. Psychophys.* 14, 263–272. doi: 10.3758/BF03212388

Stickney, G. S., Zeng, F. G., Litovsky, R., and Assmann, P. (2004). Cochlear implant speech recognition with speech maskers. *J. Acoust. Soc. Am.* 116, 1081–1091. doi: 10.1121/1.1772399

Stone, M. A., and Canavan, S. (2016). The near non-existence of "pure" energetic masking release for speech: extension to spectro-temporal modulation and glimpsing. *J. Acoust. Soc. Am.* 140, 832–842. doi: 10.1121/1.4960483

Sutton, S., Braren, M., Zubin, J., and John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science.* 150, 1187–1188. doi: 10.1126/science.150.3700.1187

Talsma, D., and Woldorff, M. G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *J. Cogn. Neurosci.* 17, 1098–1114. doi: 10.1162/0898929054475172

Tao, D. D., Fu, Q. J., Galvin, J. J. 3rd, and Yu, Y. F. (2017). The development and validation of the Closed-set Mandarin Sentence (CMS) test. *Speech Comm.* 92, 125–131. doi: 10.1016/j.specom.2017.06.008

Tao, D. D., Liu, Y. W., Fei, Y., Galvin, J. J. 3rd, Chen, B., and Fu, Q. J. (2018). Effects of age and duration of deafness on Mandarin speech understanding in competing speech by normal-hearing and cochlear implant children. *J. Acoust. Soc. Am.* 144, EL131. doi: 10.1121/1.5051051

Van Yper, L. N., Dhooge, I. J. M., Vermeire, K., De Vel, E. F. J., and Beynon, A. J. (2020). The P300 auditory event-related potential as a method to assess the benefit of contralateral hearing aid use in bimodal listeners: a proof-of-concept. *Int. J. Audiol.* 59, 73–80. doi: 10.1080/14992027.2019.1656346

Wedekind, A., Távora-Vieira, D., Nguyen, A. T., Marinovic, W., and Rajan, G. P. (2021). Cochlear implants in single-sided deaf recipients: near normal higher-order processing. *Clin. Neurophysiol.* 132, 449–456. doi: 10.1016/j.clinph.2020.11.038

Yoon, Y. S., Li, Y., and Fud Q. J. (2012). Speech recognition and acoustic features in combined electric and acoustic stimulation. *J. Speech Lang. Hear. Res.* 55, 105–124. doi: 10.1044/1092-4388(2011/10-0325)

Zhang, J., Wang, X., Wang, N. Y., Fu, X., Gan, T., Galvin, J. J. 3rd, et al. (2020). Tonal language speakers are better able to segregate competing speech according to talker sex differences. *J. Speech Lang. Hear Res.* 63, 2801–2810. doi: 10.1044/2020_JSLHR-19-00421

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Neuroscience

ORIGINAL RESEARCH
published: 14 June 2022
doi: 10.3389/fnins.2022.905244

Value of Preoperative Imaging Results in Predicting Cochlear Nerve Function in Children Diagnosed With Cochlear Nerve Aplasia Based on Imaging Results

*Xiuhua Chao, Ruijie Wang, Jianfen Luo, Haibo Wang, Zhaomin Fan and Lei Xu**

Department of Otolaryngology-Head and Neck Surgery, Shandong Provincial ENT Hospital, Shandong University, Jinan, China

OPEN ACCESS

Edited by:
Fei Chen,
Southern University of Science and Technology, China

Reviewed by:
Anandhan E. Dhanasingh,
MED-EL, Austria
Katherine Reinshagen,
Harvard Medical School, United States
Omid Majdani,
Wolfsburg Hospital, Germany

***Correspondence:**
Lei Xu
sdphxl@126.com

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *26 March 2022*
Accepted: *29 April 2022*
Published: *14 June 2022*

Citation:
Chao X, Wang R, Luo J, Wang H, Fan Z and Xu L (2022) Value of Preoperative Imaging Results in Predicting Cochlear Nerve Function in Children Diagnosed With Cochlear Nerve Aplasia Based on Imaging Results. Front. Neurosci. 16:905244. doi: 10.3389/fnins.2022.905244

This study aimed to assess the function of the cochlear nerve using electrically evoked compound action potentials (ECAPs) for children with cochlear implants who were diagnosed with cochlear nerve aplasia and to analyze the correlation between preimplantation imaging results and ECAP responses. Thirty-five children diagnosed with cochlear nerve aplasia based on magnetic resonance imaging (MRI) were included. Preimplantation MRI and high-resolution computed tomography (HRCT) images were reconstructed, and the width of the bone cochlear nerve canal (BCNC), the diameter of the vestibulocochlear nerve (VCN), and the diameter of the facial nerve (FN) were measured. ECAP input/output (I/O) functions were measured at three electrode locations along the electrode array for each participant. The relationship between ECAP responses (including ECAP threshold, ECAP maximum amplitude, and slope of ECAP I/O function) and sizes of the BCNC and VCN was analyzed using Pearson's correlation coefficients. Our analysis revealed that ECAP responses varied greatly among individual participants. Overall, ECAP thresholds gradually increased, while maximum amplitudes and ECAP I/O function slopes gradually decreased, as the electrode location moved from the basal to the apical direction in the cochlea. ECAP responses exhibited no significant correlations with BCNC width or VCN diameter. The ratio of the VCN to FN diameters was significantly correlated with the slope of the ECAP I/O function and the maximum amplitude. BCNC width could not predict the function of the cochlear nerve. Compared with the absolute size of the VCN, the size of the VCN relative to the FN may represent an indicator for predicting the functional status of the cochlear nerve in children diagnosed with cochlear nerve aplasia based on imaging results.

Keywords: cochlear nerve deficiency, cochlear implantation, imaging, electrically evoked compound action potential, cochlear nerve aplasia

INTRODUCTION

Cochlear nerve aplasia or hypoplasia is defined as an absent (aplasia) or a small (hypoplasia) cochlear nerve based on the results of magnetic resonance imaging (MRI). Clinically, children with deficient cochlear nerves often exhibit severe-to-profound sensorineural hearing loss (SNHL). The main treatment for hearing reconstruction in these children is cochlear implantation. However, children with cochlear nerve deficiency (CND) have poor outcomes after cochlear implantation. Previous studies reported that the benefits of cochlear implants (CIs) in patients with CND were worse than in other children with SNHL who had normal-sized cochlear nerves and varied greatly among individual children (Ehrmann-Muller et al., 2018; Arumugam et al., 2020). Although only a few patients can achieve simple open-set speech perception skills, most patients only exhibit improvements in sound awareness, and a few patients may experience no benefits following implantation (Kang et al., 2010; Kutz et al., 2011; Young et al., 2012; Vincenti et al., 2014). Unfortunately, to date, there are no effective methods for predicting the benefits of CIs preoperatively. Thus, providing appropriate counseling regarding the outcomes of cochlear implantation for children with CND remains challenging, as does determining the optimal ear in cases of unilateral implantation for children with bilateral CND.

Currently, the diagnosis of CND mainly depends on imaging findings. CND is diagnosed if the cochlear nerve is absent or smaller than the adjunct facial nerve (FN) in the internal auditory canal on MRI (Casselman et al., 1997; Sennaroglu and Bajin, 2017). In addition, temporal bone high-resolution computed tomography (HRCT) can be helpful for assessing the health of the cochlear nerve. The cochlear nerve is considered hypoplastic or aplastic when the diameter of the bony cochlear nerve canal (BCNC) is <1.5 mm and the diameter of the internal auditory canal is <2 mm (Miyasaka et al., 2010; Yan et al., 2013). The electrically evoked auditory brainstem response to the application of an intracochlear testing electrode is also an important indicator of the integrity of the cochlear nerve (Lassaletta et al., 2017). However, this assessment is somewhat invasive and traumatic and requires anesthesia. Previous studies have demonstrated that the morphology of the cochlear nerve or vestibulocochlear nerve (VCN) and the width of the BCNC might predict the degree of CND. Several studies have demonstrated that patients with aplastic cochlear nerves tend to perform worse than those with hypoplastic cochlear nerves following cochlear implantation (Kutz et al., 2011; Wu et al., 2015; Peng et al., 2017). Additionally, the width of the BCNC is positively correlated with the diameter of the cochlear nerve, and a narrower BCNC has been associated with more severe hearing loss and lower speech discrimination scores than a wider BCNC (Purcell et al., 2015). Thus, the size of the cochlear nerve or VCN and the width of the BCNC on imaging might be indicators of the severity of CND. However, conflicting results have been reported in previous studies regarding the relationship between preoperative imaging results and postoperative CI outcomes for children with CND. Although some studies have reported better auditory performance in children with normal BCNC than in those with BCNC stenosis (Chung et al., 2018; Kang et al., 2019), other studies have not revealed any predictive value of BCNC width for auditory or speech performance in children with CND (Warren et al., 2010; Tahir et al., 2020). In addition, the results of some studies have indicated an association between a larger VCN size in relation to the FN and better CI outcomes (Yamazaki et al., 2015; Han et al., 2019), while others have found no correlation between the size of the cochlear nerve or VCN and postimplantation auditory performance (Chao et al., 2016; Jain et al., 2020).

Although the health of the cochlear nerve is a critical factor affecting the postoperative effects of cochlear implantation, implantation age, history of hearing aid use, cognitive ability, parental socioeconomic status, and language training are also factors contributing to the outcomes of cochlear implantation. Thus, the relationships between preoperative imaging results and cochlear nerve function remain unclear. A better understanding of the value of preoperative imaging in predicting the degree of the cochlear nerve lesion could enable us to better assess CI candidacy and provide appropriate patient counseling for the benefits of CI for individual children with deficient cochlear nerves.

Recently, electrically evoked compound action potentials (ECAPs) have been widely used to evaluate cochlear nerve function in patients with implants (He et al., 2017). The ECAP response is generated by a group of auditory nerve fibers that are activated by electrical stimuli. It could be recorded using the "reverse" telemetry function implemented in current CI devices. Previous studies have shown that the slope of the ECAP input/output (I/O) function and the ECAP amplitude evoked by the most comfortable level are associated with the density of the surviving neural population, with steeper slopes and larger amplitudes suggesting a larger number of residual neurons (Miller et al., 2008; Pfingst et al., 2015). In addition, the ECAP threshold, which refers to the lowest stimulation level that could evoke an ECAP response, may also reflect the neural population to some extent (Ramekers et al., 2014). Thus, the aims of this study were to assess the function of the cochlear nerve using ECAP responses for individual children diagnosed with cochlear nerve aplasia based on imaging results and to analyze the correlation between imaging results (width of the BCNC and size of the VCN) and ECAP responses.

MATERIALS AND METHODS

Ethics Statements

This study was approved by the Ethical Committee of Shandong Provincial ENT Hospital affiliated with Shandong University (No. XYK20170906). Informed consent was obtained from the legal guardians of participants prior to participation.

Study Design and Population

This cohort study included 35 children (CND1–CND35; 11 boys and 24 girls), with cochlear nerve aplasia diagnosed by MRI. All participants were implanted with the Cochlear® Nucleus device

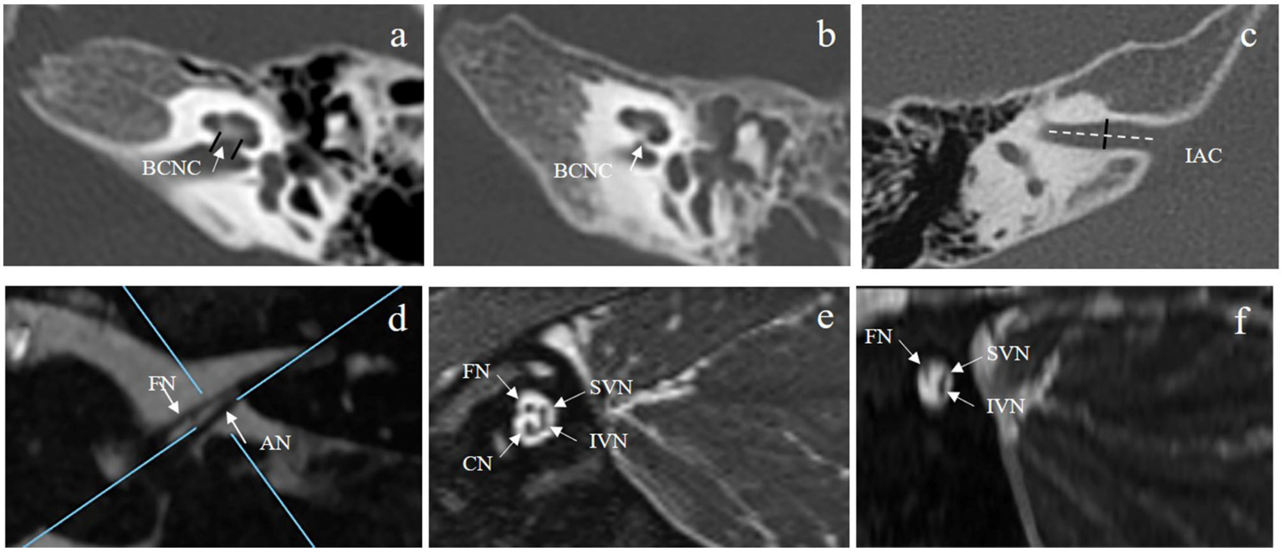

FIGURE 1 | (A,B) show the normal BCNC at the mid-modiolar level on the axial plane of HRCT **(A)** (two short black lines and white arrow) and a case of BCNC stenosis **(B)**. **(C)** shows the IAC diameter measured at the middle of the IAC on the axial plane of HRCT (short black line). **(D)** shows the vestibulocochlear nerve and the facial nerve at the cerebellopontine angle on the axial plane of MRI. The blue lines illustrate the plane prescribed for oblique plane sagittal images obtained perpendicular to the nerves of the IAC. **(E)** shows the cochlear, facial, superior vestibular, and inferior vestibular nerves on a reconstructed image in a patient with a normal cochlear nerve. **(F)** shows that the cochlear nerve could not be observed for one participant with cochlear nerve aplasia. BCNC, bone cochlear nerve canal; HRCT, high-resolution computed tomography; IAC, inner auditory canal.

(Cochlear Ltd.) in one or both ears. Children were included only if they had raw HRCT and MRI data that could be reconstructed and reanalyzed.

Radiological Assessment

All participants underwent MRI and HRCT for evaluation of the cochlear nerve status and other inner ear malformations before the operation in accordance with the previously described protocols (Chao et al., 2016). HRCT was performed using a 64-slice multidetector CT scanner (Somatom Sensation Cardiac 64; Siemens, Munich, Germany) using a standard temporal bone protocol. The main parameters for HRCT were as follows: the tube voltage was 120 kV; the tube current was automated tube current modulation (CareDose4D, Siemens); the slice thickness was 0.6 mm; the window width and window level were 4,000 HU and 600 HU, respectively. Axial and coronal images were obtained. Then, the axial images were reconstructed parallel to the lateral semicircular canal in a standard plane. The BCNC was evaluated between the anterior lower part of the bottom of the inner ear canal and the cochlear axis on axial HRCT, and the width of the BCNC was measured in the middle of the BCNC (**Figure 1A**). BCNC stenosis was diagnosed when BCNC width was <1.5 mm (**Figure 1B**), and the absence of the BCNC on any plane on HRCT was defined as atresia (Purcell et al., 2015; Lim et al., 2018). The width of the midportion of the inner auditory canal (IAC) was also measured at the level of the porus acusticus, from its posterior margin to the anterior wall of the IAC along a line orthogonal to the long axis of the IAC (**Figure 1C**) (Purcell et al., 2015). IAC stenosis was defined as an IAC width <3 mm. In addition, the structure of the inner ear was evaluated for any malformations.

Magnetic resonance imaging was performed using a clinical 3.0T MRI system (MAGNETOM Verio; Siemens) equipped with a 64-channel array head and neck coil. MRI sequences included axial T1-weighted and T2-weighted (T2W) imaging and three-dimensional fast spin-echo T2W sequences. Cochlear nerves and VCNs were evaluated on T2W axial and three-dimensional fast spin-echo T2-weighted sequences. The main parameters for the T2W sequences were as follows: the field of view was 162 × 82 mm; the repetition time was 1,200 ms; the echo time was 125 ms; the image matrix was 320 × 164; and the slice thickness was 0.5 mm. The main parameters for the three-dimensional fast spin-echo T2W sequences were as follows: the field of view was 220 × 220 mm; the repetition time was 1,200 ms; the echo time was 129 ms; the image matrix was 320 × 164; and the slice thickness was 0.2 mm. The diameters of the VCN and FN were measured at the cerebellopontine angle (**Figure 1D**), and the ratio of the VCN diameter to the FN diameter (VCN/FN ratio) was calculated. The VCN was characterized as hypoplastic if the nerve diameter was <1.5 mm or smaller than the FN with normal function (Sennaroglu, 2010). Direct oblique sagittal images perpendicular to the long axis of the IAC were reconstructed to show the cochlear, facial, superior vestibular, and inferior vestibular nerves in the IAC (**Figure 1E**). Cochlear nerve hypoplasia was diagnosed if the diameter of the cochlear nerve was smaller than that of the adjacent FN or the cochlear nerve in the contralateral ear. Cochlear nerve aplasia was diagnosed if the cochlear nerve could not be identified on any plane of the MRI (**Figure 1F**). All imaging results were reviewed by two experienced radiologists and an otologist. The results were defined as the average values measured by three persons.

Measurement of ECAPs

Electrically evoked compound action potentials were measured using the advanced neural response telemetry function implanted in the Custom Sound EP (version 4.3) software (Cochlear Ltd., Sydney, Australia). For each participant, the maximum comfort level was tested for each electrode before the ECAP recording. This level was defined as the largest stimulation level at which the participants felt comfortable. For children who could not provide a behavioral response, this level was defined as the largest stimulation level that would not cause discomfort. Two pulse forward masking methods were used to record the ECAP waveforms in this study. The stimulation and recording parameters used to record the ECAP were selected according to a previously described protocol (He et al., 2020). The stimulus was a single cathodic-leading biphasic charge-balanced pulse. The masker-to-probe interval was 400 μs, the probe rate was 15 Hz, the pulse width varied across individuals from 37 to 75 μs/phase, and the inter-phase gap was 7 μs. The recording electrode was placed two or three electrodes away from the stimulating electrode in the basal direction with a sampling delay of 98–142 μs. These parameters were adjusted for each participant to minimize artifacts and obtain optimized ECAP morphologies. First, ECAP responses were recorded from each electrode along the electrode array. Second, the ECAP I/O function was measured at three electrode locations where the ECAP waveforms could be recorded. Electrodes 3, 12, and 21 were selected for participants whose ECAPs could be recorded

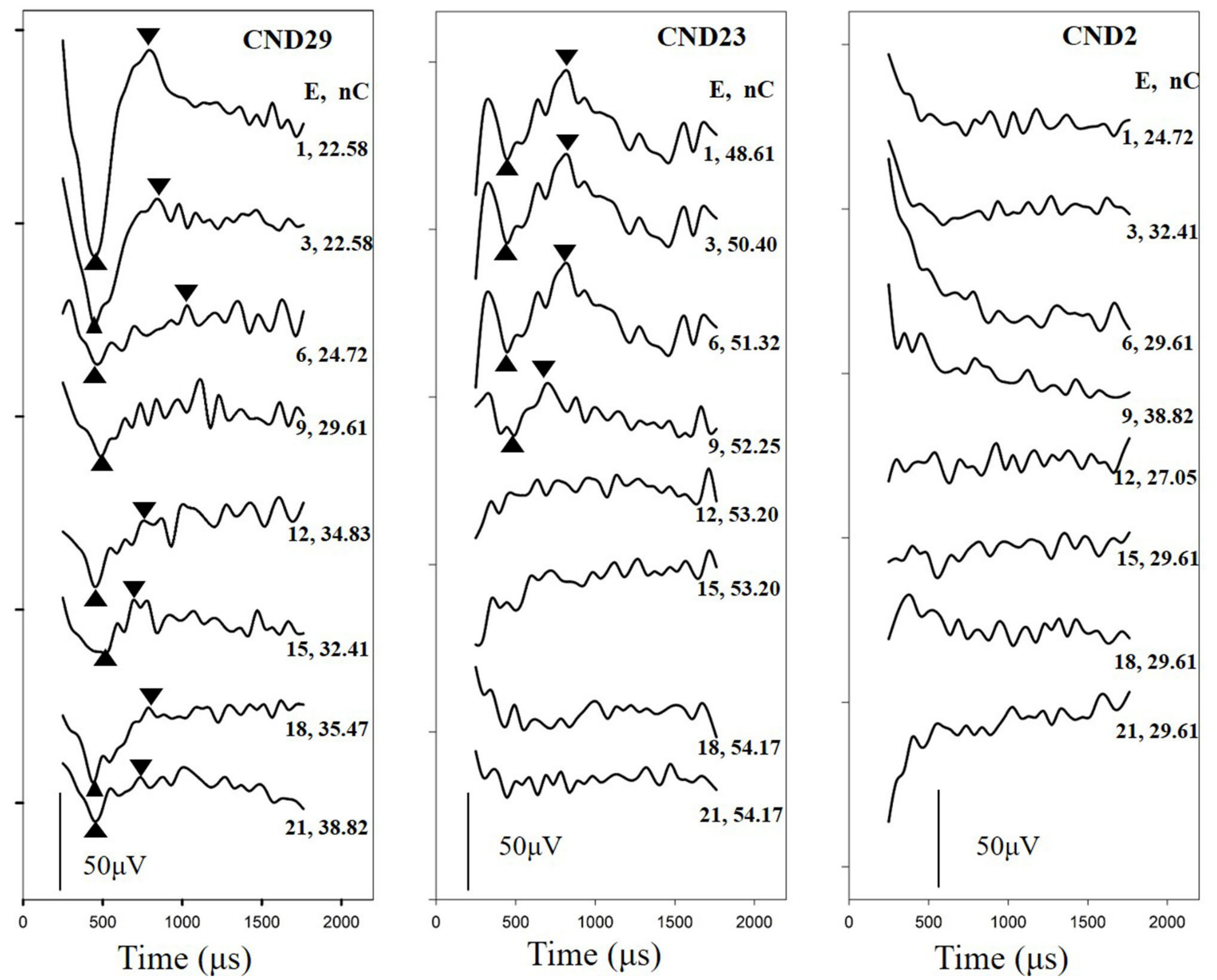

FIGURE 2 | Waveforms of ECAP responses were recorded from multiple electrodes in three children with cochlear nerve aplasia diagnosed based on MRI images. For CND29, the ECAP was recorded from all electrodes; for CND23, the ECAP was recorded from electrodes 1–9; and for CND2, the ECAP could not be recorded from any electrodes. Tested electrodes and stimulation levels are shown on the right of each panel. E, tested electrode; nC, stimulation level; ECAP, electrically evoked compound action potential.

at all electrode locations. For participants in whom ECAPs could only be recorded at some electrode locations, the selected electrodes were extended to the most apical electrode location with a measurable ECAP, and the testing electrodes were equally separated. The selected electrodes were defined as the basal, middle, and apical electrodes in this study. **Figure 2** shows the

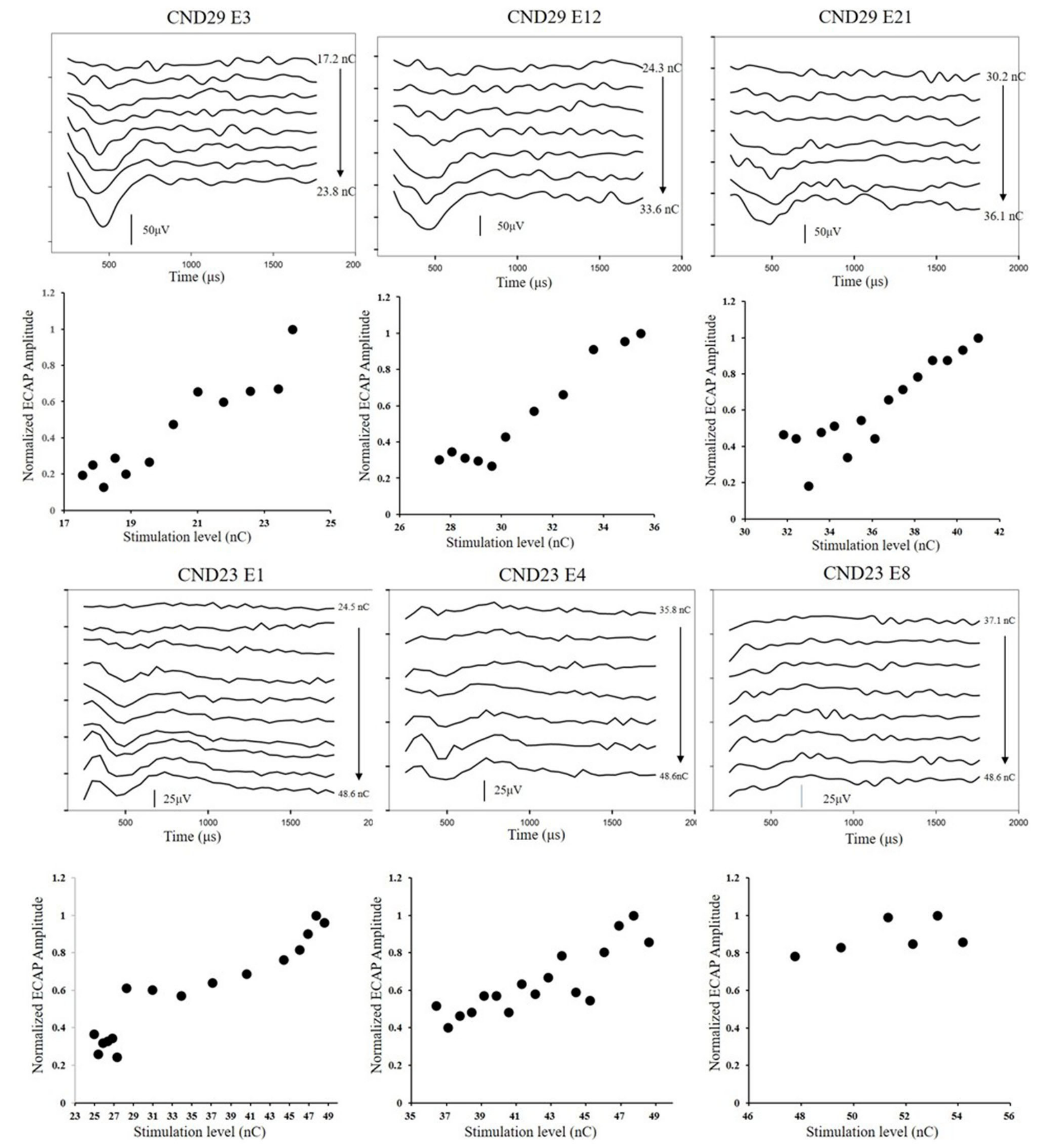

FIGURE 3 | ECAP response traces and ECAP I/O functions measured at three electrode locations in CND29 (upper) and CND23 (lower). Participants and electrode numbers are displayed in each panel. The round dots represent normalized ECAP amplitudes measured at different stimulation levels. E, electrode; ECAP, electrically evoked compound action potential; CND, cochlear nerve deficiency; I/O, input/output.

waveforms of the ECAP response from multiple electrodes for three participants. For CND29, the ECAP was recorded from all electrodes; for CND23, ECAPs were recorded from electrodes 1–9; and for CND2, the ECAP could not be recorded from any electrodes.

For the ECAP I/O function, the probe level was started at the C level and decreased in steps of two to three current levels (CLs) until no response could be visually identified and was subsequently increased in steps of 1 CL until five continuous ECAPs were measured. The ECAP threshold was defined as the lowest stimulation level that could evoke an ECAP with an amplitude ≥5 μV. Another five continuous ECAP traces below the ECAP thresholds were tested using a step size of 1 CL. For each participant, it took approximately 2–3 h to collect all the ECAP threshold data.

Data and Statistical Analyses

All ECAP thresholds were determined based on a mutual agreement between two audiologists who reviewed the data independently. As the pulse widths used among the participants were different, ECAP thresholds were converted to units of electrical charge per phase (nC). The ECAP amplitude was defined as the difference in the amplitude between the N1 and P2 peaks of the response. The slope of the ECAP I/O function was estimated using a sigmoidal regression function, as illustrated in previous studies (He et al., 2018). **Figure 3** displays ECAP response traces and ECAP I/O functions measured at three electrode locations in CND29 (upper) and CND23 (lower). ECAP amplitudes were normalized to the ECAP response tested at the maximum stimulation level. For both participants, ECAP thresholds gradually increased, while the maximum amplitudes gradually decreased, from basal to apical electrode sites. In this study, we measured ECAP responses that can represent cochlear nerve function, including the ECAP threshold, ECAP maximum amplitude, and slope of the ECAP I/O function.

The ECAP results tested at different electrode locations within participants were compared using the repeated-measures analyses of variance (ANOVA) test. In this study, the ECAP threshold, maximum ECAP amplitude, and slope of the ECAP I/O function for individual participants were defined as the mean values of the results tested at the three electrodes along the electrode array. The correlations between ECAP responses and radiological findings for all participants were analyzed using the Pearson correlation test. All statistical analyses were conducted using SPSS Statistics 21 (IBM Corp., Armonk, NY, USA). Statistical significance was set at $p < 0.05$.

RESULTS

Demographic and Clinical Characteristics

Four children underwent bilateral cochlear implantation, and all other children underwent unilateral cochlear implantation. Except for CND33, all participants were implanted with a contour electrode array, either a 24RE[CA] or CI512, in the test ear. CND33 underwent bilateral cochlear implantation and was implanted with the CI422 in the right ear and CI512 in the left ear. The electrode arrays were fully inserted in each ear. The participants' age at implantation ranged from 0.9 to 7.9 (mean: 2.8; standard deviation [SD]: 1.6) years. The test age ranged from 1.9 to 12.1 (mean: 4.9; SD: 2.6) years. All participants had normal FN function on implanted sides. In addition, no participants exhibited severe developmental delay or genetic-related hearing loss syndrome. The detailed demographic characteristics of the participants are shown in **Table 1**.

Imaging Results

Imaging findings for the individual participants are shown in **Table 1**. All tested ears, except for one, had a BCNC width <1.5 mm. BCNC atresia was observed in two ears, for which the width was defined as 0 mm. The mean BCNC width was 0.76 (SD: 0.28; range: 0–1.62) mm. On MRI, the cochlear nerve was absent on the reconstructed scans traversing the IAC in a perpendicular orientation and any plane of axial T2W sequence in all tested ears. All tested ears except for two had two nerve bundles, namely, the VCN and FNs, in the IAC on axial MRI imaging. The mean diameter of the VCN was 1.46 (SD: 0.36; range: 0.45–2.17) mm, and the mean diameter of the FN was 1.45 (SD: 0.40; range: 0.30–1.99) mm. The diameter of the VCN was smaller than that of the adjacent FN for 17 ears (44%), and the mean ratio of the VCN/FN diameter was 1.07 (SD: 0.32; range: 0.69–2.14). The ratio of the VCN/FN diameter was <1.5 in 31 ears (79%). In addition, the mean diameter of the IAC was 3.47 (SD: 1.10; range: 1.48–5.53) mm. Overall, 12 ears (31%) had IAC stenosis, with the diameter of the IAC ranging from 1.48 to 2.96 mm, while the other 27 ears (69%) had a normal IAC. In addition, five participants had Mondini malformation, and all the others had normal cochlear formation. Nine participants had vestibular or/and semicircular canal malformation, and all the others had normal vestibular and semicircular formation.

Electrically Evoked Compound Action Potential Responses

Electrically evoked compound action potential responses were recorded at all activated electrodes in 22 ears (56%) but could not be recorded at any activated electrodes in two ears (6%). For the remaining 15 ears (38%), the ECAP response could only be recorded at some of the electrodes. The proportion of electrodes with measurable ECAP responses was 74.13%. Electrodes with ECAP responses and the tested electrodes for each participant are displayed in **Table 1**. **Figure 4** shows the ECAP threshold, the maximum ECAP amplitude, and the slope of the ECAP I/O function at the three electrode locations and the mean values. At the basal, middle, and apical electrode locations, the mean ECAP thresholds were 18.80 (SD: 5.95; range: 10.50–36.77) nC, 23.96 (SD: 6.40; range: 13.28–38.32) nC, and 27.26 (SD: 8.58; range: 11.71–47.74) nC, respectively; the mean maximum ECAP amplitudes of ECAP were 74.28 (SD: 74.13; range: 9.73–456.45) μV, 48.02 (SD: 20.95; range: 106.56–24.08) μV, and 41.64 (SD: 26.52; range: 146.52–8.2) μV, respectively; and the mean slopes of ECAP I/O function were 2.33 (SD: 1.94; range: 8.70–0.11), 2.04 (SD: 1.44; range: 6.46–0.02), and 1.55 (SD: 1.36; range: 5.03–0.01), respectively. The ECAP results for each electrode are included in the **Supplementary Material 1**. As the electrode location moved from the basal to apical direction, the ECAP

TABLE 1 | Demographic information, imaging results and tested electrodes for each participant.

Number	Gender	Ear tested	Age at implantation	Age at testing	Electrode array	Width of the BCNC (mm)	Diameter of VCN nerve (mm)	Diameter of facial nerve (mm)	Width of the IAC (mm)	Electrodes with ECAP	Tested electrodes
CND1	F	L	3.36	5.4	24RECA	0.37	1.94	1.73	3.66	1–8	3, 10, 18
CND2	F	L	0.94	2.4	24RECA	0.57	1.9	1.86	3.98	0	None
CND3	M	L	2.81	5.5	24RECA	0.9	1.86	1.48	4.53	1–5, 19–21	1, 4, 21
CND4	F	R	1.55	4.0	24RECA	1.11	1.7	1.7	4.82	1–22	3, 12, 21
CND5	F	R	1.95	3.4	24RECA	0.8	0.99	1.36	1.96	9–10, 17–19, 21	9, 18, 21
CND6	F	R	1.30	2.6	24RECA	0.75	1.62	1.8	1.59	0	None
CND7	M	R	1.86	2.3	24RECA	0.77	0.9	0.5	3.1	1–22	3, 12, 21
CND7	M	L	1.86	2.4	24RECA	0.69	1.6	1.4	3	1–15	1, 7, 14
CND8	M	R	2.10	4.1	24RECA	0.63	1.64	1.86	2.5	1–8	1, 5, 8
CND9	M	L	1.24	2.3	24RECA	0.68	1.6	1	5.24	1–22	3, 12, 21
CND10	F	L	1.31	3.8	24RECA	0.93	1.1	1.3	5.53	1–7	1, 4, 7
CND11	F	L	5.95	8.0	24RECA	0	1.45	1.42	2.96	1–22	3, 12, 21
CND12	F	R	2.36	6.5	24RECA	0.82	2.08	1.35	4.83	1–22	3, 12, 21
CND13	M	L	5.60	11.2	24RECA	0	1.6	1.1	4.34	1–22	3, 12, 21
CND14	M	R	1.34	2.2	24RECA	0.56	1.32	1.51	3.54	1–22	3, 12, 21
CND14	M	L	1.34	2.0	24RECA	0.38	1.54	1.99	3.68	1–22	3, 12, 21
CND15	F	R	4.01	6.9	24RECA	1.04	1.1	0.9	3.96	1–22	3, 12, 21
CND16	F	L	4.44	8.1	24RECA	1.62	0.81	1.05	1.48	1–7	1, 3, 7
CND17	F	L	4.65	6.5	24RECA	0.97	1.53	1.87	3.48	1–22	3, 12, 21
CND18	M	L	7.93	12.1	24RECA	0.83	1.02	1.24	2.65	1–17	3, 10, 17
CND19	M	L	3.85	6.1	24RECA	0.77	1.7	1.5	3.81	1–7	1, 4, 7
CND20	M	R	1.93	4.1	24RECA	0.77	1.57	1.42	5.35	1–22	3, 12, 21
CND21	F	R	1.40	4.1	24RECA	1.07	1.3	1.7	5.22	1–9	1, 5, 9
CND22	M	R	2.61	6.8	24RECA	1.07	1.15	1.52	2.45	1–22	3, 12, 21
CND23	F	L	4.06	7.3	24RECA	0.53	1.6	1.64	4.73	1–8	1, 4, 8
CND24	F	R	1.20	2.2	24RECA	0.7	1.43	1.56	3.66	1–8	1, 4, 8
CND25	F	L	2.57	4.3	24RECA	0.76	1.65	1.51	3.04	1–11	1, 5, 11
CND26	F	R	2.93	8.5	24RECA	0.72	2.17	1.54	3.69	1–8	1, 4, 8
CND27	F	R	1.52	4.0	24RECA	0.91	1.22	0.57	3.2	1–22	3, 12, 21
CND27	F	L	1.52	2.1	24RECA	1.1	0.5	0.72	3.16	1–22	3, 12, 21
CND28	F	R	2.98	4.1	24RECA	0.7	1.7	1.4	3.05	1–22	3, 12, 21
CND29	M	L	1.94	2.5	24RECA	0.9	1.55	1.61	4.52	1–22	3, 12, 21
CND30	F	R	1.77	3.7	24RECA	0.69	1.42	1.7	4.02	1–22	3, 12, 21
CND31	F	R	1.91	3.4	24RECA	0.8	2.05	1.57	3.6	1–22	3, 12, 21
CND32	F	R	4.64	7.8	24RECA	0.87	VCN aplasia		1.86	1–22	3, 12, 21
CND33	F	R	2.63	3.4	CI422	0.9	VCN aplasia		2	1–22	3, 12, 21
CND33	F	L	2.63	3.7	CI512	0.8	1.5	0.8	1.95	1–15	3, 9, 15
CND34	F	L	3.06	4.2	24RECA	0.6	1.2	1.54	2.8	1–22	3, 12, 21
CND35	F	L	1.36	1.9	24RECA	0.72	0.45	0.3	2.36	1–22	3, 12, 21

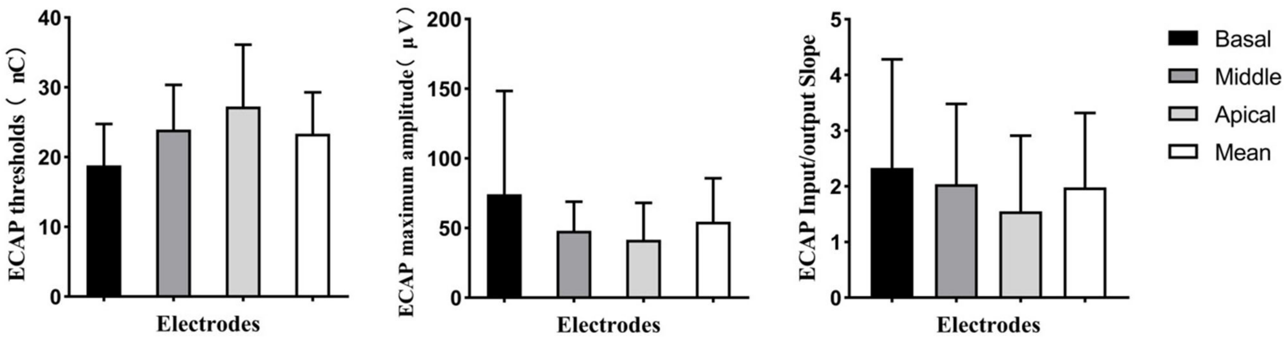

FIGURE 4 | Means and standard deviations for ECAP threshold, maximum ECAP amplitude, and slope of the ECAP input/output function for basal, middle, and apical electrodes. The mean value of the ECAP at these three electrodes is also shown. ECAP, electrically evoked compound action potential.

thresholds gradually increased and the maximum amplitudes and slopes of the ECAP I/O function gradually decreased. The repeated-measures ANOVA indicated that the electrode location had a significant effect on the ECAP thresholds ($F = 30.34$, $p < 0.01$), the maximum ECAP amplitude ($F = 5.91$, $p < 0.01$), and the slope of ECAP I/O function ($F = 4.95$, $p = 0.01$). When analyses were performed according to the participant, the mean ECAP threshold was 23.34 (SD: 5.93; range: 12.55–40.70) nC, mean maximum amplitude was 54.65 (SD: 31.04; range: 14.86–133.87) μV, and the mean slope of ECAP I/O function was 1.98 (SD: 1.34; range: 0.29–5.65).

Relationship Between Imaging Results and ECAP Responses

Table 2 shows the relationships of the BCNC width, VCN diameter, and VCN/FN ratio with the ECAP threshold, the ECAP maximum amplitude, and the slope of the ECAP I/O function. BCNC width was not significantly correlated with the slope of the ECAP I/O function ($r = -0.03$, $p > 0.05$), the maximum amplitude ($r = -0.06$, $p > 0.05$), or ECAP thresholds ($r = 0.15$, $p > 0.05$). There were also no significant correlations between the VCN diameter and the slope of the ECAP I/O function ($r = -0.9$, $p > 0.05$), the maximum amplitude ($r = 0.08$, $p > 0.05$), or ECAP thresholds ($r = 0.09$, $p > 0.05$). The VCN/FN ratio was significantly correlated with the slope of the ECAP I/O function ($r = 0.65$, $p < 0.01$) and the maximum amplitude ($r = 0.61$, $p < 0.01$) but not with the ECAP threshold ($r = -0.74$, $p > 0.05$).

TABLE 2 | Correlation of bone cochlear nerve canal (BCNC) width, vestibulocochlear nerve (VCN) diameter and VCN to facial nerve (FN) ratio to ECAP threshold, ECAP maximum amplitude and slope of ECAP Input/Output function.

	Width of BCNC ($n = 37$)	Diameter of VCN ($n = 35$)	VCN/FN ratio ($n = 35$)
ECAP thresholds	$r = 0.15$ $P = 0.36$	$r = 0.09$ $P = 0.59$	$r = -0.07$ $P = 0.67$
ECAP maximum amplitude	$r = -0.06$ $P = 0.72$	$r = -0.08$ $P = 0.96$	$r = 0.61$** $P < 0.01$
Slope of ECAP input/output function	$r = -0.03$ $P = 0.85$	$r = -0.09$ $P = 0.58$	$r = 0.65$** $P < 0.01$

** presents $p < 0.01$.

DISCUSSION

We first aimed to investigate cochlear nerve function using ECAP responses in individual children diagnosed with cochlear nerve aplasia based on imaging. In this study, the ECAP response was recorded in all but two participants. In participants with measurable ECAP responses, the percentage of electrodes with ECAP, ECAP thresholds, maximum amplitude, and slopes of the ECAP I/O function varied greatly among individual children, highlighting the variability of cochlear nerve function in individuals with cochlear nerve aplasia. Such variability may contribute to the various outcomes of cochlear implantation observed in children with cochlear nerve aplasia (Birman et al., 2016; Ehrmann-Muller et al., 2018; Yousef et al., 2021). Furthermore, in our patients, the slope of the ECAP I/O function and the maximum ECAP amplitude tended to gradually decrease, while the ECAP threshold tended to gradually increase, from basal to apical electrodes. This finding demonstrates that the responses of the cochlear nerve to electrical stimulation gradually decreased as the electrode location moved to a more apical location. These results are consistent with those of previous studies, which indicated that the degree of CND gradually worsens from the basal to the apical part of the cochlea in patients with CND (He et al., 2018; Xu et al., 2020). These specific characteristics of the cochlear nerve response to electrical stimulation for children with imaging-diagnosed cochlear nerve aplasia differ from those observed in children with normal-sized cochlear nerves diagnosed by MRI images. Previous studies have demonstrated that the electrode location exerts no significant effect on ECAP responses and that variations in ECAP results are much smaller among children with normal-sized cochlear nerves based on MRI (He et al., 2018). However, it was challenging to evaluate the function of the cochlear nerve for participants with no ECAP response in this study. Of the two participants without ECAP responses, one experienced some improvement in the ability to detect sound, while the other developed some close-set speech discrimination ability. Therefore, the absence of an ECAP response does not indicate the absence of a neural response in children diagnosed with cochlear nerve aplasia based on imaging results. A previous study indicated that the small ECAP responses for children with CND may be contaminated by artifacts related to electrical stimulation (He et al., 2020). In this study, we excluded the two children without ECAP responses from the correlation analysis.

The second aim of this study was to evaluate the possibility of using preoperative imaging results to predict the functional status of the cochlear nerve in children with imaging-diagnosed cochlear nerve aplasia. Research has indicated that the width of the BCNC is significantly smaller in children with imaging-diagnosed cochlear nerve aplasia than in other children with SNHL (Purcell et al., 2015). In this study, all but one child had BCNC stenosis, which further confirms that BCNC stenosis is a positive indicator of CND. Our results are consistent with those of previous reports. One such report indicated that ∼84% of ears with BCNC stenosis had a deficient cochlear nerve, while all ears with BCNC atresia had CND (Tahir et al., 2020). Although BCNC stenosis or atresia is used to diagnose CND, BCNC width was not significantly correlated with ECAP responses in our study, indicating that it cannot be used to predict the degree of CND. Several previous studies have reported worse CI outcomes among patients with BCNC stenosis than among patients with a normal BCNC, while these studies performed group comparisons between patients with BCNC stenosis (BCNC width <1.5 or 1.4 mm) and other groups (such as patients with a BCNC width >1.5 mm) (Kang et al., 2016, 2019; Chung et al., 2018). However, almost all participants in our study had BCNC. Whether patients with more severe BCNC stenosis perform worse postimplantation has seldom been reported. Chao et al. reported no significant relationship between BCNC width and auditory/speech performance in 10 children with CND (Chao et al., 2016). Furthermore, ECAP responses were recorded in participants with BCNC atresia in this study, which indicated that there should be some functional spiral ganglion neurons innervating the cochlea even in these patients. Previous studies have also reported that patients with BCNC atresia may experience some benefits with CIs (Warren et al., 2010; Tahir et al., 2020). Our results and those of previous studies indicate that the absence of the BCNC does not preclude the presence of the cochlear nerve. Thus, the collected results indicate that BCNC width is not related to cochlear nerve function in patients with BCNC stenosis who have been diagnosed with cochlear nerve aplasia based on imaging.

Furthermore, we investigated the influence of VCN size on the functional status of the cochlear nerve. Since the cochlear nerve could not be precisely assessed on MR images in children with imaging-diagnosed cochlear nerve aplasia, the size of the VCN was evaluated at the cerebellopontine angle on the axial plane of MRI. The diameter of the VCN tested in this study is significantly smaller than in other children with SNHL reported in previous studies (Nadol and Xu, 1992). Previous histological studies have indicated that the diameter of the VCN is significantly correlated with the number of spiral ganglion neurons (Nadol and Xu, 1992). In this study, there was no significant correlation between VCN diameter and ECAP responses, which indicates that the absolute size of the VCN does not predict cochlear nerve function in children with cochlear nerve aplasia. However, the VCN/FN ratio exhibited a significant correlation with the slope of the ECAP I/O function and ECAP maximum amplitude, indicating that the VCN diameter in relation to that of the FN may predict the functional status of the cochlear nerve. Previous histological studies have also highlighted great variability in the diameter of the VCN or FN in humans with normal hearing and those with hearing loss (Nadol and Xu, 1992; Nakamichi et al., 2013). Therefore, the diameter of the VCN is not suitable for predicting the number of residual spiral ganglion neurons. Herein, all participants had normal FN function, although the range for the diameter of the FN was large. We believe that the ratio of VCN to FN diameter can eliminate the influence of the variation in VCN diameter among patients. Previous studies have also investigated the relationship of the relative size of the VCN with CI outcomes in children with CND (Yamazaki et al., 2015; Chao et al., 2016; Han et al., 2019). Studies by Han et al. and Yamazaki et al. have demonstrated a significant correlation between the relative size of the VCN and Categories of Auditory Performance (CAP) scores. However, Chao et al. reported no significant correlation between relative VCN size and CAP scores. Further analysis showed that only 10 patients were included in Chao et al.'s study and the follow-up time was short, which would affect the differences in outcomes among patients. Overall, the relative size of the VCN may represent a sensitive indicator for predicting cochlear nerve function in children with imaging-diagnosed cochlear nerve aplasia.

Limitations

The main limitation of this study is that only three electrodes were used for each participant. In theory, the average ECAP results of all electrodes in the cochlea should be considered when examining the function of the cochlear nerve in each participant. However, electrodes exhibiting ECAP responses among participants were inconsistent, and it was time-consuming to test each electrode. Therefore, three representative cochlear electrodes were selected for this study. In addition to our study, previous studies have demonstrated that deficiency of the cochlear nerve progresses as a gradual increase from the basal to the apical region of the cochlea in children with CND (He et al., 2018; Xu et al., 2020). Thus, the average ECAP results from the basal, middle, and apical electrodes can roughly represent the function of the cochlear nerve. These three representative electrode sites were also used to estimate the function of the cochlear nerve in a previous study (Skidmore et al., 2020). In addition, patients with cochlear nerve hypoplasia have not been included in this study. It remains unclear whether there are some correlations between the relative size of the VCN and cochlear nerve function in children with cochlear nerve hypoplasia.

CONCLUSION

Children diagnosed with cochlear nerve aplasia based on MRI imaging exhibit variations in the functional status of the cochlear nerve. For children with cochlear nerve aplasia, the width of the BCNC does not predict cochlear nerve function, and the absence of the BCNC does not preclude the presence of the cochlear nerve. Compared with the absolute size of the VCN, the size of the VCN relative to the FN may represent an indicator for predicting the functional status of the cochlear nerve in children with imaging-diagnosed cochlear nerve aplasia. For these children, a larger VCN relative to the size of the FN may be associated with better CI outcomes.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of Shandong Provincial ENT Hospital affiliated with Shandong University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

XC participated in data collection and patient testing, prepared the initial draft of this manuscript, provided critical comments, and approved the final version of this manuscript. RW and JL participated in data analysis, provided critical comments, and approved the final version of this manuscript. ZF and HW provided critical comments and approved the final version of this manuscript. LX designed the study, participated in data collection and patient testing, and drafted and approved the final version of this manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from the National Natural Science Foundation of China (Nos. 81800905 and 82071053) and the Natural Science Foundation of Shandong Province Grant (No. ZR2016QZ007).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.905244/full#supplementary-material

Supplementary Material 1 | All the raw data for the ECAP responses are displayed in the Supplementary Material (S1).

REFERENCES

Arumugam, S. V., Nair, G., Paramasivan, V. K., Goyal, S., Murali, S., and Kameswaran, M. (2020). A study of outcome of pediatric cochlear implantation in patients with cochleovestibular nerve deficiency. *J. Int. Adv. Otol.* 16, 147–152. doi: 10.5152/iao.2020.8466

Birman, C. S., Powell, H. R., Gibson, W. P., and Elliott, E. J. (2016). cochlear implant outcomes in cochlea nerve aplasia and hypoplasia. *Otol. Neurotol.* 37, 438–445. doi: 10.1097/MAO.0000000000000997

Casselman, J. W., Offeciers, F. E., Govaerts, P. J., Kuhweide, R., Geldof, H., Somers, T., et al. (1997). Aplasia and hypoplasia of the vestibulocochlear nerve: diagnosis with MR imaging. *Radiology* 202, 773–781. doi: 10.1148/radiology.202.3.9051033

Chao, X., Luo, J., Fan, Z., Shi, H., Han, Y., Wang, R., et al. (2016). Usefulness of radiological findings for predicting cochlear implantation outcomes in children with cochlear nerve deficiency: a pilot study. *Acta Otolaryngol* 136, 1051–1057. doi: 10.1080/00016489.2016.1179788

Chung, J., Jang, J. H., Chang, S. O., Song, J. J., Cho, S. W., Kim, S. Y., et al. (2018). Does the width of the bony cochlear nerve canal predict the outcomes of cochlear implantation? *Biomed. Res. Int.* 2018, 5675848. doi: 10.1155/2018/5675848

Ehrmann-Muller, D., Kuhn, H., Matthies, C., Hagen, R., and Shehata-Dieler, W. (2018). Outcomes after cochlear implant provision in children with cochlear nerve hypoplasia or aplasia. *Int. J. Pediatr. Otorhinolaryngol.* 112, 132–140. doi: 10.1016/j.ijporl.2018.06.038

Han, J. J., Suh, M. W., Park, M. K., Koo, J. W., Lee, J. H., and Oh, S. H. (2019). A predictive model for cochlear implant outcome in children with cochlear nerve deficiency. *Sci. Rep.* 9, 1154. doi: 10.1038/s41598-018-37014-7

He, S., Chao, X., Wang, R., Luo, J., Xu, L., Teagle, H. F. B., et al. (2020). Recommendations for measuring the electrically evoked compound action potential in children with cochlear nerve deficiency. *Ear. Hear.* 41, 465–475. doi: 10.1097/AUD.0000000000000782

He, S., Shahsavarani, B. S., McFayden, T. C., Wang, H., Gill, K. E., Xu, L., et al. (2018). Responsiveness of the electrically stimulated cochlear nerve in children with cochlear nerve deficiency. *Ear. Hear.* 39, 238–250. doi: 10.1097/AUD.0000000000000467

He, S., Teagle, H. F. B., and Buchman, C. A. (2017). The electrically evoked compound action potential: from laboratory to clinic. *Front. Neurosci.* 11, 339. doi: 10.3389/fnins.2017.00339

Jain, S., Sharma, V., Patro, S. K., Khera, P., Yadav, T., Tiwari, S., et al. (2020). Correlation of cochlear nerve cross-sectional area and auditory performance after cochlear implantation in prelingual children with bilateral profound hearing loss. *Int. J. Pediatr. Otorhinolaryngol.* 137, 110173. doi: 10.1016/j.ijporl.2020.110173

Kang, B. C., Lee, J. Y., Kim, Y., Park, J. W., Kang, W. S., Ahn, J. H., et al. (2019). Outcome of cochlear implantation in children with narrow bony cochlear nerve canal. *Otol. Neurotol.* 40, e679–e685. doi: 10.1097/MAO.0000000000002263

Kang, D. H., Lee, M. J., Lee, K. Y., Lee, S. H., and Jang, J. H. (2016). Prediction of cochlear implant outcomes in patients with prelingual deafness. *Clin. Exp. Otorhinolaryngol.* 9, 220–225. doi: 10.21053/ceo.2015.01487

Kang, W. S., Lee, J. H., Lee, H. N., and Lee, K. S. (2010). Cochlear implantations in young children with cochlear nerve deficiency diagnosed by MRI. *Otolaryngol. Head Neck Surg.* 143, 101–108. doi: 10.1016/j.otohns.2010.03.016

Kutz, J. W. Jr., Lee, K. H., Isaacson, B., Booth, T. N., Sweeney, M. H., and Roland, P. S. (2011). Cochlear implantation in children with cochlear nerve absence or deficiency. *Otol. Neurotol.* 32, 956–961. doi: 10.1097/MAO.0b013e31821f473b

Lassaletta, L., Polak, M., Huesers, J., Diaz-Gomez, M., Calvino, M., Varela-Nieto, I., et al. (2017). Usefulness of electrical auditory brainstem responses to assess the functionality of the cochlear nerve using an intracochlear test electrode. *Otol. Neurotol.* 38, e413–e420. doi: 10.1097/MAO.0000000000001584

Lim, C. H., Lim, J. H., Kim, D., Choi, H. S., Lee, D. H., and Kim, D. K. (2018). Bony cochlear nerve canal stenosis in pediatric unilateral sensorineural hearing loss. *Int. J. Pediatr. Otorhinolaryngol.* 106, 72–74. doi: 10.1016/j.ijporl.2018.01.015

Miller, C. A., Brown, C. J., Abbas, P. J., and Chi, S. L. (2008). The clinical application of potentials evoked from the peripheral auditory system. *Hear Res.* 242, 184–197. doi: 10.1016/j.heares.2008.04.005

Miyasaka, M., Nosaka, S., Morimoto, N., Taiji, H., and Masaki, H. (2010). CT and MR imaging for pediatric cochlear implantation: emphasis on the relationship between the cochlear nerve canal and the cochlear nerve. *Pediatr. Radiol.* 40, 1509–1516. doi: 10.1007/s00247-010-1609-7

Nadol, J. B. Jr., and Xu, W. Z. (1992). Diameter of the cochlear nerve in deaf humans: implications for cochlear implantation. *Ann. Otol. Rhinol. Laryngol.* 101, 988–993. doi: 10.1177/000348949210101205

Nakamichi, R., Yamazaki, M., Ikeda, M., Isoda, H., Kawai, H., Sone, M., et al. (2013). Establishing normal diameter range of the cochlear and facial nerves with 3D-CISS at 3T. *Magn. Reson. Med. Sci.* 12, 241–247. doi: 10.2463/mrms.2013-0004

Peng, K. A., Kuan, E. C., Hagan, S., Wilkinson, E. P., and Miller, M. E. (2017). cochlear nerve aplasia and hypoplasia: predictors of cochlear implant success. *Otolaryngol. Head Neck Surg.* 157, 392–400. doi: 10.1177/0194599817718798

Pfingst, B. E., Hughes, A. P., Colesa, D. J., Watts, M. M., Strahl, S. B., and Raphael, Y. (2015). Insertion trauma and recovery of function after cochlear implantation: evidence from objective functional measures. *Hear Res.* 330, 98–105. doi: 10.1016/j.heares.2015.07.010

Purcell, P. L., Iwata, A. J., Phillips, G. S., Paladin, A. M., Sie, K. C., and Horn, D. L. (2015). Bony cochlear nerve canal stenosis and speech discrimination in pediatric unilateral hearing loss. *Laryngoscope* 125, 1691–1696. doi: 10.1002/lary.25087

Ramekers, D., Versnel, H., Strahl, S. B., Smeets, E. M., Klis, S. F., and Grolman, W. (2014). Auditory-nerve responses to varied inter-phase gap and phase duration of the electric pulse stimulus as predictors for neuronal degeneration. *J. Assoc. Res. Otolaryngol.* 15, 187–202. doi: 10.1007/s10162-013-0440-x

Sennaroglu, L. (2010). Cochlear implantation in inner ear malformations–a review article. *Cochlear Implants Int.* 11, 4–41. doi: 10.1002/cii.416

Sennaroglu, L., and Bajin, M. D. (2017). Classification and current management of inner ear malformations. *Balkan Med. J.* 34, 397–411. doi: 10.4274/balkanmedj.2017.0367

Skidmore, J., Xu, L., Chao, X., Riggs, W. J., Pellittieri, A., Vaughan, C., et al. (2020). Prediction of the functional status of the cochlear nerve in individual cochlear implant users using machine learning and electrophysiological measures. *Ear Hear.* 42, 180–192. doi: 10.1097/AUD.0000000000000916

Tahir, E., Cinar, B. C., Ozkan, H. B., Yarali, M., Boke, B., and Sennaroglu, L. (2020). Successful use of a cochlear implant in a patient with bony cochlear nerve canal atresia. *J. Int. Adv. Otol.* 16, 271–273. doi: 10.5152/iao.2020.3766

Vincenti, V., Ormitti, F., Ventura, E., Guida, M., Piccinini, A., and Pasanisi, E. (2014). Cochlear implantation in children with cochlear nerve deficiency. *Int. J. Pediatr. Otorhinolaryngol.* 78, 912–917. doi: 10.1016/j.ijporl.2014.03.003

Warren, F. M. 3rd, Wiggins, R. H. 3rd, Pitt, C., Harnsberger, H. R., and Shelton, C. (2010). Apparent cochlear nerve aplasia: to implant or not to implant? *Otol. Neurotol.* 31, 1088–1094. doi: 10.1097/MAO.0b013e3181eb3272

Wu, C. M., Lee, L. A., Chen, C. K., Chan, K. C., Tsou, Y. T., and Ng, S. H. (2015). Impact of cochlear nerve deficiency determined using 3-dimensional magnetic resonance imaging on hearing outcome in children with cochlear implants. *Otol. Neurotol.* 36, 14–21. doi: 10.1097/MAO.0000000000000568

Xu, L., Skidmore, J., Luo, J., Chao, X., Wang, R., Wang, H., et al. (2020). The effect of pulse polarity on neural response of the electrically stimulated cochlear nerve in children with cochlear nerve deficiency and children with normal-sized cochlear nerves. *Ear Hear* 41, 1306–1319. doi: 10.1097/AUD.0000000000000854

Yamazaki, H., Leigh, J., Briggs, R., and Naito, Y. (2015). Usefulness of MRI and EABR testing for predicting CI outcomes immediately after cochlear implantation in cases with cochlear nerve deficiency. *Otol. Neurotol.* 36, 977–984. doi: 10.1097/MAO.0000000000000721

Yan, F., Li, J., Xian, J., Wang, Z., and Mo, L. (2013). The cochlear nerve canal and internal auditory canal in children with normal cochlea but cochlear nerve deficiency. *Acta Radiol.* 54, 292–298. doi: 10.1258/ar.2012.110596

Young, N. M., Kim, F. M., Ryan, M. E., Tournis, E., and Yaras, S. (2012). Pediatric cochlear implantation of children with eighth nerve deficiency. *Int. J. Pediatr. Otorhinolaryngol.* 76, 1442–1448. doi: 10.1016/j.ijporl.2012.06.019

Yousef, M., Mesallam, T. A., Garadat, S. N., Almasaad, A., Alzhrani, F., Alsanosi, A., et al. (2021). Audiologic outcome of cochlear implantation in children with cochlear nerve deficiency. *Otol. Neurotol.* 42, 38–46. doi: 10.1097/MAO.0000000000002849

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

ORIGINAL RESEARCH
published: 23 June 2022
doi: 10.3389/fnins.2022.895560

Check for updates

Machine Learning-Based Prediction of the Outcomes of Cochlear Implantation in Patients With Cochlear Nerve Deficiency and Normal Cochlea: A 2-Year Follow-Up of 70 Children

Simeng Lu[1†], Jin Xie[2†], Xingmei Wei[1], Ying Kong[3], Biao Chen[1], Jingyuan Chen[1], Lifang Zhang[1], Mengge Yang[1], Shujin Xue[1], Ying Shi[1], Sha Liu[3], Tianqiu Xu[3], Ruijuan Dong[3], Xueqing Chen[3], Yongxin Li[1] and Haihui Wang[2]**

[1] Key Laboratory of Otolaryngology Head and Neck Surgery, Department of Otorhinolaryngology Head and Neck Surgery, Beijing Tongren Hospital, Ministry of Education, Capital Medical University, Beijing, China, [2] Laboratory of Haihui Data Analysis, School of Mathematical Sciences, Beihang University, Beijing, China, [3] Beijing Tongren Hospital, Beijing Institute of Otolaryngology, Capital Medical University, Beijing, China

OPEN ACCESS

Edited by:
Jing Chen,
Peking University, China

Reviewed by:
Alexandros-Apostolos A. Boulogeorgos,
University of Piraeus, Greece
Anandhan E. Dhanasingh,
MED-EL, Austria
Wei Liu,
Uppsala University, Sweden

***Correspondence:**
Yongxin Li
entlyx@sina.com
Haihui Wang
whhmath@buaa.edu.cn

† These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *14 March 2022*
Accepted: *31 May 2022*
Published: *23 June 2022*

Citation:
Lu S, Xie J, Wei X, Kong Y, Chen B, Chen J, Zhang L, Yang M, Xue S, Shi Y, Liu S, Xu T, Dong R, Chen X, Li Y and Wang H (2022) Machine Learning-Based Prediction of the Outcomes of Cochlear Implantation in Patients With Cochlear Nerve Deficiency and Normal Cochlea: A 2-Year Follow-Up of 70 Children. Front. Neurosci. 16:895560. doi: 10.3389/fnins.2022.895560

Cochlear nerve deficiency (CND) is often associated with variable outcomes of cochlear implantation (CI). We assessed previous investigations aiming to identify the main factors that determine CI outcomes, which would enable us to develop predictive models. Seventy patients with CND and normal cochlea who underwent CI surgery were retrospectively examined. First, using a data-driven approach, we collected demographic information, radiographic measurements, audiological findings, and audition and speech assessments. Next, CI outcomes were evaluated based on the scores obtained after 2 years of CI from the Categories of Auditory Performance index, Speech Intelligibility Rating, Infant/Toddler Meaningful Auditory Integration Scale or Meaningful Auditory Integration Scale, and Meaningful Use of Speech Scale. Then, we measured and averaged the audiological and radiographic characteristics of the patients to form feature vectors, adopting a multivariate feature selection method, called stability selection, to select the features that were consistent within a certain range of model parameters. Stability selection analysis identified two out of six characteristics, namely the vestibulocochlear nerve (VCN) area and the number of nerve bundles, which played an important role in predicting the hearing and speech rehabilitation results of CND patients. Finally, we used a parameter-optimized support vector machine (SVM) as a classifier to study the postoperative hearing and speech rehabilitation of the patients. For hearing rehabilitation, the accuracy rate was 71% for both the SVM classification and the area under the curve (AUC), whereas for speech rehabilitation, the accuracy rate for SVM classification and AUC was 93% and 94%, respectively. Our results identified that a greater number of nerve bundles and a larger VCN area were associated with better CI outcomes. The number of nerve bundles and VCN area can predict CI outcomes in patients with CND. These findings can help surgeons in selecting the side for CI and provide reasonable expectations for the outcomes of CI surgery.

Keywords: cochlear nerve deficiency, cochlear implantation, machine learning, stability selection, support vector machines

INTRODUCTION

Cochlear nerve deficiency (CND) is defined as a small or absent cochlear nerve (CN) (Adunka et al., 2007). When the CN is small, it is referred to as cochlear nerve hypoplasia (CNH). When the CN is absent, it is referred to as cochlear nerve aplasia (CNA). The estimated prevalence of CND is 18% among children with congenital sensorineural hearing loss (SNHL) (Jallu et al., 2015).

Cochlear implant (CI) was an effective treatment to restore hearing for patients with SNHL. The mechanism of cochlear implantation (CI) involves converting acoustic signals into electrical signals, directly stimulating the spiral ganglion neurons (SGNs), and transmitting the signals through the CN fibers to the auditory brainstem. Recent years, optics has been proposed to stimulate CN such as optical wireless CI and all-optical CI (Trevlakis et al., 2019, 2020). These architectures could convert acoustic to optical signals which improved the reliability and the efficiency of the transcutaneous link (Boulogeorgos et al., 2021).

In CND patients, due to the decrease in the absolute number of CN fibers, SGNs are insufficient and the stimuli that can be received are limited. In early studies, CND was considered a contraindication for CI (Shelton et al., 1989). However, a large number of studies have shown that some patients with CND can benefit from CI (Kang et al., 2010; Wu et al., 2015; Wei et al., 2017). When compared with patients without CND, patients with CND need higher stimulation to induce a CN response (Yousef et al., 2021). Generally, children with CND perform worse than those without CND (Wei et al., 2017; Yousef et al., 2021) and some patients even experience no benefit at all (Colletti et al., 2004). Due to the low incidence of CND and the uncertainty regarding the effects of CI surgery, there has been no study involving a large sample of patients with CND. At the same time, patients with CND were far more likely to exhibit inner ear malformations than patients without CND (Wu et al., 2015), which affected the number of SGNs and further limited the CI outcomes, making it difficult to determine whether the surgical results were different due to the differences in the surgical methods and electrode positions (Shi et al., 2019). Some studies have revealed the predictive role of radiographic information for CI outcomes in CND patients (Wu et al., 2015), but they did not include the patients with inner ear malformations.

Over the past 5 years, machine learning has been increasingly used to automate intelligent processes and improve the efficiency of medical processes. For example, cochlear implants can be enhanced by adopting machine learning techniques, which have been applied to create predictive models (Crowson et al., 2020; Velde et al., 2021) (see Velde et al., 2021 for a recent review). In addition, machine learning algorithms have been used to predict cochlear implantation (CI) outcomes. The prediction of postoperative CI performance from preoperative data may allow practitioners to evaluate implantation candidacy, estimate performance expectations from non-modifiable predictors, and optimize the procedure by intervening in modifiable predictors (Crowson et al., 2020). Han et al. (2019) established a multiple regression model for 25 CND patients with normal cochlea to predict CI postoperative outcomes, explaining 66% variance of the Categories of Auditory Performance (CAP) scores for patients with 2-year CIs. They concluded that the postoperative effect of CI in patients with CND was related to the preoperative auditory brainstem response (ABR) and the area ratio of the vestibulocochlear nerve (VCN) to the facial nerve (FN). Preoperative counseling based on this model helped determine the treatment modalities for hearing rehabilitation.

In this study, we analyzed the CI surgery-related factors in CND patients with normal cochlea, based on a relatively larger sample. We used a data-driven multivariate approach based on machine learning to evaluate postoperative hearing and speech rehabilitation in patients with CND and the influencing factors. We retrospectively analyzed the clinical data of CI surgery from 70 patients with CND and normal cochlea. Then, we measured and averaged audiological and radiological features of patients with CND to form feature vectors. Data-driven methods (i.e., stability selection and SVM) were applied to data from patients with CND to relate various factors to the CI outcomes and to build the corresponding predictive models. In addition, stability selection, a machine learning method that identifies highly consistent and representative features, was used to examine the factors that best differentiate the effects of postoperative hearing and speech rehabilitation in patients with CND.

MATERIALS AND METHODS

Participants

This study was approved by the Research Ethics Board of Tongren Hospital, Beijing, China. We considered 70 CI pediatric recipients (37 males and 33 females; ages 7–54 months) with CND who were diagnosed using three-dimensional MRI and who underwent CI between January 2012 and August 2018 in this study. All children failed to pass the newborn hearing screening sequence and the ear with better residual hearing was selected to undergo CI. Thirty-four patients were implanted with Med-El (Innsbruck, Austria) devices, 22 with Cochlear (Sydney, Australia) devices, 14 with AB (California, United States) devices. **Table 1** lists the demographic details with quantitative variables shown as count, mean, standard deviation, minimum, maximum,

TABLE 1 | Descriptive statistics of patients.

	Count	Mean	Standard deviation	Min.	25%	50%	75%	Max.
Age (months)	70.00	27.31	13.92	7.00	14.00	25.50	38.00	54.00
Residual hearing (dB)	70.00	108.17	14.07	81.00	97.50	106.88	124.69	125.00
Bony cochlear nerve canal diameter (mm)	70.00	0.83	0.58	0.01	0.35	0.83	1.17	2.28
Internal auditory canal diameter (mm)	70.00	2.47	0.85	0.41	1.91	2.51	2.96	4.42
Number of nerve bundles	70.00	1.71	0.80	1.00	1.00	2.00	2.00	4.00
Vestibulocochlear nerve area (mm^2)	70.00	1.32	0.55	0.30	0.91	1.27	1.71	2.78
Area ratio of the vestibulocochlear nerve to the facial nerve	70.00	1.50	0.48	0.44	1.22	1.47	1.69	2.63
40-Hz auditory-evoked related potential	70.00	–	–	–	–	–	–	–
Auditory brainstem responses	70.00	–	–	–	–	–	–	–
Cochlear microphonics	70.00	–	–	–	–	–	–	–
Distortion product otoacoustic emissions	70.00	–	–	–	–	–	–	–
Acoustic immittance	70.00	–	–	–	–	–	–	–
CAP	70.00	4.10	1.32	2.00	3.00	5.00	5.00	7.00
SIR	70.00	1.87	0.92	1.00	1.00	2.00	3.00	4.00
MAIS	70.00	25.14	10.47	3.00	18.00	26.50	33.75	40.00
MUSS	70.00	11.96	10.32	0.00	3.00	8.00	21.75	33.00

Not available values (–) indicate discrete variables without mean, standard deviation, maximum, minimum, and quantile.

and quantile, and qualitative variables shown as count. Hearing impairment was classified according to the World Health Organization classification (Olusanya et al., 2019) into mild (26–40 dB), moderate (41–60 dB), severe (61–80 dB), and profound (81 dB or greater). The inclusion criteria were as follows: (1) the diameter of the CN smaller than that of the FN or less than four nerve bundles within the internal auditory canal (IAC), (2) bilateral severe to profound SNHL, (3) no inner ear malformation or other congenital syndromes, (4) history of CI, and (5) completion of 2-year follow-up after CI.

Radiographic Examinations

High-resolution computed tomography was used to evaluate inner ear malformations according to Sennaroglu's classification (Sennaroğlu and Bajin, 2017). Normal cochlea was diagnose when normal cochlear appearance shown with current MRI and CT. The diameter of the bony cochlear nerve canal (BCNC) with the width of the canal at the midportion of the IAC fundus (**Figure 1A**) and the widest diameter of the IAC (**Figure 1B**) were measured on computed tomography images. The CN traverses along the fundus of the IAC to the base of the modiolus through the BCNC. Magnetic resonance imaging (MRI) was performed to determine the condition of the CN using a 1.5 Tesla scanner or a 3.0 Tesla scanner. The scan sequence is a 3D FIESTA water imaging sequence. Oblique sagittal reconstruction was performed perpendicular to the plane of the IAC. The areas of the VCN and FN at the cerebellopontine angle (CPA) were evaluated, as the cross-sections of these nerves were well visualized as this level. The area ratio of the VCN to the FN was evaluated at the CPA using MRI (**Figure 1C**). In addition, the number of nerve bundles was counted within the IAC (**Figure 2**).

Preoperative Auditory Evaluation

The diagnostic protocol for children with suspected hearing loss incorporated behavioral testing, acoustic emittance, distortion product otoacoustic emission (DPOAE), ABR, cochlear microphonics (CM), and 40-Hz auditory-evoked related potential (40-Hz AERP). The average hearing threshold was assumed to be 5 dB HL greater than the maximum output of the audiometer and was averaged across 0.5, 1.0, 2.0, and 4.0 kHz of pure-tone or behavioral testing.

Cochlear Implantation Device and Activation

The CI device was selected by the parents with the support and counseling by the CI team. Typically, the first mapping was initiated at 3–4 weeks after the surgery. During the programming sessions, observation and conditioned behavioral audiometry techniques were used to determine the electrical threshold and comfortable listening levels. Usually, a stable map can be achieved at 3–6 months after the initial stimulation.

Evaluation of Cochlear Implantation Outcomes

Postoperative speech evaluation was performed at 3, 6, 12, 18, and 24 months after CI. Since most of the patients had stable outcomes at 2 years, we selected the 2-year outcomes as the predicted results. The CAP, Speech Intelligibility Rating (SIR), Infant-Toddler Meaningful Auditory Integration Scale (IT-MAIS, for patients aged $<$ 3 years) or Meaningful Auditory Integration Scale (MAIS, for patients aged $>$ 3 years), and Meaningful Use of Speech Scale (MUSS) were used to evaluate hearing and speech in the patients 24 months after CI surgery. The CAP has eight levels of sound perception (0–7), ranging *from no awareness of the environment* (0) to *use of telephone with known users* (7). The CAP is intended to reflect the real-life auditory capabilities of children. The SIR is a highly reliable and time-effective measure of children's speech production in real-life situations and ranks children's spontaneous speech into five categories, ranging from connected speech is unintelligible (1) to connected speech is intelligible to all listeners (5). To distinguish the degree of patients' auditory performance and speech perception, we divided the patients into two groups according to CAP and SIR. **Figure 1** shows the initial distributions of CAP

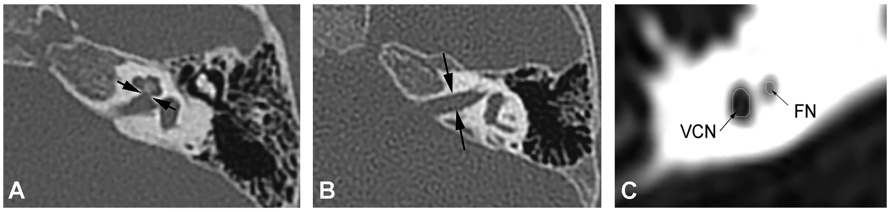

FIGURE 1 | (A) Measurement of the bony cochlear nerve canal diameter using high-resolution computed tomography (distance between the two black arrows). **(B)** Measurement of the internal auditory canal diameter using high-resolution computed tomography (distance between the two black arrows). **(C)** Measurement of the area ratio of the vestibulocochlear nerve to the facial nerve at the cerebellopontine angle using magnetic resonance imaging (black arrows).

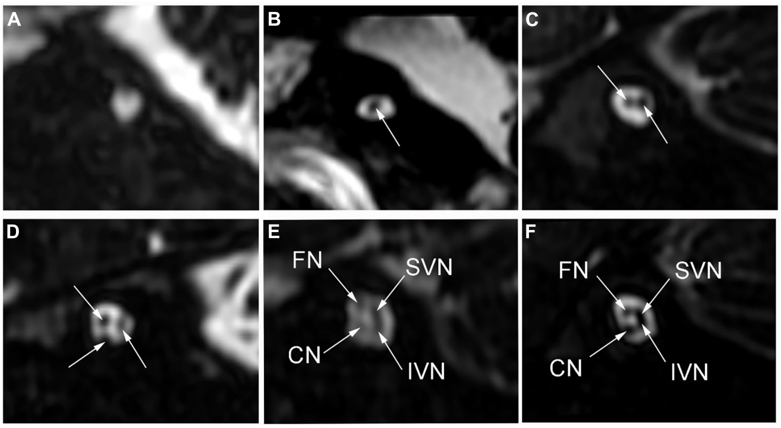

FIGURE 2 | The number of nerve bundles (indicated by the white arrows) in the internal auditory canal on oblique sagittal magnetic resonance imaging. **(A)** No nerve bundle, **(B)** one nerve bundle, **(C)** two nerve bundles, **(D)** three nerve bundles, **(E)** four nerve bundles (thin), **(F)** four nerve bundles (normal); CN: cochlear nerve, FN: facial nerve, IVN: inferior vestibular nerve, SVN: superior vestibular nerve.

(**Figure 3A**) and SIR (**Figure 3C**) and the distributions of CAP (**Figure 3B**) and SIR (**Figure 3D**) after grouping. For CAP, the patients were divided into spoken language understanding (CAP of 5–7) and no spoken language understanding (CAP of 0–4). For SIR, the patients were divided into intelligible speech (SIR of 2–5) and unintelligible speech (SIR of 1).

Feature Selection

Feature selection is essential in feature engineering as it aims to find an optimal subset of features and eliminate redundant features for classification. The effectiveness of hearing and speech rehabilitation after receiving a CI depends on various complex and interdependent factors. By considering the experience of doctors, we collected and measured these influencing factors. We measured the radiological and audiological characteristics of patients with CND multiple times and averaged the results. Radiological features include the bony cochlear nerve canal diameter, internal auditory canal diameter, number of nerve bundles, VCN area and area ratio of the VCN to the FN. Audiological features include residual hearing, 40-Hz AERP, ABR, CM, DPOAE, and acoustic immittance. In addition to radiological and audiological features, we considered the implantation age of patients with CND. To build a simple model, we applied feature selection to these 12 features. For feature selection, the variance threshold and stability selection are important methods. In brief, the variance threshold removes all features whose variance does not meet a certain threshold. By default, it removes all zero-variance features, that is, features with the same value across samples.[1] In this study, we performed preliminary feature selection using the variance threshold, graphical method and correlation method and then applied stability selection to the remaining features. These features were used as inputs to an SVM classifier and stability selection coupled with SVM. As usual for classifiers, we applied min-max normalization to the data before classification and stability selection to ensure that all features had a common scale and range.

Stability Selection

A robust model should be sufficiently complete to allow generalization and interpretation. Hence, the most salient discriminating features consistent across a range of model parameters should be selected. Stability selection achieves state-of-the-art feature selection while preventing overfitting and enabling data interpretability. In general, representative features do not score 0 for similar features or associative features. We used randomized logistic regression and the randomized least absolute shrinkage and selection operator (LASSO) for stability

[1] https://scikit-learn.org/stable/modules/feature_selection.html#feature-selection

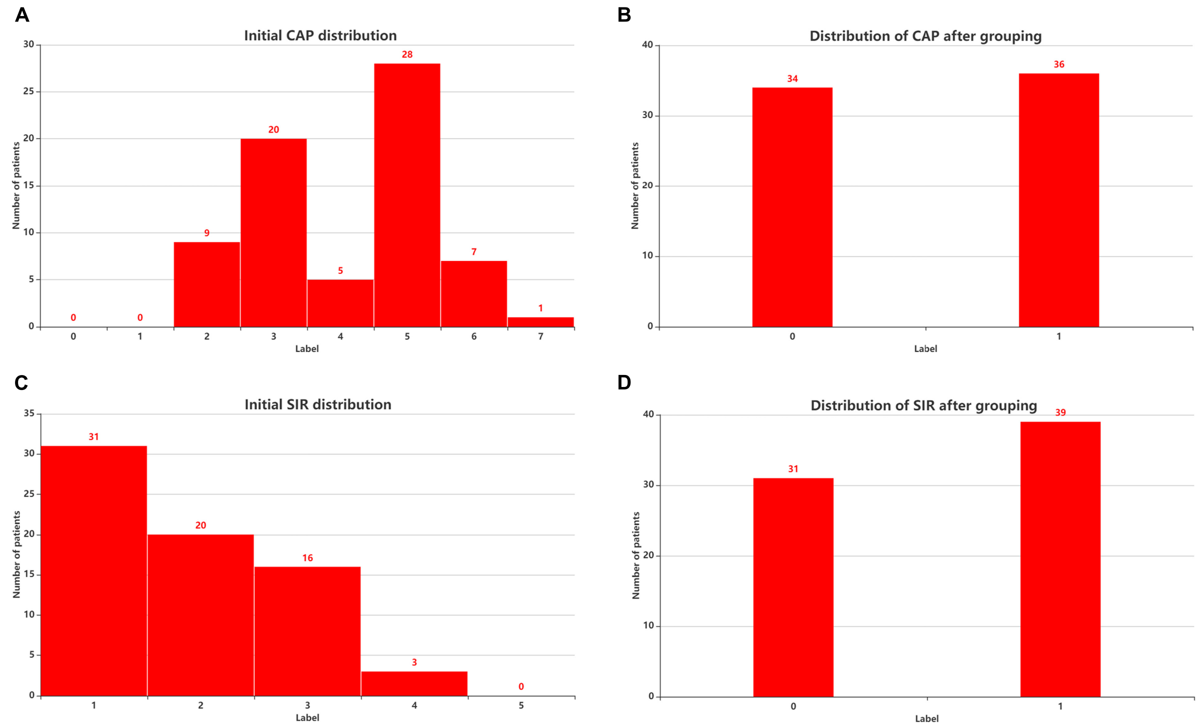

FIGURE 3 | Patient distribution before and after considering CAP and SIR grouping. **(A)** Initial CAP distribution. **(B)** Distribution of CAP after grouping. **(C)** Initial SIR distribution. **(D)** Distribution of SIR after grouping. CAP, categories of auditory performance; SIR, speech intelligibility rating.

selection. Labels CAP and SIR are discrete variables. Labels MAIS and MUSS are continuous variables. Therefore, the remaining 6 features and the label CAP or SIR were entered into the randomized logistic regression model; the remaining 6 features and the label MAIS or MUSS were entered into the randomized LASSO model. The stability score of the features to the labels was obtained, and the feature selection was made according to the stability score.

Stability selection used a randomized logistic regression algorithm, which worked by subsampling the training data and fitting a L1-penalized logistic regression model. By performing this double randomization several times (running logistic regression algorithms on different subsets of data and features), the method assigned high scores to features that were repeatedly selected across randomizations. In short, the features selected more often were considered as representative features (Meinshausen and Bühlmann, 2010). Stability selection used a randomized LASSO algorithm, which worked by subsampling the training data and computing a Lasso estimate (Meinshausen and Bühlmann, 2010). In stability selection, the feature stability increases as a feature is increasingly selected over repeated subsampling processes (Nogueira et al., 2017). As stability selection includes internal randomization over many interactions, it yields a more reliable and consistent feature set than conventional filtering or other multivariate approaches (Mahmud et al., 2020).

We considered the regularization parameter C of 1, the scaling parameter of 0.5, a sample fraction of 0.75 and 200 resampling processes to implement randomized logistic regression. The scaling parameter was used to randomly scale the features (Meinshausen and Bühlmann, 2010). We considered regularization parameter alpha = "aic," sample fraction = 0.75, scaling = 0.5, number of resamples = 200 in our implementation of randomized LASSO. This was not the alpha parameter in the stability selection article which was scaling (Meinshausen and Bühlmann, 2010). Randomized LASSO was able to select the optimal alpha based on "AIC." The feature scores were scaled between 0 and 1, where 0 was the lowest score (i.e., irrelevant feature) and 1 was the highest score (i.e., most representative or stable feature). Over 200 resampling processes, stability selection provided the overall feature scores (0–1) based on the selection frequency, and a variable was selected. The stability scores were ranked to identify the most important, consistent, stable, and invariant features (i.e., demographic, audiological, and radiological features) over a range of model parameters. We used the ranked features and corresponding class labels in an SVM classifier. Based on the input stable features, the SVM classified patients with CND for different stability thresholds.

Support Vector Machine Classification

Data-driven multivariate analysis is widely used for modeling complex data and understanding relations between the considered variables. Parameter-optimized SVM classifiers can provide robust discriminative models with small sample sizes, being suitable for human neuroimaging studies (Tan et al., 2015;

Feng et al., 2018; Skidmore et al., 2020). The classification performance is greatly affected by the choice of kernel functions, which can map non-linearly separable data onto a linearly separable space. Other tunable parameters, such as the kernel, regularization coefficient C, and γ (γ is an argument having the RBF function as the kernel), also determine the performance. Thus, we used grid search to find the optimal kernel, C, and γ. For kernel functions, we considered the linear function and radial basis function, whereas for C, we considered values from 1 to 10 in increments of 1, and for γ, we considered values of 0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.2, 0.3, 0.4, and 0.5. Both C and γ were evaluated using a radial basis function kernel. We randomly split the data into training and test sets containing 80% and 20% of the available samples, respectively.

During training, we fine-tuned parameters C and γ to find the values that maximally distinguish observations from the CI postoperative CAP and SIR in good and poor recovery groups. The SVM learned the support vectors from the training data containing the attributes (e.g., age in months, residual hearing) and class labels (e.g., spoken language understanding). The resulting hyperplanes were fixed with maximum margin of separation between classes and used to predict unseen test data by providing the unlabeled attributes to the model. The classification performance was evaluated using common measures: accuracy, F1-score, and area under the receiver operating characteristic curve (AUC) (Saito and Rehmsmeier, 2015). The AUC describes the degree to which a model can distinguish between classes. An excellent model has an AUC close to 1, indicating high separability, whereas a poor model has an AUC close to 0, indicating poor separability.

Technology Roadmap

The experimental process is shown in **Figure 4**. We input the CND dataset. First we made a preliminary feature selection. As shown in **Figure 5**, we removed three features according to the variance threshold. As shown in **Figure 6**, we removed two features according to the effect of the features on the labels, that is, the graphical method. As shown in **Figure 7**, we removed one feature according to the correlation. Stability selection was made for the remaining six features. We entered 6 features and a label CAP or SIR into a randomized logistic regression model. We entered 6 features and a label MAIS or MUSS into a randomized LASSO model. We got the stability score of the features on the labels and sorted the stability scores as shown in **Figure 8**. According to the sorted features, the features were added to the SVM model in turn, and the models labeled CAP and SIR were established, respectively. Accuracy and AUC of each model were output, and the best models were selected, respectively, as the prediction models for predicting auditory and speech performance after CI.

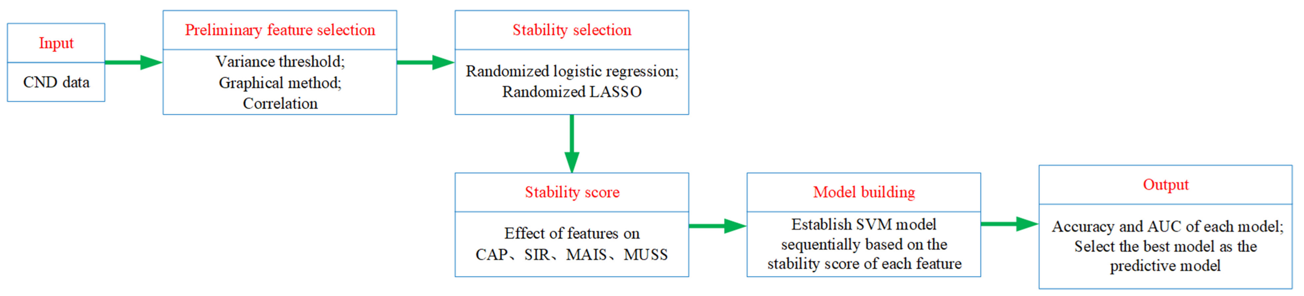

FIGURE 4 | Technology Roadmap.

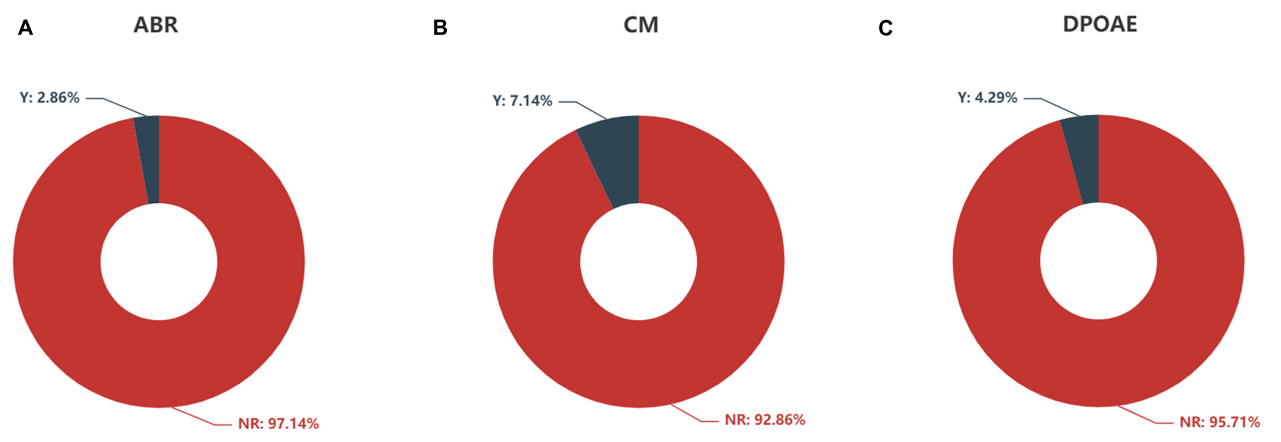

FIGURE 5 | Feature distribution of **(A)** ABR, **(B)** CM, and **(C)** DPOAE according to auditory response. ABR, auditory brainstem response; CM, cochlear microphonics; DPOAE, distortion product otoacoustic emission; Y, response; NR, no response.

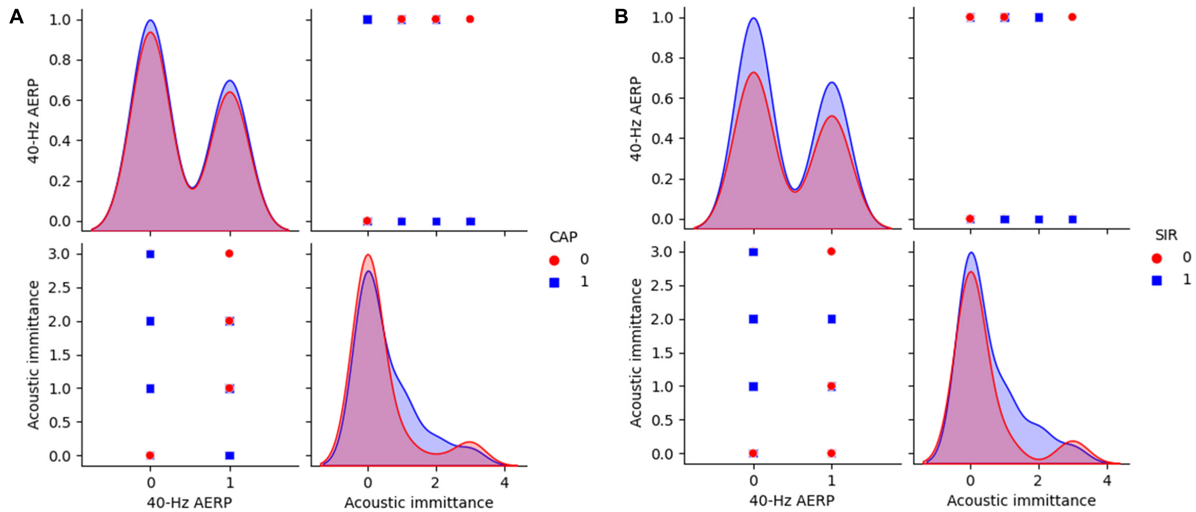

FIGURE 6 | 40-Hz AERP and acoustic immittance for CAP and SIR. **(A)** CAP: 0 indicates no spoken comprehension, and 1 indicates spoken comprehension. **(B)** SIR: 0 indicates unintelligible connected speech, and 1 indicates intelligible connected speech. CAP, categories of auditory performance; SIR, speech intelligibility rating.

RESULTS

Feature Selection

The distributions of ABR, CM and DPOAE are shown in **Figure 5**. For ABR, 68 patients showed no hearing response (NR), and 2 patients had a hearing response (Y), while for CM, 65 patients had NR and 5 had Y, and for DPOAE, 67 patients had NR and 3 had Y. Variance selection removed the ABR, CM, and DPOAE. The influences of the 40-Hz AERP and acoustic immittance on the labels are shown in **Figure 6**. As these two features had less influence on the SIR and CAP labels, they were removed. The correlation coefficient matrix is shown in **Figure 7**. Considering that the correlation coefficient between area ratio of the VCN to the FN and VCN area is 0.63, it has a high correlation. Area ratio of the VCN to the FN is removed.

The six remaining features and the corresponding labels were processed using stability selection to obtain the most representative factors affecting the postoperative hearing and speech rehabilitation of patients with CND and CI. **Figure 8** illustrates the importance of stability selection. Among the factors affecting the postoperative CAP and SIR in patients with CND, VCN area and number of nerve bundles were highly stable and important. Therefore, these features were selected to establish a prediction model of postoperative CAP and SIR in patients with CND. Among the factors affecting the postoperative MAIS in patients with CND, VCN area was the most stable, and the stability scores of the number of nerve bundles, residual hearing, and internal auditory canal diameter were similar. Among the factors affecting the postoperative MUSS in patients with CND, VCN area was the most stable, with a stability score of 1, followed by the number of nerve bundles. Overall, characteristic VCN area and number of nerve bundles were more stable and showed the greatest influence on the postoperative hearing and speech rehabilitation of patients with CND.

Support Vector Machine Classification of Hearing and Speech Rehabilitation Effects Using Vestibulocochlear Nerve Area and Number of Nerve Bundles

We only used VCN area and number of nerve bundles to analyze the effects of postoperative hearing and speech rehabilitation in patients with CND. These features and the corresponding category labels were used to train the SVM. In addition, we applied sevenfold cross-validation and grid search during training to determine the optimal SVM parameters. The optimal parameters for the maximum classification performance listed in **Table 2** were $C = 5$ and $\gamma = 0.02$ for CAP and $C = 6$ and $\gamma = 0.2$ for SIR.

We then selected the best model and performance measures from the predicted class labels, which were obtained from the unseen test data and corresponding ground truth. We applied the SVM classifier using VCN area and number of nerve bundles to identify the effects of hearing and speech rehabilitation. **Table 2** shows that for the hearing rehabilitation effect considering VCN area and number of nerve bundles, the accuracy of spoken language understanding prediction after CI surgery in patients with CND was 71%. For the speech rehabilitation effect considering those two features, the accuracy of intelligible connected speech prediction after CI surgery in patients with CND was 93%. These results indicate suitable prediction of the effects of postoperative hearing and speech rehabilitation after receiving a CI.

The SVM classification results on the test dataset are shown in **Figure 9**. As a correct prediction is shown with a black circle, a model with fewer red circles is preferable, whereas numerous red circles indicate a low generalization ability of the SVM classifier. In fact, each red circle indicates a misclassified patient with CND. **Figure 9A** shows four red circles, indicating four patients with

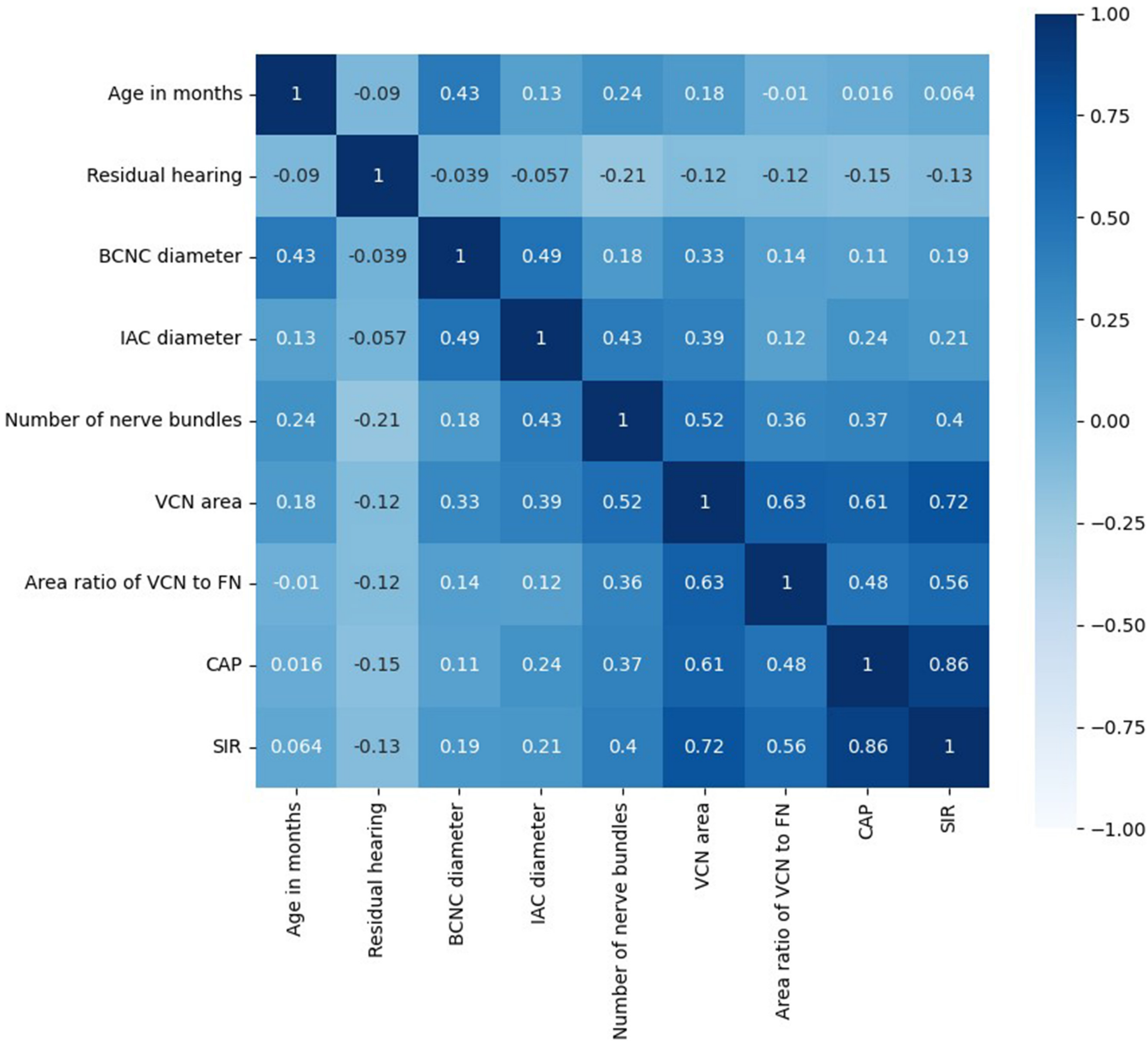

FIGURE 7 | Correlation coefficient matrix.

misclassified CAP and a classification accuracy of 71%. **Figure 9B** shows one red circle, indicating that the SIR of only one patient was misclassified and a classification accuracy of 93%.

Stability Selection for Support Vector Machine Training

We then used stability selection to identify the most representative stable features to separate groups without overfitting. We evaluated stability thresholds yielding different classification performances. The effect of the stability selection threshold on the classification performance is shown in **Figure 10A** for CAP and in **Figure 10B** for SIR. The histogram shows the distribution of feature scores.

The feature scores for stability selection were first determined. As shown in **Figure 8**, for CAP, stability in descending order was obtained for VCN area, number of nerve bundles, internal auditory canal diameter, bony cochlear nerve canal diameter, residual hearing, and age in months; for SIR, this order was obtained for the VCN area, number of nerve bundles, bony cochlear nerve canal diameter, residual hearing, age in months, and internal auditory canal diameter.

The features obtained by stability selection were used in the SVM, whose performance depended on the stability threshold. For CAP and SIR, 66.7% of features scored between 0 and 0.1. Hence, most features were selected less than 10% of the time over 200 model iterations and thus carried near-zero importance for separating groups. Therefore, 66.7% of the features were not related to the grouping of CAP or SIR.

For CAP, the maximum classification performance with 71% accuracy, 71% AUC, and F1-score of 67% was achieved for a stability score threshold of 0.10. For this threshold, two out of

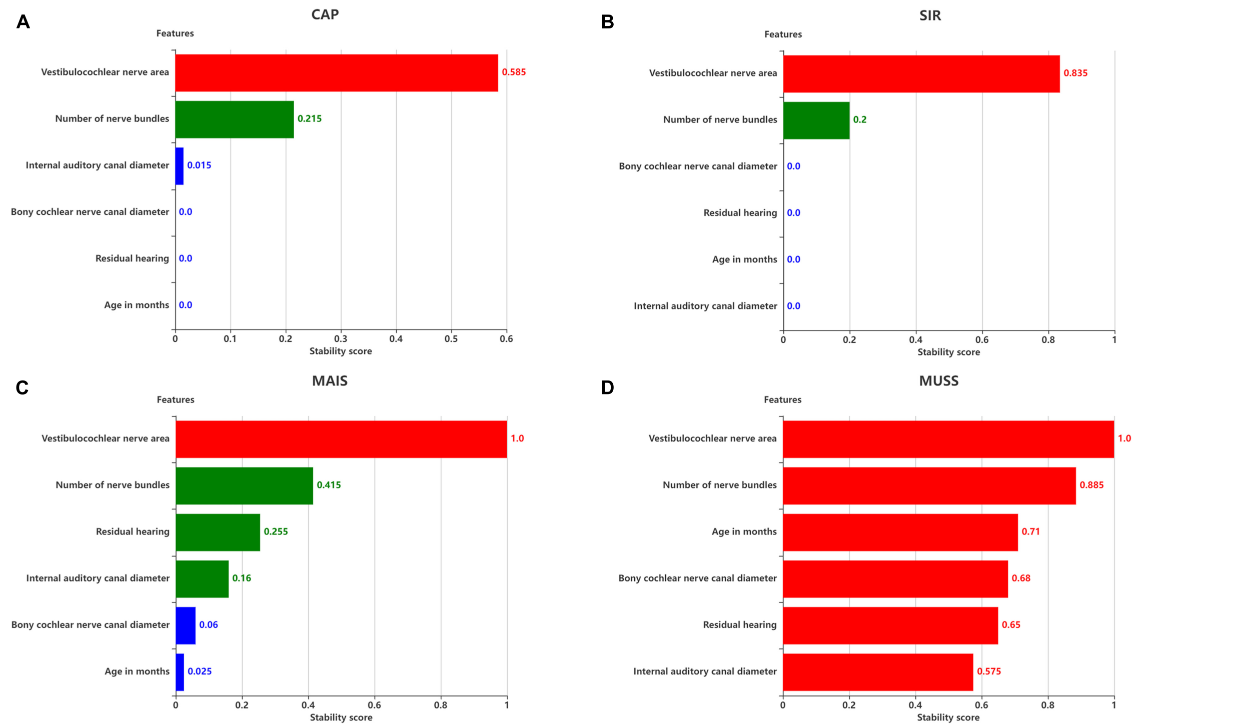

FIGURE 8 | Stability scores for feature selection. Effect of features on **(A)** SIR, **(B)** CAP, **(C)** MAIS, and **(D)** MUSS. The stability score in 0–1 is shown as range per bin.

TABLE 2 | Maximum performance (%) of SVM classifier for distinguishing hearing and speech rehabilitation effects (good and poor).

Measure	CAP	SIR
Accuracy	71	93
AUC	71	94
F1-score	67	93

F1-score = 2 (precision × recall)/(precision + recall).

the six features were selected. For SIR, stability selection provided two out of the six features (33.3%), reaching an accuracy of 93%, AUC of 94%, and F1-score of 93%. Below the optimal threshold of 0.2, the classifier performance reduced owing to the inclusion of unrelated features, while above the threshold, some representative features for distinguishing the effects of hearing and speech rehabilitation were discarded. Even when choosing a stability threshold of 0.5 as a conservative selection, CAP classification reached 64% accuracy with one selected feature, and SIR classification reached 57% accuracy with one selected feature. Thus, predicting the effects of CIs on postoperative hearing and speech rehabilitation in patients with CND may require only a few representative features to notably outperform the random level of classification.

DISCUSSION

The present study included 70 patients diagnosed with CND with normal cochlea according to computed tomography and MRI findings. All subjects underwent unilateral CI. We summarized the age at operation, preoperative audiological findings, and preoperative imaging characteristics of CND patients with normal cochlea and analyzed their correlation with the 2-year postoperative outcome of CI. In our study, the overall mean scores of the CAP, SIR, IT-MAIS/MAIS, and MUSS after 2-year CI activation were 4.10 ± 1.32, 1.87 ± 0.92, 25.14 ± 10.47 and 11.96 ± 10.32, respectively. Among the 70 participants, 36 (51.4%) achieved an understanding of common phrases or the ability to carry on a conversation without lip-reading (CAP 5–7) and 39 (55.7%) patients achieved intelligible speech (SIR > 1). We also obtained a prediction model of the CAP and SIR scores at 2 years after surgery based on these associated factors. We observed that the CAP and SIR scores at 2 years after CI surgery were strongly correlated with the number of nerve bundles and the VCN area.

In our study, the mean age at CI was 27.31 ± 13.92 months (range: 7–54 months). All children failed to pass the newborn hearing screening sequence. Age at CI is known to potentially influence the CI outcomes (Peng et al., 2017). Cochlear implantation provides a unique opportunity to study cortical plasticity associated with long-term deafness and restoration of the auditory modality (Lee et al., 2006). Children who underwent implantation at a younger age have been reported to demonstrate greater gains in speech perception over time than those who underwent implantation at an older age (Zwolan et al., 2004). In our study, we did not find a strong correlation between age at CI and CI performance. This might be due to the limited amount

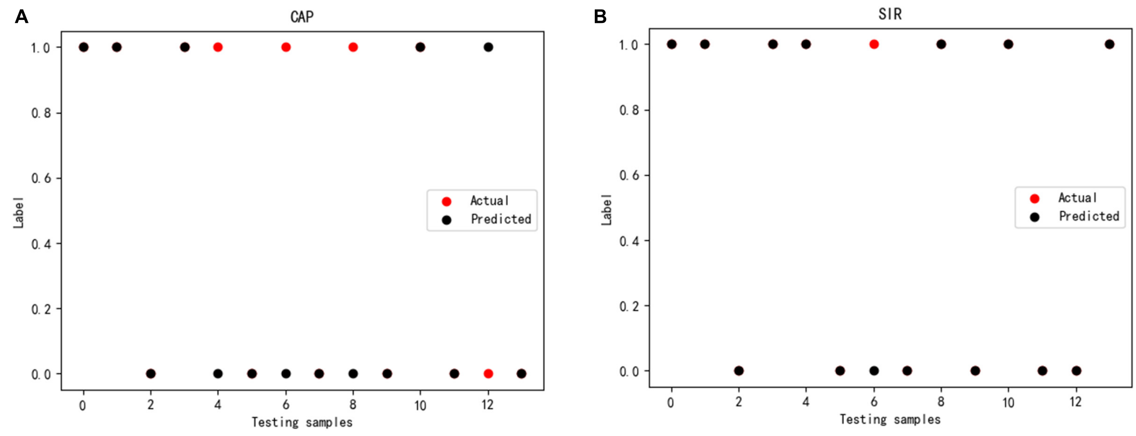

FIGURE 9 | Prediction using SVM classifier for **(A)** CAP and **(B)** SIR.

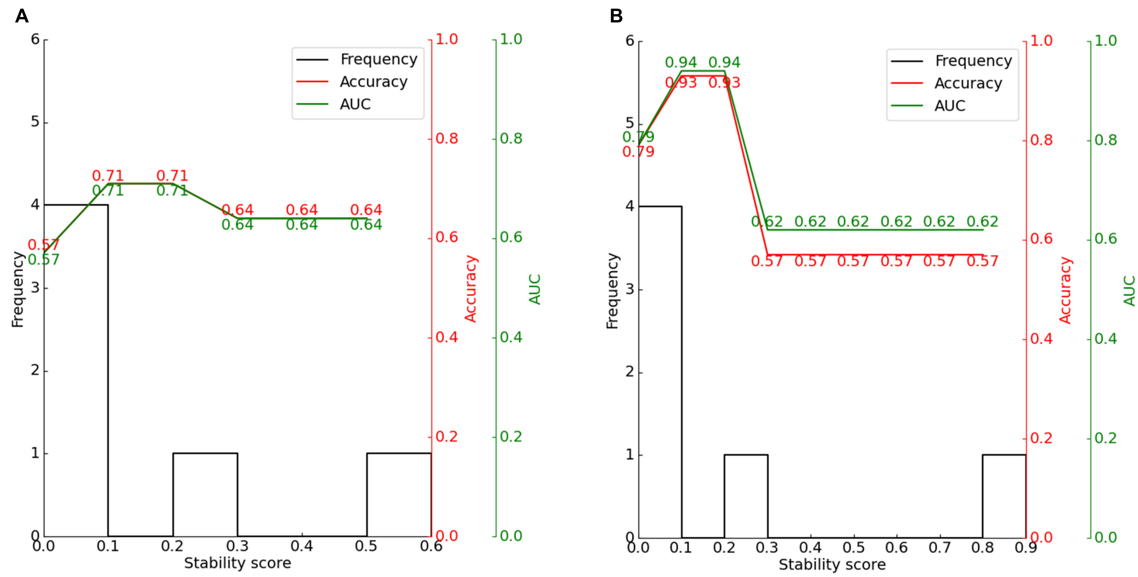

FIGURE 10 | Effect of stability score threshold on model performance for **(A)** CAP and **(B)** SIR. The stability score ranges from 0 to 1.

of CN and auditory stimulation. The results were consistent with those of previous studies (Birman et al., 2016; Han et al., 2019).

The residual hearing threshold is one of the most important prognostic factors correlated with the CI outcomes (Chiossi and Hyppolito, 2017). It represents the number of SGNs and the integrity of neural pathways including the SGNs and the CN. Patients with residual hearing displayed significant improvements in language development (Carlson et al., 2015) after CI surgery. However, we did not find a significant correlation between the average residual hearing threshold and CI performance, which is consistent with the finding in a previous study (Han et al., 2019). This might be attributed to the poor residual hearing of these CND patients. Moreover, we assumed the average hearing threshold of the absence of measurable response in pure-tone audiometry or behavior test to be equal to the maximum output or 5 dB greater than the maximum output of the audiometer for the purpose of evaluation. The calculated mean residual hearing threshold was 108.17 $\pm$ 14.07 dB (range: 81–125 dB), which might have reduced the difference between cases with and without residual hearing.

Auditory brainstem response (ABR) represents the efficacy of hearing aids and cortical development with acoustic stimulation before CI. It was significantly correlated with CI performance in a previous study (Han et al., 2019). However, we did not observe any significant correlation between the ABR and CI outcomes. This finding might be due to the fact that only 2 patients (2/70, 2.86%) exhibited an ABR. Therefore, the sample size was too small to obtain reliable results. Three patients (3/70, 4.27%) exhibited the presence of DPOAE and 5 patients (5/70, 7.14%) exhibited the presence of CM with absent ABR, which suggested a gross discrepancy between the measures of cochlear and neural function in the auditory system diagnosed with auditory neuropathy (AN) (Buchman et al., 2006). The responses of DPOAE as well as CM did not show a statistically significant correlation with CI outcomes, probably due to the small sample size.

Due to the limitations of the current imaging techniques, it is difficult to measure the CN parameters directly. Clinically, the IAC diameter, BCNC diameter, area of the VCN, area ratio of the VCN to the FN, and IAC grade can indirectly determine the condition of the CN. These parameters have been reported to predict the effect of CI after surgery (Shelton et al., 1989; Minami et al., 2015; Birman et al., 2016; Wei et al., 2017; Chung et al., 2018; Han et al., 2019). The maximum diameters of the IAC and the BCNC can indirectly reflect the number of CN fibers and are generally considered to be related to CND (Shelton et al., 1989; Chung et al., 2018). In a previous study (Clemmens et al., 2013), BCNC had a sensitivity of 84% and a specificity of 98% for predicting CND, while IAC had a specificity of 98% and a sensitivity of 44%. In our study, the mean IAC diameter was 2.47 ± 0.85 mm and the mean BCNC diameter was 0.83 ± 0.58 mm. The IAC and BCNC diameters exhibited a weak correlation with CI performance.

Cochlear nerve deficiency (CND) diagnosis mainly relies on MRI (Cerini et al., 2006). Measuring the CN parameters on MRI is the most direct way to determine the condition of the CN. However, due to the limited resolution of the currently used MRI devices, the CN is not clearly visualized. In some cases, although CN fibers are present, they are not reflected in the data regarding CN diameter measurements on MRI and the CN cannot even be visualized. VCN contains all the CN fibers. Hence, some scholars defined CND as VCN deficiency or the absence or thin branches of the VCN (Yamazaki et al., 2015). Measuring the VCN diameter at the CPA indirectly reflects the number of CN fibers. In our study, the VCN area showed a strong correlation with CI performance.

Since direct measurement of the VCN diameter is difficult, some scholars have opted to measure the area ratio of the VCN to the FN. The size of the VCN is generally 1.5–2 times the size of the FN and the size of the CN is similar to that of the FN (Giesemann et al., 2012). Minami et al. (Minami et al., 2015) reported the relationship between the relative sizes of the VCN and FN after CI. They observed that 83% of the patients with IAC stenosis had an FN larger than the VCN. Patients whose FN was larger than the VCN had an average score of 1.1 for auditory behavior after CI, while patients whose FN was smaller than the VCN exhibited an average CAP score of 4.1 after CI. Han et al. (2019) found that the area ratio of the VCN to the FN was significantly correlated with the CAP and IT-MAIS scores at 2 years after CI. In our study, this ratio showed a strong correlation with the CAP and SIR scores, but the correlation was weaker than VCN area. However, due to the presence of thinner FN in some patients, the area ratio of the VCN to the FN was still large despite a thin VCN, which interfered with the accuracy of the results.

Oblique plane sagittal IAC views can show four nerve bundles on MRI: CN, FN, inferior vestibular nerve, and superior vestibular nerve (Govaerts et al., 2003). Since it is difficult to distinguish the CN from other nerves on MRI, Birman et al. (2016) suggested classifying CND according to the number of nerves within the IAC. From oblique plane sagittal IAC views on MRI, IAC nerve grades 0, I, II, and III represent no nerves, one, two, and three nerve bundles, respectively, inside the IAC. These grades correspond to CNA. Grade IV represents four nerves and a thin CN and corresponds to CNH. Previous studies have reported that the IAC nerve grading system was significantly related to the postoperative effect of CI (Birman et al., 2016; Wei et al., 2017; Han et al., 2019). In our study, a higher number of nerve bundles was associated with higher CAP and SIR scores in CND patients. The number of nerve bundles showed strong correlations with the CAP and SIR scores.

In the stability selection analysis for CI outcomes in patients with CND, the VCN area and the number of nerve bundles within the IAC were most representative stable features affecting the CAP and SIR scores at 2 years after CI. For CAP, the accuracy rate was 71% for both the SVM classification and the AUC, whereas for SIR, the accuracy rate for SVM classification and AUC was 93% and 94%, respectively. Only about half of the CND patients in our study were expected to show relative good outcomes (CAP 5-7 or SIR > 1). Our models can help surgeons select the appropriate side for CI and at the same time, provide reasonable expectations regarding the effects of CI surgery. For patients who show inadequate benefit following CI, auditory brainstem implantation should be considered despite the risk of serious complications such as cerebrospinal fluid leakage, meningitis, intracranial bleeding, stroke, cranial nerve damage, and even death (Freeman and Sennaroglu, 2018). In our model, we included only the CND patients with normal cochlea and patients with cochlear malformations and other systemic complications were excluded. Therefore, the application of this model requires complete audiological and imaging evaluations before surgery.

CONCLUSION

CI in CND with normal cochlea is associated with variable outcomes. We observed that postoperative CAP and SIR scores of CND patients showed a strong correlation with the VCN area and the number of nerve bundles within the IAC. However, age at implantation and residual hearing did not show any strong correlation. Results from our study can help surgeons select the appropriate side for CI and provide reasonable expectations regarding the outcomes of CI surgery.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethics Board of Tongren Hospital, Beijing, China. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

SLu designed the experiments and collected the data. JX analyzed the data. SLu and JX drafted the manuscript. All authors edited and revised the manuscript and approved the final version of manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (grant no. 81670923) and Beijing Municipal Natural Science Foundation (grant no. 7212015).

ACKNOWLEDGMENTS

We would like to thank Wiley Editing Services for English language editing of this manuscript.

REFERENCES

Adunka, O. F., Jewells, V., and Buchman, C. A. (2007). Value of computed tomography in the evaluation of children with cochlear nerve deficiency. *Otol. Neurotol.* 28, 597–604. doi: 10.1097/01.mao.0000281804.36574.72

Birman, C. S., Powell, H. R. F., Gibson, W. P. R., and Elliott, E. J. (2016). Cochlear implant outcomes in cochlea nerve aplasia and hypoplasia. *Otol. Neurotol.* 37, 438–445. doi: 10.1097/MAO.0000000000000997

Boulogeorgos, A. A. A., Trevlakis, S. E., and Chatzidiamantis, N. D. (2021). Optical wireless communications for in-body and transdermal biomedical applications. *IEEE Commun. Mag.* 59, 119–125.

Buchman, C. A., Roush, P. A., Teagle, H. F., Brown, C. J., Zdanski, C. J., and Grose, J. H. (2006). Auditory neuropathy characteristics in children with cochlear nerve deficiency. *Ear Hear.* 27, 399–408. doi: 10.1097/01.aud.0000224100.30525.ab

Carlson, M. L., Sladen, D. P., Haynes, D. S., Driscoll, C. L., DeJong, M. D., Erickson, H. C., et al. (2015). Evidence for the expansion of pediatric cochlear implant candidacy. *Otol. Neurotol.* 36, 43–50.

Cerini, R., Faccioli, N., Cicconi, D., Schenal, G., Cugini, C., Giarbini, N., et al. (2006). Role of CT and MRI in the preoperative evaluation of auditory brainstem implantation in patients with congenital inner ear pathology. *Radiol. Med.* 111, 978–988. doi: 10.1007/s11547-006-0096-2

Chiossi, J. S. C., and Hyppolito, M. A. (2017). Effects of residual hearing on cochlear implant outcomes in children: a systematic-review. *Int. J. Pediatr. Otorhinolaryngol.* 100, 119–127. doi: 10.1016/j.ijporl.2017.06.036

Chung, J., Jang, J. H., Chang, S. O., Song, J. J., Cho, S. W., Kim, S. Y., et al. (2018). Does the width of the bony cochlear nerve canal predict the outcomes of cochlear implantation? *Biomed Res. Int.* 2018:5675848. doi: 10.1155/2018/5675848

Clemmens, C. S., Guidi, J., Caroff, A., Cohn, S. J., Brant, J. A., Laury, A. M., et al. (2013). Unilateral cochlear nerve deficiency in children. *Otolaryngol. Head Neck Surg.* 149, 318–325.

Colletti, V., Fiorino, F. G., Carner, M., Miorelli, V., Guida, M., and Colletti, L. (2004). Auditory brainstem implant as a salvage treatment after unsuccessful cochlear implantation. *Otol. Neurotol.* 25, 485–496. doi: 10.1097/00129492-200407000-00016

Crowson, M. G., Lin, V., Chen, J. M., and Chan, T. C. (2020). Machine learning and cochlear implantation—a structured review of opportunities and challenges. *Otol. Neurotol.* 41, e36–e45. doi: 10.1097/MAO.0000000000002440

Feng, G., Ingvalson, E. M., Grieco-Calub, T. M., Roberts, M. Y., Ryan, M. E., Birmingham, P., et al. (2018). Neural preservation underlies speech improvement from auditory deprivation in young cochlear implant recipients. *Proc. Natl. Acad. Sci. U.S.A.* 115, E1022–E1031. doi: 10.1073/pnas.1717603115

Freeman, S. R., and Sennaroglu, L. (2018). Management of cochlear nerve hypoplasia and aplasia. *Adv. Otorhinolaryngol.* 81, 81–92. doi: 10.1159/000485542

Giesemann, A. M., Kontorinis, G., Jan, Z., Lenarz, T., Lanfermann, H., and Goetz, F. (2012). The vestibulocochlear nerve: aplasia and hypoplasia in combination with inner ear malformations. *Eur. Radiol.* 22, 519–524. doi: 10.1007/s00330-011-2287-z

Govaerts, P. J., Casselman, J., Daemers, K., De Beukelaer, C., Yperman, M., and De Ceulaer, G. (2003). Cochlear implants in aplasia and hypoplasia of the cochleovestibular nerve. *Otol. Neurotol.* 24, 887–891. doi: 10.1097/00129492-200311000-00011

Han, J. J., Suh, M. W., Park, M. K., Koo, J. W., Lee, J. H., and Oh, S. H. (2019). A predictive model for cochlear implant outcome in children with cochlear nerve deficiency. *Sci. Rep.* 9:1154. doi: 10.1038/s41598-018-37014-7

Jallu, A. S., Jehangir, M., Ul Hamid, W., and Pampori, R. A. (2015). Imaging evaluation of pediatric sensorineural hearing loss in potential candidates for cochlear implantation. *Ind. J. Otolaryngol. Head Neck Surg.* 67, 341–346. doi: 10.1007/s12070-015-0819-6

Kang, W. S., Lee, J. H., Lee, H. N., and Lee, K. S. (2010). Cochlear implantations in young children with cochlear nerve deficiency diagnosed by MRI. *Otolaryngol. Head Neck Surg.* 143, 101–108. doi: 10.1016/j.otohns.2010.03.016

Lee, H.-J., Giraud, A.-L., Kang, E., Oh, S.-H., Kang, H., Kim, C.-S., et al. (2006). Cortical activity at rest predicts cochlear implantation outcome. *Cereb. Cortex* 17, 909–917. doi: 10.1093/cercor/bhl001

Mahmud, M. S., Ahmed, F., Al-Fahad, R., Moinuddin, K. A., Yeasin, M., Alain, C., et al. (2020). Decoding hearing-related changes in older adults' spatiotemporal neural processing of speech using machine learning. *Front. Neurosci.* 14:748. doi: 10.3389/fnins.2020.00748

Meinshausen, N., and Bühlmann, P. (2010). Stability selection. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 72, 417–473. doi: 10.1111/j.1467-9868.2010.00740.x

Minami, S. B., Takegoshi, H., Shinjo, Y., Enomoto, C., and Kaga, K. (2015). Usefulness of measuring electrically evoked auditory brainstem responses in children with inner ear malformations during cochlear implantation. *Acta Otolaryngol.* 135, 1007–1015. doi: 10.3109/00016489.2015.1048377

Nogueira, S., Sechidis, K., and Brown, G. (2017). On the stability of feature selection algorithms. *J. Mach. Learn. Res.* 18, 6345–6398.

Olusanya, B. O., Davis, A. C., and Hoffman, H. J. (2019). Hearing loss grades and the International classification of functioning, disability and health. *Bull. World Health Organ.* 97, 725–728. doi: 10.2471/BLT.19.230367

Peng, K. A., Kuan, E. C., Hagan, S., Wilkinson, E. P., and Miller, M. E. (2017). Cochlear nerve aplasia and hypoplasia: predictors of cochlear implant success. *Otolaryngol. Head Neck Surg.* 157, 392–400. doi: 10.1177/0194599817718798

Saito, T., and Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One* 10:e0118432. doi: 10.1371/journal.pone.0118432

Sennaroğlu, L., and Bajin, M. D. (2017). Classification and current management of inner ear malformations. *Balkan Med. J.* 34, 397–411. doi: 10.4274/balkanmedj.2017.0367

Shelton, C., Luxford, W. M., Tonokawa, L. L., Lo, W. W. M., and House, W. F. (1989). The narrow internal auditory canal in children: a contraindication to cochlear implants. *Otolaryngol. Head Neck Surg.* 100, 227–231. doi: 10.1177/019459988910000310

Shi, Y., Li, Y., Gong, Y., Chen, B., and Chen, J. (2019). Cochlear implants for patients with inner ear malformation: experience in a cohort of 877 surgeries. *Clin. Otolaryngol.* 44, 702–706. doi: 10.1111/coa.13360

Skidmore, J., Xu, L., Chao, X., Riggs, W. J., Pellittieri, A., Vaughan, C., et al. (2020). Prediction of the functional status of the cochlear nerve in individual cochlear implant users using machine learning and electrophysiological measures. *Ear Hear.* 42, 180. doi: 10.1097/AUD.0000000000000916

Tan, L., Holland, S. K., Deshpande, A. K., Chen, Y., Choo, D. I., and Lu, L. J. (2015). A semi-supervised support vector machine model for predicting the language outcomes following cochlear implantation based on pre-implant brain fMRI imaging. *Brain Behav.* 5:e00391. doi: 10.1002/brb3.391

Trevlakis, S. E., Boulogeorgos, A. A., Chatzidiamantis, N. D., and Karagiannidis, G. K. (2020). All-Optical Cochlear Implants. *IEEE Trans. Mol. Biol. Multi-Scale Commun.* 6, 13–24.

Trevlakis, S. E., Boulogeorgos, A.-A., Sofotasios, P. C., Muhaidat, S., and Karagiannidis, G. K. (2019). Optical wireless cochlear implants. *Biomed. Opt. Express* 10, 707–730. doi: 10.1364/BOE.10.000707

Velde, H., Rademaker, M., Damen, J., Smit, A., and Stegeman, I. (2021). Prediction models for clinical outcome after cochlear implantation: a systematic review. *J. Clin. Epidemiol.* 137, 182–194. doi: 10.1016/j.jclinepi.2021.04.005

Wei, X., Li, Y., Chen, B., Gong, Y., Fu, Q. J., Liu, T., et al. (2017). Predicting auditory outcomes from radiological imaging in cochlear implant patients with cochlear nerve deficiency. *Otol. Neurotol.* 38, 685–693. doi: 10.1097/MAO.0000000000001382

Wu, C. M., Lee, L. A., Chen, C. K., Chan, K. C., Tsou, Y. T., and Ng, S. H. (2015). Impact of cochlear nerve deficiency determined using 3-dimensional magnetic resonance imaging on hearing outcome in children with cochlear implants. *Otol. Neurotol.* 36, 14–21. doi: 10.1097/MAO.0000000000000568

Yamazaki, H., Leigh, J., Briggs, R., and Naito, Y. (2015). Usefulness of MRI and EABR Testing for Predicting CI Outcomes Immediately After Cochlear Implantation in Cases With Cochlear Nerve Deficiency. *Otol. Neurotol.* 36, 977–984. doi: 10.1097/MAO.0000000000000721

Yousef, M., Mesallam, T. A., Garadat, S. N., Almasaad, A., Alzhrani, F., Alsanosi, A., et al. (2021). Audiologic outcome of cochlear implantation in children with cochlear nerve deficiency. *Otol. Neurotol.* 42, 38–46. doi: 10.1097/MAO.0000000000002849

Zwolan, T. A., Ashbaugh, C. M., Alarfaj, A., Kileny, P. R., Arts, H. A., El-Kashlan, H. K., et al. (2004). Pediatric cochlear implant patient performance as a function of age at implantation. *Otol. Neurotol.* 25, 112–120. doi: 10.1097/00129492-200403000-00006

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Neuroscience

ORIGINAL RESEARCH
published: 24 June 2022
doi: 10.3389/fnins.2022.885263

Check for updates

The Effectiveness of Targeted Electrical Stimulation *via* Cochlear Implant on Tinnitus-Perceived Loudness

Walter Di Nardo[1], Tiziana Di Cesare[1]*, Angelo Tizio[1], Gaetano Paludetti[1] and Anna Rita Fetoni[2]

[1] Dipartimento di Scienze dell'Invecchiamento, Neurologiche, Ortopediche e della Testa-Collo, UOC di Otorinolaringoiatria, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy, [2] Dipartimento di Neuroscienze, Sezione di Audiologia, Universitá Federico II, Naples, Italy

Introduction: The cause of tinnitus improvement in cochlear implant (CI) users is not understood. On the basis that a spatially limited dysfunction in the auditory pathway could cause tinnitus, we used single-channel stimulation to evaluate any variation of tinnitus-perceived loudness and identify the cochlear regions involved.

Materials and Methods: It was an observational prospective case-crossover study. After the first mapping, 21 adults with unilateral CI and chronic tinnitus expressed their tinnitus loudness based on the Visual Analog Scale (VAS) score (0–10) at baseline (L^0), during a 10 s single-channel stimulation with C-level of electric current (L^1) and 30 min after CI activation (L^2). Tinnitus reduction [$R^T = (L^0 - L^1) \times 100/L^0$] > 50% was considered significant. VAS outcomes were compared between baseline (L0) and (each) single-channel stimulation (L1) to find the channel with the greatest R^T (suppressive channel-SC), whose frequency range revealed the cochlear region involved. Seven patients with asymmetric hearing loss underwent the pitch-matching test to identify the actual frequency evoked by the SC. We compared selective (L^1) and non-selective (L^2) intracochlear stimulation using paired *t*-test. Preoperative Tinnitus Handicap Inventory (THI) score was compared with those at 1, 6, and 12 months with paired *t*-tests to evaluate long-term tinnitus perception.

Results: We observed a significant reduction of tinnitus loudness during the experimental procedure [L^0 (6.4 ± 2.4) vs. L^1 (1.7 ± 2.7), $p = 0.003$]. A total of 15/21 patients (71.4%) had a significant ($R^T > 50\%$) and selective improvement, reporting a mean L^1 of 0.4 ± 2.0 ($p = 0.0001$). In 10/15 (66.6%) patients, the SC was in the apical turn, within 1,000 Hz; in 5/15 patients (33.4%) within 4,000 Hz. The cochlear region 125–313 Hz was the most affected by tinnitus improvement ($p = 0.0074$). Targeted stimulation was more effective than non-selective stimulation [L^1 vs. L^2 (4.3 ± 2.5), $p = 0.0022$]. In 3/7 patients, the perceived pitch did not fall within the SC frequency

OPEN ACCESS

Edited by:
*Fei Chen,
Southern University of Science and Technology, China*

Reviewed by:
*Agnieszka J. Szczepek,
University Medical Center Utrecht, Netherlands*

****Correspondence:***
*Tiziana Di Cesare
tizianadicesare90@gmail.com
orcid.org/0000-0001-9756-1880*

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *27 February 2022*
Accepted: *23 May 2022*
Published: *24 June 2022*

Citation:
Di Nardo W, Di Cesare T, Tizio A, Paludetti G and Fetoni AR (2022) The Effectiveness of Targeted Electrical Stimulation via Cochlear Implant on Tinnitus-Perceived Loudness. Front. Neurosci. 16:885263. doi: 10.3389/fnins.2022.885263

ranges. All patients with selective attenuation described tinnitus as monotone. Patients with non-selective attenuation had polyphonic tinnitus and better THI results after 1 year.

Conclusion: Targeted intracochlear electrical stimulation improved chronic tinnitus perception, especially in monotone tinnitus, and the apical region was mainly involved. Our results provide new insights into the pathophysiological mechanisms of tinnitus and targets for innovative therapeutic strategies.

Keywords: cochlear implant, tinnitus, intracochlear electrical stimulation, cochlear regions, pitch match

INTRODUCTION

Subjective tinnitus, consisting of the perception of sounds without a corresponding acoustic stimulus, is a very common and disabling condition with severe effects on health and wellbeing, imposes a substantive economic burden, and has no known cure (Lockwood et al., 2002). Up to 50% of adults report to have experienced transient tinnitus following noise exposure, while 5–15% of people living in industrialized societies suffer from chronic tinnitus with negative effects on their quality of life (Gallus et al., 2015; Bhatt et al., 2016). Hearing loss is a common cause of tinnitus and is experienced by up to 86% of adult cochlear implant (CI) candidates (Quaranta et al., 2004).

Knowledge of the pathophysiological mechanisms that trigger and maintain chronic tinnitus is one of the major challenges of tinnitus research whose ultimate purpose is to find a cure (Haider et al., 2018). Peripheral lesions, including loss of hair cells, dysregulated endocochlear potential, and cochlear spontaneous overactivity, could explain the temporary tinnitus occurring immediately after an acute noise trauma (Norena and Eggermont, 2003). In contrast, the finding that bilateral auditory nerve sectioning does not always eliminate tinnitus (Pulec, 1984) suggests that a peripheral lesion is not sufficient to maintain tinnitus and rather represents the trigger of a cascade of neuroplastic changes involving retro-cochlear auditory structures (Henry et al., 2014; Fetoni et al., 2015). Neural reorganization can occur within multiple levels of the central auditory pathway, including the dorsal cochlear nucleus, ventral cochlear nucleus, and inferior colliculus of the brainstem, whose hyperactivity resulting from the downregulation of inhibitory signals has been extensively studied in tinnitus models (Mulders and Robertson, 2009; Zeng et al., 2009). In addition, the medial geniculate body of the thalamus, a major gate of sensory signals to the cortex, the limbic system (Rauschecker et al., 2010), and the primary auditory cortex itself were deemed centers of tinnitus due to the reorganization of the tonotopic map (Eggermont and Roberts, 2004; Knipper et al., 2013; Noreña, 2015). Increasing evidence shows that auditory deprivation leads to chronic subjective tinnitus caused by the overrepresentation of adjacent cortical areas with similar characteristic frequency, known as "edge frequencies," due to the lack of lateral inhibition phenomena (Eggermont and Roberts, 2004). Thus, maladaptive plastic changes in the central auditory pathways may be involved in maintaining tinnitus in a sort of "vicious circle" (Fetoni et al., 2015). Therapeutic implications are significant since it was first supposed that only peripheral tinnitus could be masked by sounds as opposed to tinnitus powered by central generators (Haider et al., 2018). Hence, the pathogenesis is still unclear and multiple mechanisms at various levels of the auditory system are likely to concur.

To date, only a few tinnitus treatments are available, but there is no pharmacological approach approved by the major drug agencies. Electrical stimulation delivered both transcutaneously (Steenerson and Cronin, 2003; Aydemir et al., 2006) and transtympanically (Konopka et al., 2001; Rubinstein et al., 2003; Di Nardo et al., 2009) has been proposed as a promising approach to suppress peripheral tinnitus. It has been suggested that electric current could act both at a presynaptic level, with the reduction in the spontaneous release of neurotransmitters from inner hair cells (Konishi et al., 1970), and with a postsynaptic mechanism by reducing the opening of voltage-gated sodium channels, with a direct effect on the membrane potential of the cochlear fibers (Shepherd and Javel, 1999). On this basis, central tinnitus, which becomes independent of cochlear residual spontaneous activity, should not respond to this treatment (Noreña et al., 2015). Interestingly, in the experimental model, brain stimulation induced by the anodal transcranial direct current affects the structural plasticity of the auditory cortex and compensates for the effects of sensory deprivation following cochlear damage by increasing dendritic spine numbers and rearranging synaptic networks (Paciello et al., 2018) in the primary auditory cortex. Therefore, attempts have been made to relieve chronic intractable tinnitus by delivering different electrical stimuli directly to the auditory cortex (De Ridder et al., 2006; Seidman et al., 2008) with doubtful results and many possible side effects related to the invasiveness of the method.

Nowadays, the different methods used and the uncertainty of clinical efficacy have made it impossible to guide technological development toward a real therapy (Assouly et al., 2021). Considering this scenario, the intracochlear electrical stimulation *via* CI could represent a putative approach to tinnitus treatment. Moreover, the evaluation of the effects of electrical stimulation should improve the general knowledge of tinnitus mechanisms. Beneficial effects on tinnitus have been previously reported in many CI users (Quaranta et al., 2004; Di Nardo et al., 2007b; Ramakers et al., 2015; Peter et al., 2019), but the cause is still not understood: experimental studies on animals suggest that CI can restore a certain degree of normal discharge in the cochlear nerve through the inhibition of spontaneous activity or even a reflex increase in microcirculation in the auditory pathway (Runge-Samuelson et al., 2004). In contrast, the long-term stable effects of the CI on tinnitus would require the reorganization of the central auditory cortex. The masking theory, which is secondary to hearing relief, is unlikely since most patients report persistent improvement of their tinnitus even at night when CI is off (Quaranta et al., 2004). It has been proposed that auditory stimulation could reverse the tinnitus-related central changes, but the presence of degenerated cochlear fibers hinders the process of restoring the pre-hearing loss distribution of sensory inputs to the auditory centers (Noreña, 2015). In addition, recent attempts to optimize intracochlear electrical stimulation to reduce tinnitus have not led to clear results (Arts et al., 2015, 2016), suggesting that the characteristics of the stimulus for tinnitus reduction are highly subject-specific. Furthermore, the most effective target of intracochlear electrical stimulation was not investigated by previous studies.

Assuming that the stimulation of a limited area of the cochlea involved in triggering and maintaining chronic tinnitus could improve tinnitus in CI users, as opposed to the non-specific stimulation of the entire cochlea, we aimed to further investigate the intracochlear electrical stimulation and identify the electrical channel(s) of the CI array that suppress/attenuate tinnitus.

Objectives

The major aim of our study was to measure in each patient the variation of the subjectively perceived tinnitus loudness during short-term single-channel stimulation and to find the best-performer channel (i.e., the channel that caused the greatest tinnitus reduction). Furthermore, the correlation between the reduction of tinnitus loudness and the position of the channels inside the cochlea has been evaluated in all patients to define the cochlear regions involved in tinnitus improvement.

We also aimed (i) to compare the impact of short-term single-channel electrical stimulation with the normal functioning of the whole CI in the early phase of activation in terms of tinnitus improvement, (ii) study long-term results of CI on the subjective perception of tinnitus, and (iii) evaluate whether qualitative characteristics of tinnitus (monotone vs. polyphonic tinnitus) could influence the results of short- and long-term intracochlear electrical stimulation on tinnitus loudness.

MATERIALS AND METHODS

Subjects

It was an observational prospective case-crossover study. Subjects were enrolled from January 2021, and the study lasted for 1 year. Before enrollment, all patients received complete and comprehensible information regarding the tests administered and gave their written consent, in agreement with the ethical standards of the Declaration of Helsinki. The study was approved by our institution's ethics committee under protocol no. 0023756/21.

All candidates for unilateral cochlear implantation surgery at our institution preoperatively underwent an accurate medical history focused on duration, type and cause of hearing loss, and onset and characteristics of tinnitus. A complete audiological evaluation was performed, including otoscopy, tympanometry, and acoustic reflex measurement (Grason Stadler Tympstar), as well as standard pure-tone audiometry, testing conventional frequency ranging from 0.25 to 8 kHz (Amplaid 319 audiometer, Amplaid Inc.) in a double-walled, soundproof room. Preoperative pure tone average (PTA) (average of hearing threshold levels at 500, 1,000, 2,000, and 4,000 Hz) was measured on both ears in all patients. All patients had bilateral sensorineural HL, which was severe to profound (PTA > 70 dB HL) in the worst ear and slight to profound in the other ear (**Table 1**). Inclusion criteria were as follows: age $\geq$ 18 years; chronic (at least 6 months) tinnitus perceived in the worst hearing-impaired ear; intracochlear placement of the implant through the (extended) round window membrane; and the ability to read, understand, and fill in the assigned questionnaires and sign an informed consent form.

Patients with tinnitus onset after surgery were not enrolled. Other exclusion criteria were pulsatile tinnitus, congenital malformation of the auditory system detected with preoperative CT and MR imaging of the inner ear and brain, history of vestibular schwannoma, active middle ear disease, and complications during or after surgery (i.e., flap necrosis, improper electrode placement, facial nerve problems, infection, facial nerve stimulation, vertigo). Cases with incomplete/difficult insertion of the array into the cochlea were also excluded. The insertion of the CI in the cochlea was demonstrated in all patients with intraoperative X-ray static fluoroscopy (Garaycochea et al., 2020) to avoid possible extracochlear array misplacement (e.g., semicircular canal, vestibule, middle ear), tip rollover, kinking, or lopping. All patients underwent the intraoperative electrophysiological test to verify the neural response from all the CI's electrical channels.

Patients with a history of psychiatric disorders, depression, and use of antidepressant treatments, as well as patients affected with neurodegenerative diseases, especially Alzheimer's and Parkinson's diseases, were also excluded from the study.

Study Design

Study Questionnaires

All patients underwent the assessment of tinnitus characteristics by using self-administered questionnaires as follows:

- *The Tinnitus Characteristics Questionnaire for CI recipients* used by Wang et al. (2017) was translated into Italian and administered immediately before surgery to define the qualitative characteristics of hearing loss and tinnitus (cause, laterality, grading, duration), typology (subjective, objective), year of onset, localization (bilateral, unilateral right or left, central), components (monotone or polyphonic, intermittent, or continuous), subjectively defined type of tinnitus (cicadas, roar, crackle, rain, wind, hum, whistle, music), and aspect of greatest influence on daily life (sleep, hearing, emotion, work, memory).

- *The Tinnitus Handicap Inventory (THI)* in its validated Italian version (Monzani et al., 2008) was administered before surgery, 1 month (immediately before CI activation), 6 months, and 1 year after surgery to study long-term CI effect on tinnitus. THI was administered according

TABLE 1 | Patients' demographic characteristics, causes, and duration of hearing loss (HL) and preoperative audiometric data.

Patient	Sex	Age (yrs)	Hearing Loss (HL) cause	HL, duration (yrs)	PTA R (db)	PTA L (db)
N1	Female	53	Ménière's disease	16	45	114
N2	Female	65	Otosclerosis	43	68	120
N3	Male	68	Idiopathic	18	120	99
N4	Male	62	Idiopathic	53	84	103
N5	Female	22	Sudden HL	10	102	99
N6	Male	58	Idiopathic	58	115	115
N7	Female	68	Otosclerosis	50	120	114
N8	Female	58	Cogan syndrome	24	63	120
N9	Male	53	Otosclerosis	30	102	102
N10	Female	56	Idiopathic	13	120	120
N11	Male	20	CMV	16	120	120
N12	Female	59	Iatrogenic (aminoglycosides)	46	93	101
N13	Male	55	Auditory neuropathy	13	67	77
N14	Female	46	Hereditary genetics	10	62	72
N15	Male	64	Otosclerosis	19	68	91
N16	Female	47	Idiopathic	20	120	106
N17	Female	38	Idiopathic	14	76	76
N18	Female	63	Otosclerosis	30	91	112
N19	Female	49	Sudden HL	30	84	93
N20	Female	80	Otosclerosis	30	107	120
N21	Male	40	Otosclerosis	10	99	75

PTA, pure tone average (average of hearing threshold levels at 500, 1,000, 2,000, and 4,000 Hz); R, right ear; L, left ear.

TABLE 2 | Detection of cochlear regions most involved in the attenuation/suppression of tinnitus: electrical channel with the greatest attenuation capacity on tinnitus in each patient [suppressive channel (SC)]; minimum current level suppressing tinnitus (MCLT) (μA).

Patient	SC	Frequency	MCLT (μA)	CI model	Electrodelength	Strategy of stimulation	Frequency of stimulation	Pulsewidth
N1	22	125–313	215,4	Cochlear CI612Peri-modiolar	19mm	MP3000	900	25
N2	7	3063–3563	244,4	Cochlear CI512Peri-modiolar	19mm	MP3000	900	25
N3	Nf	/	/	Cochlear CI632Peri-modiolar	17mm	ACE	900	25
N4	Nf	/	/	MED-EL Mi1200Flex28Lateral-wall	28mm	FS4	1237	25.42–30.42
N5	17	563–688	735,5	Cochlear CI512Peri-modiolar	19mm	ACE	900	25
N6	4	491–710	368	MED-EL Mi1200Flex28Lateral-wall	28mm	FS4	1660	12.08–17.92
N7	11	1813–2063	235,7	Cochlear CI512Peri-modiolar	19mm	ACE	900	25
N8	22	125–313	344,5	Cochlear CI24RE CAPeri-modiolar	19mm	ACE	900	25
N9	22	125–313	613,9	Cochlear CI512Peri-modiolar	19mm	ACE	900	25
N10	6	3563–4063	348,9	Cochlear CI422Lateral-wall	25mm	ACE	900	25
N11	22	125–313	204	Cochlear CI512Peri-modiolar	19mm	ACE	900	25
N12	22	125–313	443,5	Cochlear CI612Peri-modiolar	19mm	MP3000	900	25
N13	11	1813–2063	215,4	Cochlear CI512Peri-modiolar	19mm	ACE	900	25
N14	Nf	/	/	AB Hires 90KJLateral-wall	20mm	HRes Optima-S	1258	53
N15	Nf	/	/	Cochlear CI422Lateral-wall	25mm	ACE	900	25
N16	22	125–313	560,9	Cochlear CI24RE CAPeri-modiolar	19mm	ACE	900	25
N17	1	100–198	397	MED-EL Mi1200Flex28Lateral-wall	28mm	FS4	1277	17.08–33.75
N18	6	3563–4063	485,5	Cochlear CI612Peri-modiolar	19mm	ACE	900	25
N19	19	563–688	309	Cochlear CI612Peri-modiolar	19mm	ACE	900	25
N20	Nf	/	/	Cochlear CI512Peri-modiolar	19mm	ACE	900	25
N21	Nf	/	/	Cochlear CI512Peri-modiolar	19mm	ACE	900	25

CI, cochlear implant model; nf, not found.

to the model proposed by Newman (Newman et al., 1996) and graded according to the McCombe grading system (McCombe et al., 2001). The THI questionnaire is composed of 25 questions, each with three quantifiable answers (yes = 4, sometimes = 2, no = 0). The final total score (0–100) defines the degree of subjective perception of tinnitus in the last week: grade 1, very slight (THI score 0–16); grade 2, mild (THI 18–36); grade 3, moderate (THI 38–56); grade 4, severe (THI 58–76); grade 5, catastrophic (THI 78–100).

- *Visual Analog Scale (VAS),* which is used in clinical research to measure the intensity of symptoms, was administered to patients to assess perceived tinnitus loudness. Patients were asked to rate the perceived loudness of their tinnitus on a horizontal scale oriented from right to left, from 0 (inaudible) to 10 (loud like never before). VAS was administered before surgery (L^S); 4 weeks after surgery (immediately before CI activation) to define the baseline loudness of tinnitus (L^0); on the day of activation during the experimental procedure *via* the single electrical channel stimulation (L^1); and on the day of activation, 30 min after the whole CI was first turned on (L^2). This scale expresses the subjective perception of tinnitus, namely, absent (0–1), mild (2–3), moderate (3–6), severe (6–8), and very serious (9–10).

Study Procedures

Enrolled patients underwent the following procedures:

- *CI activation and mapping*: CI was activated 4 weeks after surgery. All channels were sequentially activated. The maximum comfort level (C-level) and mean threshold level (T-level) were determined based on subjective responses. The full CI frequency range was distributed to the different electrical channels in accordance with the CI manufactures' standards that are based on the Greenwood's function (Greenwood, 1961).

- *Experimental procedure – single electrical channel stimulation*: The short-term effect of electric current on tinnitus was evaluated during the stimulation of the different electrical channels with the C-level, for 10 s, one by one, starting from the cochlear apex toward the base, with a recovery time of 30 s between one channel and the next. The basic stimulation parameters for each brand are shown in **Table 2**. For each channel, the patients were asked to rank their perceived tinnitus during electrical stimulation on the VAS. The level of tinnitus reduction was expressed in percentiles relative to the baseline loudness and calculated using the following equation:

$$R^T = (L^0 - -L^1) \times 100/L^0$$

where R^T represents the amount of tinnitus reduction, and L^0 represents the baseline perceived tinnitus loudness before stimulation. The loudness ranked on the VAS during stimulation is denoted as L^1. A tinnitus reduction of 0% corresponds to no change in perceived tinnitus loudness, while positive values correspond to tinnitus reduction and negative values correspond to a worsening of tinnitus.

According to Arts et al. (2016), R^T was considered significant when > 50%, and it was graduated as follows, namely, complete suppression of tinnitus ($90\% < R^T \leq 100\%$), relevant attenuation ($50\% < R^T \leq 90\%$), mild attenuation ($10\% < R^T \leq 50\%$), and no effect on tinnitus ($R^T \leq 10\%$).

R^T was measured in each patient for each channel of the CI. The attenuation/suppression effect of electrical stimulation on tinnitus perception was considered "selective" when it was possible to find an electrical channel with a significantly higher tinnitus-reducing effect than the other channels [suppressive channel (SC)]; otherwise, the effect was considered "non-selective." The frequency ranges assigned to the SCs were considered to establish their location inside the cochlea. Multiple comparisons between mean L^0 and mean L^1 measured for each frequency range have been done to find the cochlear area most affected by tinnitus improvement in our sample.

Once the channel with the highest attenuation effect on tinnitus was identified, the intensity of the current was gradually reduced until tinnitus occurred again in the patient to identify the minimum current level suppressing tinnitus (MCLT) (μA).

- *Pitch-matching procedure:* The variability in the cochlear size and, above all, the different insertion depths of the array can create a mismatch between the default frequencies assigned to each channel and their actual location in the cochlea (Di Nardo et al., 2007a, Di Nardo et al., 2010). Patients with hearing residues in the non-implanted ear underwent the procedure of mismatch evaluation (Di Nardo et al., 2011), consisting of the pitch comparison between the electrical and acoustic stimuli sent independently to the two ears. Each pitch comparison trial consisted in first stimulating an electrical channel of the CI with a 5 s electric stimulus, followed by a 5 s acoustic pure tone presented to the other ear to find the corresponding frequency. The patient was asked to verbally report the tone sent to the other ear as higher, lower, or similar in pitch. The frequency of the acoustic stimulus was adaptively changed by 1/12th of an octave up or down, depending on the subject's response. A minimum of three matching attempts were conducted. Both stimuli could be presented in a random order where either electrode stimulus or acoustic tone was presented first. The procedure can be repeated for all channels, but for the purpose of the study it was sufficient to find the SC pitch. The procedure requires a good degree of auditory rehabilitation; therefore, it was performed 6 weeks after CI activation.

Statistical Analysis

Statistical analysis was performed using Microsoft Excel (Microsoft Corporation, Redmond, WA, United States). Continuous values, such as the THI score, are expressed as mean ± standard deviation (SD). Qualitative variables were summarized with absolute and percentage frequency tables.

The primary objective of this study was achieved by calculating the greatest reduction of the perceived loudness of tinnitus (R^T) on a VAS in each patient during short-term single-channel stimulation, with the abovementioned formula [$R^T = (L^0 – L^1) \times 100/L^0$], considering R^T to be significant when it was > 50%. The electrical channel with the greatest R^T compared to the other channels was found in each patient.

Multiple t-test comparisons between mean L^0 and mean L^1 measured for each frequency range have been done to find the cochlear area most affected by tinnitus improvement in our sample. The secondary objective, consisting of the comparison between continuous VAS scores obtained during single-channel stimulation and after CI activation, was achieved using the t-test for paired data. Long-term CI results on tinnitus perception were measured through THI before surgery, 1 month, 6 months, and 1 year after surgery; mean THI score obtained for each time point was compared to the preoperative score using t-test for paired data. The intergroup comparison considering the qualitative characteristics of tinnitus (incidence of monotone vs. polyphonic tinnitus) was performed using the chi-square test. The results were considered significant for p values < 0.05.

RESULTS

Patients

A total of 21 adult patients suffering from profound sensorineural hearing loss (HL) and subjective chronic tinnitus undergoing unilateral cochlear implantation were finally included (**Table 1**): 13 women (61.9%) and 8 men (38.1%) aged between 20 and 80 years (mean: 53.5 ± 15). **Table 1** summarizes the patients' demographic characteristics, side of deafness, grading and causes of hearing loss, as well as preoperative audiometric data.

The implanted devices used were Cochlear (Cochlear Ltd, Melbourne, Australia) (17/21), MED-EL (MED-EL Corp., Innsbruck, Austria) (3/21), and Advanced Bionics (Advanced Bionics LLC, Valencia, CA) (1/21). **Table 2** describes the implanted devices, characteristics of the electrodes, strategy, pulse width, and frequency of stimulation used.

Tinnitus Characteristics

The patients in question had been suffering from chronic tinnitus for an average of 18.5 (± 13.6) years. The questionnaire developed by Wang et al. (2017) regarding the characteristics of tinnitus showed that 8/21 (38.1%) patients had unilateral tinnitus (ipsilateral to the CI) while 13/21 (61.9%) patients had bilateral tinnitus.

Tinnitus was described as monotone in 16/21 (76.2%) patients and polyphonic in 5/21 (23.8%). A total of 21/21 (100%) patients were affected by continuous tinnitus.

Tinnitus loudness assessed using the VAS before surgery ($L^S = 6.5 \pm 2.5$) and 4 weeks after surgery (before the activation) ($L^0 = 6.4 \pm 2.4$) did not change significantly ($p > 0.05$). Also, 30 min after CI activation and mapping (CI turned on), the mean VAS value (L^2) was 4.3 ± 2.5. CI activation induced a significant reduction of subjective perception of tinnitus as measured by the VAS (L^0 vs. L^2; $p = 0.0095$), even in the early stage of stimulation (**Figure 1**). Specifically, 10/21 (47.6%) patients reported a significant improvement in VAS ($R^T > 50\%$) after CI activation, whereas in 2/21 (9.5%) the tinnitus worsened when the CI was turned on. In 2/21 (9.5%) patients, CI activation had no effect on subjective perception of tinnitus.

Short-Term Effect of Single-Channel Stimulation and Detection of Cochlear Regions Involved in the Attenuation/Suppression of Tinnitus

The experimental procedure showed that 8/21 patients (38%) experienced complete suppression of tinnitus ($90\% < R^T \leq 100\%$) through the activation of a specific channel (SC); 7/21 (33.3%) experienced a significant but not full attenuation of tinnitus ($50\% < R^T \leq 90\%$) with the SC (**Figure 2**).

In 4/21 patients, the tinnitus effect was non-selective: 3/21 (14.3%) reported a significant ($50\% < R^T \leq 90\%$) but non-selective attenuation, 1/21 (4.7%) reported a mild and non-selective attenuation ($10\% < R^T \leq 50\%$). 2/21 (9.5%) had no tinnitus effect ($R^T \leq 10\%$). No one reported worsened tinnitus symptoms during or after stimulation.

Considering the mean baseline loudness L^0 6.4 ± 2.4, the mean VAS score reported by patients during the experimental procedure (L^1) was 1.7 ± 2.7, with a significant difference ($p = 0.003$) and a mean reduction in loudness of 74.9%. Consequently, short-term targeted stimulation was more effective than the stimulation with the whole CI in our sample (L^1 vs. L^2, $p = 0.0002$) in the early phase of activation (**Figure 1**).

A total of 15/21 patients (71.4%) had a significant R^T *via* the stimulation of a specific channel, reporting a mean L^1 of 0.4 ± 2.0 (L^0 vs. L^1, $p = 0.0001$), with an average reduction in loudness of 93%. Thus, for most patients, tinnitus improvement was caused by the stimulation of a narrow area of the cochlea. The characteristics of the electrodes (12/16 perimodiolar-positioned and 3/6 lateral wall) and their length, as shown in **Table 2**, did not influence tinnitus perception. We identified the best-performer channel as reported in **Figure 2** and **Table 2**. Namely, for Cochlear implants (17/21), the selective stimulation on channel number 22 (188–313 Hz band) significantly reduced the tinnitus perception in 6/17 (35.3%) patients, Cochlear number 19 (563–688 Hz) in 1/17 (5.9%), Cochlear number 17 (813–938 Hz) in 1/17 (5.9%), Cochlear number 11 (1,813–2,063 Hz) in 2/17 (11.8%), and Cochlear number 6 (3,563–4,063 Hz) in 3/17 (17.6%). For MED_EL CIs (3/21), channel number 1 (70–170 Hz) reduced tinnitus in 1/3 (33.3%) patients; MED-EL channel number 4 (491–710 Hz) in 1/3 (33.3%).

Seven patients with hearing residues in the non-implanted ear were able to perform the pitch-matching test. In three patients (N1, N2, N5), the pitch heard during the procedure did not fall within the frequency range empirically assigned to the SC: patient N1 (channel 22 – 125 Hz); patient N2 (channel 6 – 3,400 Hz); and patient N5 (channel 17 – 625 Hz). **Table 3** shows the results of the pitch-matching procedure.

When considered together, our results showed that the best-performer channels fell within the following frequency bands: 125–313 Hz in 7/15 (46.6%); 563–688 Hz in 3/15 (20%); 1,813–2,063 Hz in 2/15 (13.3%); 3,063–3,563 in 1/15 (6.6%); and 3,563–4,063 Hz in 2/15 (13.3%) (**Figure 2** and **Table 2**). Based on cochlear tonotopy, in 10/15 (66.6%) patients the SC was found in its apical turn, within 1,000 Hz, whereas in 5/15 patients (33.4%) it was found within 4,000 Hz.

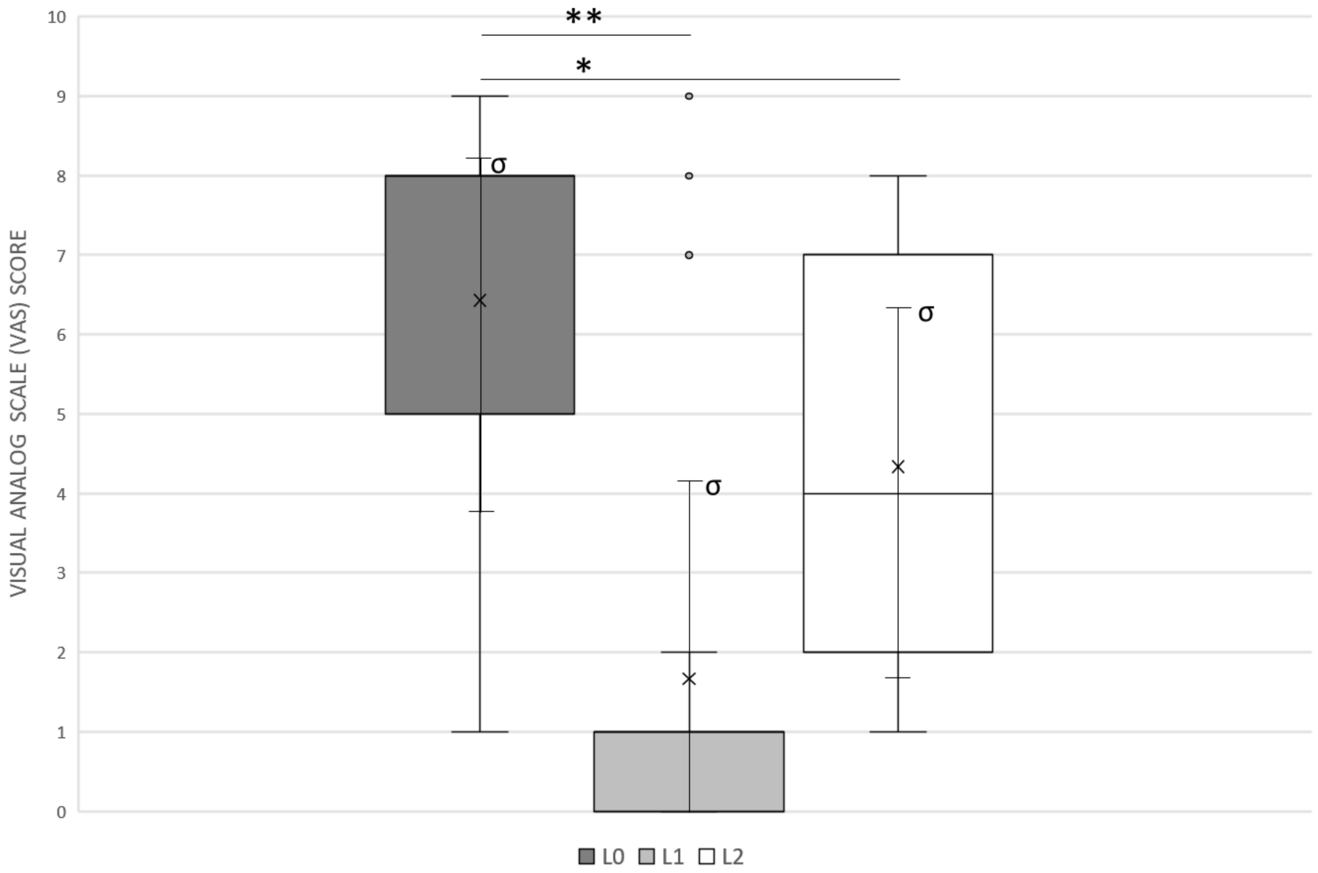

FIGURE 1 | The Visual Analog Scale (VAS) was used to assess tinnitus modification. Patients were asked to rate the perceived loudness of their tinnitus on a numeric scale, from 0 (inaudible) to 10 (loud like never before). L^0, basal loudness of tinnitus (mean 6.4 ± 2.4); L^1, minimal loudness of tinnitus at 10 s single-channel stimulation (mean 1.7 ± 2.7); L^2, loudness of tinnitus after 30 min of whole CI activation (4.3 ± 2.5). Median values are displayed as a horizontal line, mean values as x, standard deviation as σ. We used *t*-test for statistical analysis. $^{**}p < 0.01$, $^{*}p < 0.05$.

Comparing mean L^0 with mean L^1 measured for each frequency band, the one corresponding to 125–313 Hz resulted as the cochlear area most affected by tinnitus improvement in our sample ($p = 0.0074$).

To assess the intensity of electric current necessary to suppress tinnitus, for each SC, the MCL value was identified: it was on average 388.5 ± 173.4 μA. For the different frequency bands, 381 μA (125–313 Hz); 522.2 μA (563–688 Hz); 368 μA (813–938 Hz); 225.5 μA (1,813–2,063 Hz); 244.4 μA (3,063–3,563 Hz); 417.2 μA (3,563–4,063 Hz) (**Table 2**). We observed that the intensity of electric current was not related to the baseline loudness of tinnitus measured with the VAS (Pearson $R = 0.03$), nor to the THI (Pearson $R = 0.03$). Conversely, it was shown to be independent of the subjective perception of tinnitus. **Table 2** shows the results of the test for each patient.

All patients undergoing the experimental procedure were retested to verify that the best-performer channel described in the early phase of activation was the same after 2 weeks. The repeatability rate of our experimental procedure, calculated as the percentage of patients whose measures were found to be the same after 2 weeks, was 95%, obtaining the same results in 20/21 patients. The results only changed between the two settings in one patient (N5) [channel number 21 (range 313–438 Hz) instead of channel 17].

Tinnitus Effect on Daily Life and Long-Term Effect of Intracochlear Electrical Stimulation

When asked to complete the Qian Wang questionnaire, 8/21 (38%) patients reported that the most serious impairment was related to quality of sleep, 6/21 (28.6%) to hearing and speech perception, 4/21 (19%) to emotional state, 2/21 (9.5%) to memory and ability to concentrate, and 1/21 (4.7%) to work activity.

The mean preoperative THI was 39.7 ± 27.2. Four weeks after surgery, the mean THI was 40.3 ± 26.7 (before activation), with no statistically significant difference ($p > 0.05$); in 6/21 (28.6%) patients, the tinnitus worsened in the immediate postoperative period showing an increased grading class, while 5/21 (23.8%) patients reported tinnitus improvement following surgery with a decreased grading class.

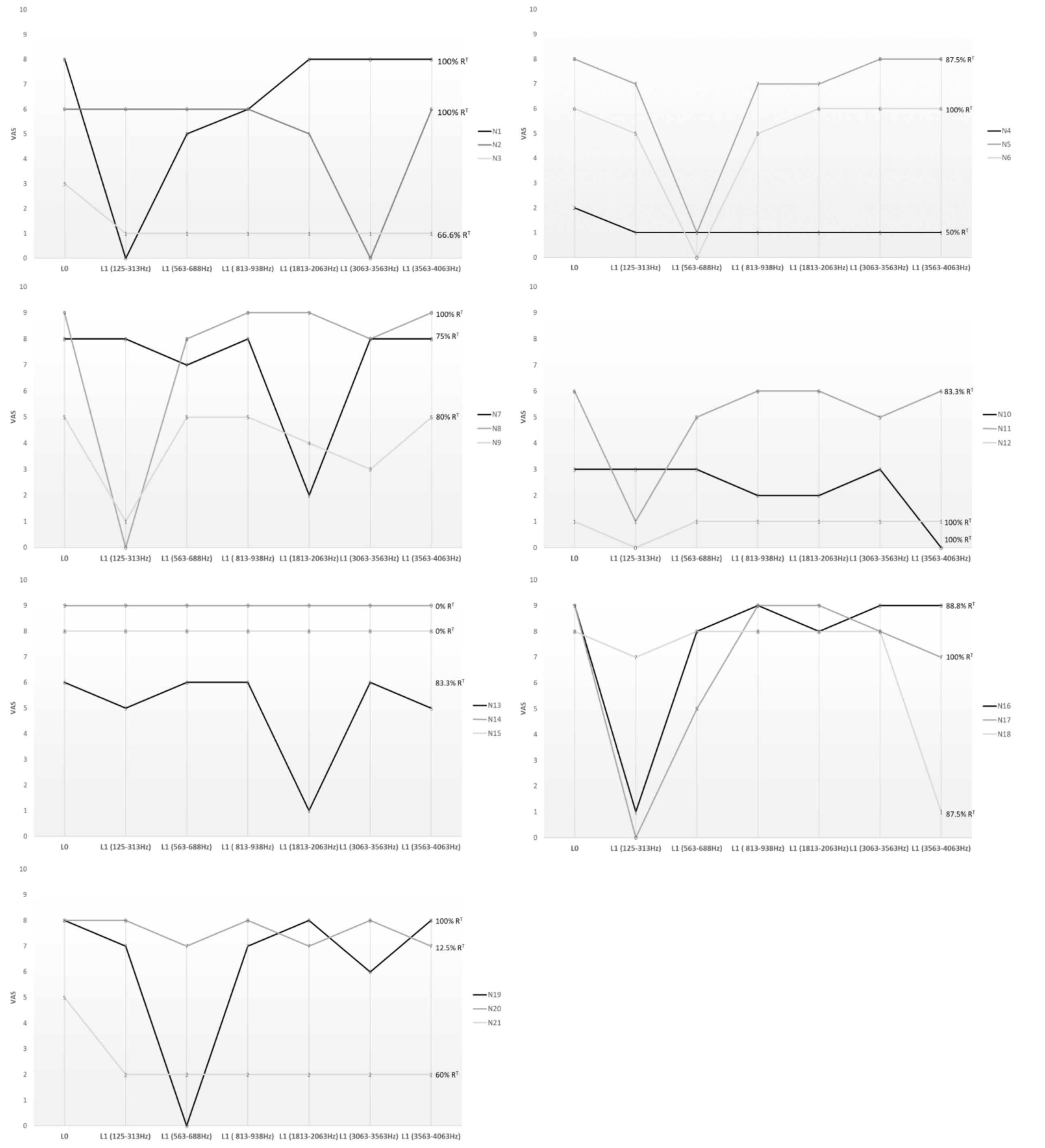

FIGURE 2 | The Visual Analog Scale (VAS) was used to assess tinnitus modification. Patients were asked to rate the perceived loudness of their tinnitus on a numeric scale, from 0 (inaudible) to 10 (loud like never before), during short-term single-channel stimulation. The figure shows the cochlear regions involved in each patient. Only the frequency bands corresponding to the best-performer channels are depicted in the figure, not the entire cochlear tonotopy. L^0, basal loudness of tinnitus; L^1, loudness of tinnitus at 10 s single-channel stimulation; R^T, greatest percentage of tinnitus reduction.

Six months after activation, the THI was 27.8 ± 29.5 with a statistically significant difference compared to the preoperative value (p = 0.0325). Mean THI at 1 year (26.7 ± 28.9) was also significantly higher (p = 0.0244). There were no significant changes between 6 months and 1 year. **Figure 3** shows the trend of THI score in our sample.

TABLE 3 | Results of the pitch matching procedure.

Patient	SC	Frequency band	Real pitch
N1	22 Cochlear	188–313 Hz	**125 Hz**
N2	6 Cochlear	3563–4063 Hz	**3400 Hz**
N5	17 Cochlear	813–938 Hz	**625 Hz**
N8	22 Cochlear	188–313 Hz	225 Hz
N13	11 Cochlear	1813–2063 Hz	2000 Hz
N17	1 MED-EL	100–198Hz	170 Hz
N19	19 Cochlear	563–688 Hz	625 Hz

In three patients (N1, N2, N5), the actual pitch measured (bold values) did not fall 924 within the frequency range empirically assigned to the suppressive channel (SC).

Qualitative Characteristics of Tinnitus and Long-Term Results (Intergroup Comparison)

The qualitative characteristics of tinnitus were analyzed and compared between the group of patients who had a best-performer channel (15/21), the patients who had a non-selective attenuation (4/21), and the two patients who had no tinnitus changes during the experimental procedure (2/21).

The four patients who had a non-selective attenuation were compared to patients with selective suppression (**Table 4**). They had a more recent onset of tinnitus (11 vs. 20.8 years, $p < 0.05$), a higher incidence of subjectively defined polyphonic tinnitus (3/4 – 75% vs. 0/15 – 0%, $p < 0.05$), a lower baseline VAS with CI off (4.5 vs. 6.6, $p < 0.05$), and with CI on (2.5 vs. 4.5, $p < 0.05$), a lower THI pre-implantation (36.5 vs. 40.6, $p > 0.05$), on the day of CI activation (23 vs. 44.6, $p < 0.05$), and 1 year after surgery (18.5 vs. 30.5, $p < 0.05$).

The two patients who did not report changes during the experimental procedure, compared to patients having selective suppression, had subjectively defined polyphonic tinnitus (2/2 – 100% vs. 0/15 – 0%, $p < 0.05$), a higher VAS with CI off (8 vs. 6.6, $p < 0.05$), and on (6.5 vs. 4.5, $p < 0.05$), a higher THI 1 year after surgery (40 vs. 30.5, $p < 0.05$).

Therefore, patients in whom it was possible to obtain a short-term suppression of tinnitus with a specific channel had a subjectively defined monotone tinnitus of longer duration and a greater subjective perception of tinnitus. On the other hand, the non-specific attenuation was related to better long-term results measured with THI, possibly due to a lower perceived loudness and the possibility to attenuate the tinnitus by stimulating the entire cochlea rather than selectively stimulating a single electrical channel, which is the way the CI normally works. Patients who reported no changes in tinnitus perception during the experimental procedure also had poor long-term results with 1-year THI.

The research had no missing data for any of the measured variables.

DISCUSSION

Our results showed that the intracochlear electrical stimulation significantly reduced the subjective tinnitus perception, in agreement with previous reports (Quaranta et al., 2004; Ramakers et al., 2015; Peter et al., 2019). In detail, we observed a significant improvement of tinnitus handicap severity 6 months after CI activation as measured with the THI (from 39.7 ± 27.2 to 27.8 ± 29.5; $p = 0.0325$). It is reasonable to expect that tinnitus can be partially alleviated by increasing auditory stimulation, considering the current consensus that chronic tinnitus is the result of maladaptive plasticity of the auditory cortex consequent to sensorial deprivation (Engineer et al., 2011). Increasing evidence indicates that electrical current is a promising treatment itself, independently of auditory stimulation (Shepherd and Javel, 1999; Konopka et al., 2001; Rubinstein et al., 2003; Steenerson and Cronin, 2003; Aydemir et al., 2006; De Ridder et al., 2006; Seidman et al., 2008; Di Nardo et al., 2009; Noreña et al., 2015; Paciello et al., 2018). To define which stimulation patterns should be optimized for tinnitus relief, a deeper understanding of the mechanisms involved in tinnitus suppression is needed (Assouly et al., 2021). CI represents the main tool to investigate the effects of intracochlear electric current in patients with chronic tinnitus.

The experimental animal models revealed that the mechanism behind the efficacy of electric current on tinnitus was the reversal of the reorganization of the auditory structures involved in chronic tinnitus, rather than the shift in attention from tinnitus to external sounds (Noreña et al., 2015; Paciello et al., 2018). This mechanism, based on the effectiveness of electrical hearing, could explain the evidence of persistent tinnitus reduction in many patients after the CI stimulation is turned off (Quaranta et al., 2004). Interestingly, previous studies showed short-term tinnitus suppression *via* CI, independently of environmental sound stimulation, by interfering with the fitting software (Arts et al., 2015, 2016). Accordingly, it was suggested that the coding of environmental sounds is not required to reduce tinnitus.

However, it remains unclear whether the beneficial results on tinnitus observed with CI, in a percentage of patients that varies from 25 to 72% (Ramakers et al., 2015), depend on the undifferentiated stimulation of the whole cochlear partition or on a spatially limited alteration induced by electric current.

Therefore, we aimed to evaluate whether the single-channel stimulation of CI could improve tinnitus perception and whether it was possible to identify a region in the cochlea whose stimulation would cause beneficial effects on tinnitus. To exclude that the effects on tinnitus perception could depend on electrical hearing and avoid any possible alteration of baseline tinnitus characteristics induced from electric current, the experimental procedure was performed in the early phase of CI activation. We introduced a new procedure to be combined with CI mapping in patients affected by chronic tinnitus, showing that electric current delivered on a single channel of the CI for 10 s constituted the "trigger" of tinnitus suppression in 71.4% of our sample [L^0 (6.4 ± 2.4) vs. L^1 (1.7 ± 2.7), $p = 0.003$]. Patients in whom it was possible to identify a best-performer channel reported a significant reduction in their tinnitus-perceived loudness as measured with the VAS during stimulation ($L^1 = 0.4 \pm 2.0$, $p = 0.0001$), with an average reduction in loudness of 93%. The results suggest that tinnitus subjective perception could be partially or totally alleviated by an electrical stimulus

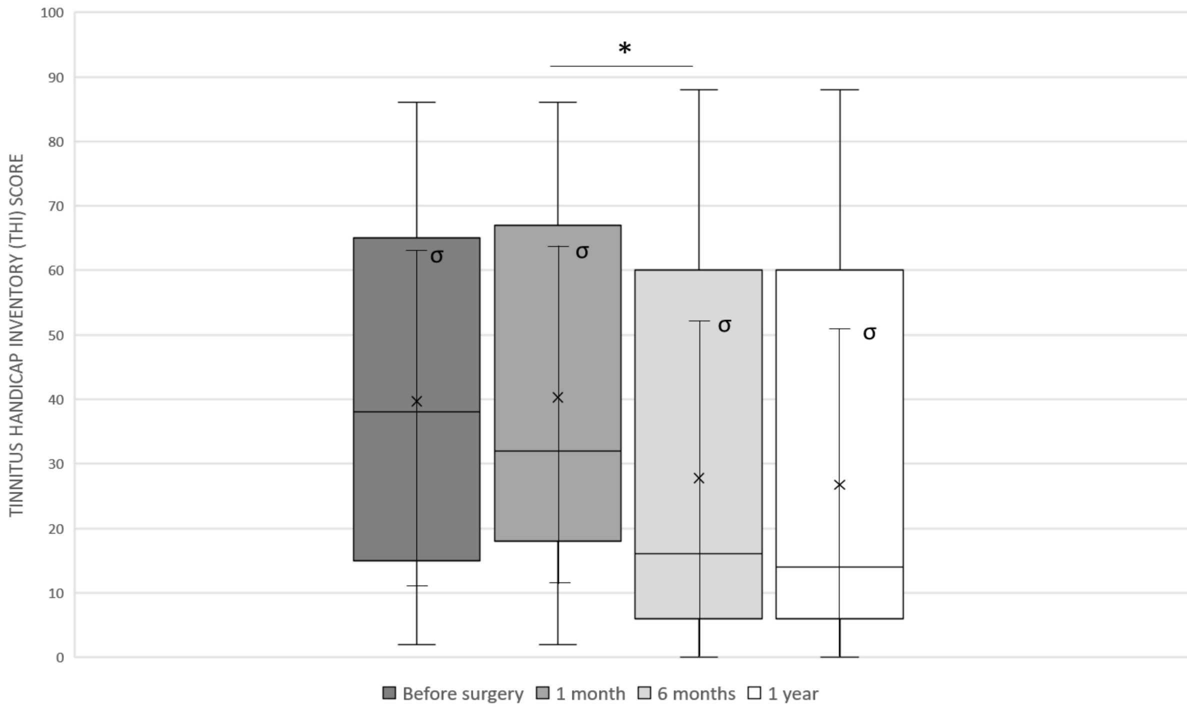

FIGURE 3 | The Tinnitus Handicap Inventory (THI) trend in our sample: before surgery (mean 39.7 ± 27.2), 4 weeks after surgery (mean 40.3 ± 26.7), after 6 months (mean 27.8 ± 29.5) and after 1 year (mean 26.7 ± 28.9). Median values are displayed as a horizontal line, mean values as x, standard deviation as σ. We used *t*-test for statistical analysis. *$p < 0.05$.

TABLE 4 | Intergroup comparison.

Patients	Polyphonic tinnitus	mean VAS (CI off – L0)	mean VAS (CI on – L2)	mean THIpre-op.	mean THI1 year
Selective attenuation of tinnitus15/21	0%	6.6	4.5	40.6	30.5
Non-selective attenuation of tinnitus4/21	75%*	4.5*	2.5*	36.5	18.5*
No effect on tinnitus2/21	100%*	8*	6.5*	72*	40*

**Statistically significant difference compared to patients with specific attenuation on tinnitus, $p < 0.05$.*

targeted to a limited region of the cochlea rather than to the entire cochlear duct.

Furthermore, short-term targeted stimulation was more effective than stimulation involving the whole CI in the early phase of activation [L^1 (1.7 ± 2.7) vs. L^2 (4.3 ± 2.5), $p = 0.0022$]. The possibility to find a more sensitive area to tinnitus attenuation even in postverbal deaf patients with an almost completely compromised cochlea suggests that tinnitus probably arises from a limited area of the basilar membrane and secondarily involves other relays of the auditory pathway. For this reason, we hypothesized that the stimulation of a specific region of the damaged cochlea could be the most involved in long-term tinnitus improvement and in the reversal of cortical reorganization maintaining chronic tinnitus (Haider et al., 2018).

Another important challenge was to understand whether and how the location of the electrical channels along the CI array and their corresponding frequency were related to the reduction of tinnitus. Several concerns affect the knowledge of the exact tonotopic position of the channels inside the cochlea, including the different electrodes used, their different insertion depths, and the variable tonotopic organization in damaged cochleae (Rak et al., 2020; Niu et al., 2021). Consequently, it is very difficult to determine exactly the regions stimulated by the channels, except in patients with asymmetric hearing thresholds. Despite these issues, the exact pitch of the best-performer channel was identified in seven patients who underwent the pitch-matching procedure, showing that the most effective stimulation frequencies were those at the apex of the cochlea, from 125 to 313 Hz ($p = 0.0074$). This finding seems to run counter to the experimental model of tinnitus consequent to acoustic trauma or ototoxic drugs and suggests that in most cases the initial dysfunction is in the basal turn of the cochlea (Ralli et al., 2014; Paciello et al., 2020). However, our patients were affected by severe to profound hearing loss, indicating extensive cochlear damage. Furthermore, since CIs stimulate the spiral ganglion fibers, it is reasonable to assume that their unpredictable number and particularly their distribution also influence the results (Dhanasingh et al., 2020). As it stands, we do not know if the cochlear region with the greatest attenuation capacity on tinnitus is that of initial cochlear damage triggering tinnitus. Considering that cochlear damage was demonstrated to determine the overrepresentation of the edge frequencies in the cortex due to the loss of lateral inhibition phenomena (Eggermont and Roberts, 2004), we hypothesized that the electrical stimulation of the overrepresented area can mask or reverse mechanisms favoring tinnitus. While some studies have localized the tinnitus pitch at the edge of hearing loss, others found that it occurs in the region of maximum hearing loss (Sereda et al., 2011). We were not aware of the real tinnitus pitch in our sample as it was only subjectively defined with the Qian Wang questionnaire because of the severe-to-profound hearing loss of the patients and the impossibility to evaluate it with the audiometry. In addition, this study, by nature, evaluates tinnitus through subjective scores (THI, VAS), by asking participants to rate their tinnitus several times during the experiments, which could have influenced the results. To select a cohort of patients with single-sided deafness or asymmetric thresholds, it should be essential in the future to define whether the tinnitus pitch falls within the frequency band assigned to the suppressive electrical channel and rate the tinnitus loudness using audiometry to further deepen our understanding of the pathophysiological mechanisms of tinnitus. Nevertheless, considering our results, the electrical stimulation of the apical side of the cochlea could represent how to turn off the mechanisms that feed tinnitus.

Coherently with our findings, it has been reported that electrical stimulation in the first 10 mm of the basal part of the scala tympani is not sufficient to reduce tinnitus. Conversely, the stimulation over the complete CI length produced an immediate tinnitus reduction (Punte et al., 2013), suggesting that the stimulation of the basal channels is ineffective for tinnitus attenuation.

Interesting results emerged from the 1-year follow-up with regard to the THI questionnaire: patients with significant tinnitus attenuation achieved through non-selective electrical stimulation, all reporting polyphonic tinnitus, showed lower THI scores compared to CI users with monotone tinnitus in whom the suppressive channel was detected. We assume that the better long-term results in patients with polyphonic tinnitus were related to the effective non-selective attenuation of tinnitus by stimulating along the cochlea through multichannel stimulation. In other words, patients with monotone tinnitus who reported immediate tinnitus attenuation from a single-channel stimulation could not achieve similar long-term results by alternatively stimulating the various frequency bands, which is how the CI normally works. A possible explanation is that a selective continuous amplification of the channel identified by the experimental procedure could be essential in those patients to obtain lasting effects on tinnitus. The long-term CI effects on tinnitus have been assessed by Arts et al. (2015) through the establishment of a sort of "tinnitus implant," that is, a CI suitably programmed in the different variables considering the parameters subjectively evaluated as more effective in the suppression of tinnitus. These authors, however, do not consider the possibility that a spatially targeted electrical stimulation could offer an additional advantage. Major findings are needed to confirm the correspondence between the stimulation of the cochlear apical turn and tinnitus improvement. Thus, the amplification of the suppressive electrical channel could represent an effective "tinnitus implant" for future clinical use.

In conclusion, the study confirmed the improvement of long-term tinnitus with intracochlear electrical stimulation in CI users. We experimented a procedure focused on the identification and characterization of the suppressive channel obtaining evidence of tinnitus reduction through a targeted electrical stimulation only in patients affected by monotone tinnitus, independently of the tinnitus pitch. We demonstrated that the cochlear apical stimulation is more effective in tinnitus suppression, opening new scenarios for the knowledge of the pathophysiology of tinnitus and future therapeutic applications. The data collected could be helpful for the future development of an implantable stimulator or optimize the CI parameters for tinnitus suppression.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comitato Etico Fondazione Policlinico Agostino Gemelli IRCCS. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

WD and AF conceived the presented idea and were in charge of overall direction and planning. TD and AT carried out the experiments, collected the data, and contributed to the interpretation of the results. GP supervised the findings of this work. All authors discussed the results and contributed to the final version of the manuscript.

ACKNOWLEDGMENTS

We acknowledge Minioto F. and Martina B. M. for technical assistance and data collection.

REFERENCES

Arts, R. A., George, E. L., Chenault, M. N., and Stokroos, R. J. (2015). Optimizing intracochlear electrical stimulation to suppress tinnitus. *Ear. Hear.* 36, 125–135. doi: 10.1097/AUD.0000000000000090

Arts, R. A., George, E. L., Janssen, M., Griessner, A., Zierhofer, C., and Stokroos, R. J. (2016). Tinnitus suppression by intracochlear electrical stimulation in single sided deafness – a prospective clinical trial: follow-up. *PLoS One.* 11:e0153131. doi: 10.1371/journal.pone.0153131

Assouly, K., Dullaart, M. J., Stokroos, R. J., van Dijk, B., Stegeman, I., and Smit, A. L. (2021). Systematic review on intra- and extracochlear electrical stimulation for tinnitus. *Brain sci.* 11, 1394. doi: 10.3390/brainsci11111394

Aydemir, G., Tezer, M. S., Borman, P., Bodur, H., and Unal, A. (2006). Treatment of tinnitus with transcutaneous electrical nerve stimulation improves patients' quality of life. *J. Laryngol. Otol.* 120, 442–445. doi: 10.1017/S0022215106000910

Bhatt, J. M., Lin, H. W., and Bhattacharyya, N. (2016). Prevalence, severity, exposures, and treatment patterns of tinnitus in the United States. *JAMA Otolaryngol. Head Neck Surg.* 142, 959–965. doi: 10.1001/jamaoto.2016.1700

De Ridder, D., De Mulder, G., Verstraeten, E., Van der Kelen, K., Sunaert, S., Smits, M., et al. (2006). Primary and secondary auditory cortex stimulation for intractable tinnitus. *ORL* 68, 48–55. doi: 10.1159/000090491

Dhanasingh, A. N., Jolly, C., Rajan, G., and van de Heyning, P. (2020). Literature Review on the distribution of spiral ganglion cell bodies inside the human cochlear central modiolar trunk. *J. Int. Adv. Otol.* 16, 104–110. doi: 10.5152/iao.2020.7510

Di Nardo, W., Cantore, I., Cianfrone, F., Melillo, P., Fetoni, A. R., and Paludetti, G. (2007a). Differences between electrode-assigned frequencies and cochlear implant recipient pitch perception. *Acta Otolaryngol.* 127, 370–377. doi: 10.1080/00016480601158765

Di Nardo, W., Cantore, I., Cianfrone, F., Melillo, P., Scorpecci, A., and Paludetti, G. (2007b). Tinnitus modifications after cochlear implantation. *Eur. Arch. Otorhinolaryngol.* 264, 1145–1149. doi: 10.1007/s00405-007-0352-7

Di Nardo, W., Cianfrone, F., Scorpecci, A., Cantore, I., Giannantonio, S., and Paludetti, G. (2009). Transtympanic electrical stimulation for immediate and long-term tinnitus suppression. *Int. Tinnitus J.* 15, 100–106.

Di Nardo, W., Scorpecci, A., Giannantonio, S., Cianfrone, F., and Paludetti, G. (2011). Improving melody recognition in cochlear implant recipients through individualized frequency map fitting. *Eur. Arch Otorhinolaryngol.* 268, 27–39. doi: 10.1007/s00405-010-1335-7

Di Nardo, W., Scorpecci, A., Giannantonio, S., Cianfrone, F., Parrilla, C., and Paludetti, G. (2010). Cochlear implant patients' speech understanding in background noise: effect of mismatch between electrode assigned frequencies and perceived pitch. *J. Laryngol. Otol.* 124, 828–834. doi: 10.1017/S0022215110000320

Eggermont, J. J., and Roberts, L. E. (2004). The neuroscience of tinnitus. *Trends Neurosci.* 27, 676–682.

Engineer, N. D., Riley, J. R., Seale, J. D., Vrana, W. A., Shetake, J. A., Sudanagunta, S. P., et al. (2011). Reversing pathological neural activity using targeted plasticity. *Nature* 470, 101–104.

Fetoni, A. R., Troiani, D., Petrosini, L., and Paludetti, G. (2015). Cochlear injury and adaptive plasticity of the auditory cortex. *Front. Aging Neurosci.* 7:8. doi: 10.3389/fnagi.2015.00008

Gallus, S., Lugo, A., Garavello, W., Bosetti, C., Santoro, E., Colombo, P., et al. (2015). Prevalence and determinants of tinnitus in the Italian adult population. *Neuroepidemiology* 45, 12–19. doi: 10.1159/000431376

Garaycochea, O., Manrique-Huarte, R., and Manrique, M. (2020). Intra-operative radiological diagnosis of a tip roll-over electrode array displacement using fluoroscopy, when electrophysiological testing is normal: the importance of both techniques in cochlear implant surgery. *Braz. J. Otorhinolaryngol.* 86(Suppl. 1), 38–40. doi: 10.1016/j.bjorl.2017.05.003

Greenwood, D. D. (1961). Critical bandwidth and the frequency coordinates of the basilar membrane. *J. Acoust. Soc. Am.* 33, 1344–1356.

Haider, H. F., Bojiæ, T., Ribeiro, S. F., Paço, J., Hall, D. A., and Szczepek, A. J. (2018). Pathophysiology of subjective tinnitus: triggers and maintenance. *Front. Neurosci.* 12:866. doi: 10.3389/fnins.2018.00866

Henry, J. A., Roberts, L. E., Caspary, D. M., Theodoroff, S. M., and Salvi, R. J. (2014). Underlying mechanisms of tinnitus: review and clinical implications. *J. Am. Acad. Audiol.* 25, 5–126. doi: 10.3766/jaaa.25.1

Knipper, M., Van Dijk, P., Nunes, I., Rüttiger, L., Zimmermann, U., et al. (2013). Advances in the neurobiology of hearing disorders: recent developments regarding the basis of tinnitus and hyperacusis. *Prog. Neurobiol.* 111, 17–33. doi: 10.1016/j.pneurobio.2013.08.002

Konishi, T., Teas, D. C., and Wernick, J. S. (1970). Effects of electrical current applied to cochlear partition on discharges in individual auditory-nerve fibers. I. Prolonged direct-current polarization. *J. Acoust. Soc. Am.* 47, 1519–1526. doi: 10.1121/1.1912083

Konopka, W., Zalewski, P., Olszewski, J., Olszewska-Ziaber, A., and Pietkiewicz, P. (2001). Tinnitus suppression by electrical promontory stimulation (EPS) in patients with sensorineural hearing loss. *Auris Nasus Larynx* 28, 35–40. doi: 10.1016/s0385-8146(00)00086-9

Lockwood, A. H., Salvi, R. J., and Burkard, R. F. (2002). Tinnitus. *N. Engl. J. Med.* 347, 904–910.

McCombe, A., Baguley, D., Coles, R., McKenna, L., McKinney, C., and Windle-Taylor, P. (2001). Guidelines for the grading of tinnitus severity: the results of a working group commissioned by the British association of otolaryngologists, head and neck surgeons, 1999. *Clin. Otolaryngol. Allied Sci.* 26, 388–393. doi: 10.1046/j.1365-2273.2001.00490.x

Monzani, D., Genovese, E., Marrara, A., Gherpelli, C., Pingani, L., Forghieri, M., et al. (2008). Validity of the Italian adaptation of the tinnitus handicap inventory; focus on quality of life and psychological distress in tinnitus-sufferers. *Acta otorhinolaryngol.* 28, 126.

Mulders, W. H., and Robertson, D. (2009). Hyperactivity in the auditory midbrain after acoustic trauma: dependence on cochlear activity. *Neuroscience* 164, 733–746. doi: 10.1016/j.neuroscience.2009.08.036

Newman, C. W., Jacobson, G. P., and Spitzer, J. B. (1996). Development of the tinnitus handicap inventory. *Arch. Otolaryngol. Head Neck Surg.* 122, 143–148.

Niu, X. M., Ping, L., Gao, R. Z., Xia, X., Fan, X. M., Chen, Y., et al. (2021). Selection of cochlear electrode array implantation lengths and outcomes in patients with bilateral sensorineural hearing loss. *Zhonghua Yi Xue Za Zhi* 101, 108–114. doi: 10.3760/cma.j.cn112137-20201005-02770

Noreña, A. J. (2015). Revisiting the cochlear and central mechanisms of tinnitus and therapeutic approaches. *Audiol. Neurootol.* 20(Suppl. 1), 53–59. doi: 10.1159/000380749

Norena, A. J., and Eggermont, J. J. (2003). Changes in spontaneous neural activity immediately after an acoustic trauma: implications for neural correlates of tinnitus. *Hear. Res.* 183, 137–153. doi: 10.1016/s0378-5955(03)00225-9

Noreña, A. J., Mulders, W. H., and Robertson, D. (2015). Suppression of putative tinnitus-related activity by extra-cochlear electrical stimulation. *J. Neurophysiol.* 113, 132–143. doi: 10.1152/jn.00580.2014

Paciello, F., Fetoni, A. R., Mezzogori, D., Rolesi, R., Di Pino, A., Paludetti, G., et al. (2020). The dual role of curcumin and ferulic acid in counteracting chemoresistance and cisplatin-induced ototoxicity. *Sci. Rep.* 10, 1063. doi: 10.1038/s41598-020-57965-0

Paciello, F., Podda, M. V., Rolesi, R., Cocco, S., Petrosini, L., Troiani, D., et al. (2018). Anodal transcranial direct current stimulation affects auditory cortex plasticity in normal-hearing and noise-exposed rats. *Brain Stimul.* 11, 1008–1023. doi: 10.1016/j.brs.2018.05.017

Peter, N., Liyanage, N., Pfiffner, F., Huber, A., and Kleinjung, T. (2019). The influence of cochlear implantation on tinnitus in patients with single-sided deafness: a systematic review. *Otolaryngol. Head Neck Surg.* 161, 576–588. doi: 10.1177/0194599819846084

Pulec, J. L. (1984). Tinnitus: surgical therapy. *Am. J. Otol.* 5, 479–480.

Punte, A. K., De Ridder, D., and Van de Heyning, P. (2013). On the necessity of full length electrical cochlear stimulation to suppress severe tinnitus in single-sided deafness. *Hear. Res.* 295, 24–29. doi: 10.1016/j.heares.2012.08.003

Quaranta, N., Wagstaff, S., and Baguley, D. M. (2004). Tinnitus and cochlear implantation. *Int. J. Audiol.* 43, 245–251.

Rak, K., Ilgen, L., Taeger, J., Schendzielorz, P., Voelker, J., Kaulitz, S., et al. (2020). Influence of cochlear parameters on the current practice in cochlear implantation: development of a concept for personalized medicine. German version. *HNO* 69, 943–951. doi: 10.1007/s00106-020-00968-0

Ralli, M., Troiani, D., Podda, M. V., Paciello, F., Eramo, S. L., de Corso, E., et al. (2014). The effect of the NMDA channel blocker memantine on salicylate-induced tinnitus in rats. *Acta Otorhinolaryngol. Ital.* 34, 198–204.

Ramakers, G. G., van Zon, A., Stegeman, I., and Grolman, W. (2015). The effect of cochlear implantation on tinnitus in patients with bilateral hearing loss: a systematic review. *Laryngoscope* 125, 2584–2592. doi: 10.1002/lary.25370

Rauschecker, J. P., Leaver, A. M., and Mühlau, M. (2010). Tuning out the noise: limbic-auditory interactions in tinnitus. *Neuron* 66, 819–826. doi: 10.1016/j.neuron.2010.04.032

Rubinstein, J. T., Tyler, R. S., Johnson, A., and Brown, C. J. (2003). Electrical suppression of tinnitus with high-rate pulse trains. *Otol. Neurotol.* 24, 478–485. doi: 10.1097/00129492-200305000-00021

Runge-Samuelson, C. L., Abbas, P. J., Rubinstein, J. T., Miller, C. A., and Robinson, B. K. (2004). Response of the auditory nerve to sinusoidal electrical stimulation: effects of high-rate pulse trains. *Hear. Res.* 194, 1–13. doi: 10.1016/j.heares.2004.03.020

Seidman, M. D., Ridder, D. D., Elisevich, K., Bowyer, S. M., Darrat, I., Dria, J., et al. (2008). Direct electrical stimulation of Heschl's gyrus for tinnitus treatment. *Laryngoscope* 118, 491–450. doi: 10.1097/MLG.0b013e31815daf5a

Sereda, M., Hall, D. A., Bosnyak, D. J., Edmondson-Jones, M., Roberts, L. E., Adjamian, P., et al. (2011). Re-examining the relationship between audiometric profile and tinnitus pitch. *Int. J. Audiol.* 50, 303–312. doi: 10.3109/14992027.2010.551221

Shepherd, R. K., and Javel, E. (1999). Electrical stimulation of the auditory nerve: II. Effect of stimulus waveshape on single fibre response properties. *Hear. Res.* 130, 171–188. doi: 10.1016/s0378-5955(99)00011-8

Steenerson, R. L., and Cronin, G. W. (2003). Tinnitus reduction using transcutaneous electrical stimulation. *Otolaryngol. Clin. North Am.* 36:337. doi: 10.1016/s0030-6665(02)00164-0

Wang, Q., Li, J. N., Lei, G. X., Chen, D. S., Wang, W. Z., Chen, A. T., et al. (2017). Interaction of tinnitus suppression and hearing ability after cochlear implantation. *Acta OtoLaryngologica.* 137, 1077–1082. doi: 10.1080/00016489.2017.1336283

Zeng, C., Nannapaneni, N., Zhou, J., Hughes, L. F., and Shore, S. (2009). Cochlear damage changes the distribution of vesicular glutamate transporters associated with auditory and nonauditory inputs to the cochlear nucleus. *J. Neurosci.* 29, 4210–4217. doi: 10.1523/JNEUROSCI.0208-09.2009

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Neuroscience

ORIGINAL RESEARCH
published: 12 July 2022
doi: 10.3389/fnins.2022.892894

Children's Neural Sensitivity to Prosodic Features of Natural Speech and Its Significance to Speech Development in Cochlear Implanted Children

Yuebo Chen[1†], Qinqin Luo[2,3†], Maojin Liang[1], Leyan Gao[4], Jingwen Yang[5,6], Ruiyan Feng[4], Jiahao Liu[1,7], Guoxin Qiu[6], Yi Li[3], Yiqing Zheng[1,7*] and Shuo Lu[3,6*]

[1] Department of Otolaryngology, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China, [2] Department of Chinese Language and Literature, The Chinese University of Hong Kong, Hong Kong, Hong Kong SAR, China, [3] School of Foreign Languages, Shenzhen University, Shenzhen, China, [4] Neurolinguistics Teaching Laboratory, Department of Chinese Language and Literature, Sun Yat-sen University, Guangzhou, China, [5] Department of Neurology, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, [6] Department of Clinical Neurolinguistics Research, Mental and Neurological Diseases Research Center, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, [7] Hearing and Speech Science Department, Guangzhou Xinhua University, Guangzhou, China

OPEN ACCESS

Edited by:
*Jing Chen,
Peking University, China*

Reviewed by:
*Kara Hawthorne,
Gallaudet University, United States
Yuxia Wang,
Shanghai Jiao Tong University, China*

***Correspondence:**
*Yiqing Zheng
zhengyiq@mail.sysu.edu.cn
Shuo Lu
Lushuo@szu.edu.cn*

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *09 March 2022*
Accepted: *14 June 2022*
Published: *12 July 2022*

Citation:
Chen Y, Luo Q, Liang M, Gao L, Yang J, Feng R, Liu J, Qiu G, Li Y, Zheng Y and Lu S (2022) Children's Neural Sensitivity to Prosodic Features of Natural Speech and Its Significance to Speech Development in Cochlear Implanted Children. Front. Neurosci. 16:892894. doi: 10.3389/fnins.2022.892894

Catchy utterances, such as proverbs, verses, and nursery rhymes (i.e., "No pain, no gain" in English), contain strong-prosodic (SP) features and are child-friendly in repeating and memorizing; yet the way those prosodic features encoded by neural activity and their influence on speech development in children are still largely unknown. Using functional near-infrared spectroscopy (fNIRS), this study investigated the cortical responses to the perception of natural speech sentences with strong/weak-prosodic (SP/WP) features and evaluated the speech communication ability in 21 pre-lingually deaf children with cochlear implantation (CI) and 25 normal hearing (NH) children. A comprehensive evaluation of speech communication ability was conducted on all the participants to explore the potential correlations between neural activities and children's speech development. The SP information evoked right-lateralized cortical responses across a broad brain network in NH children and facilitated the early integration of linguistic information, highlighting children's neural sensitivity to natural SP sentences. In contrast, children with CI showed significantly weaker cortical activation and characteristic deficits in speech perception with SP features, suggesting hearing loss at the early age of life, causing significantly impaired sensitivity to prosodic features of sentences. Importantly, the level of neural sensitivity to SP sentences was significantly related to the speech behaviors of all children participants. These findings demonstrate the significance of speech prosodic features in children's speech development.

Keywords: natural speech perception, prosodic feature, neural response, cochlear implantation, speech communication ability, temporal cortex

INTRODUCTION

Catchy utterances, such as proverbs (i.e., "No pain, no gain" in English), verses, and nursery rhymes, contain strong-prosodic (SP) features and are child-friendly in speech repeating and memorizing (Yuzawa and Saito, 2006). Prosodic features can be recognized by the variation in pitch, loudness, and duration (Everhardt et al., 2022) and play an important role in children's speech development. Behavioral studies found successful prosody perception facilitated children's speech acquisition in that they used speech prosodic information to segment words (Jusczyk et al., 1999; Johnson and Jusczyk, 2001), discriminate emotion (Scheiner et al., 2006; Flom and Bahrick, 2007), and eliminate syntactic ambiguity (Snedeker and Yuan, 2008). Considering that childhood is the critical period of neural plasticity and that neural function development for speech prosody perception can be highly related to both biological growth and the environment (Werker and Hensch, 2015), the correlation between speech development and neural functional development for speech prosody perception in children is worth investigating.

However, neither the underlying neural mechanism of children's prosody perception nor its specific relationship with children's speech development has been studied extensively. One possible reason is that the widely used neural functional imaging technique, such as functional magnetic resonance imaging (fMRI) is noisy and highly sensitive to motion artifacts, hence is particularly not applicable for young children (Soltanlou et al., 2018). Functional near-infrared spectroscopy (fNIRS), however, is well accepted as a child-friendly optical neuroimaging technique (Quaresima et al., 2012; Saliba et al., 2016). A few fNIRS studies have explored single acoustic aspects of speech prosody for children, such as rhythm perception (Kovelman et al., 2012), intonation perception (Arimitsu et al., 2011), and prosodic emotion perception (Grossmann et al., 2010), which identified the active role of regions in the right hemisphere. Besides the limited number of studies, another limitation in previous studies is that most studies used single words or artificial utterances to test the neural processing of prosodic features. Little is done targeting directly the actual catchy utterances being used in daily life. In this study, we postulated that there is a specific neural sensitivity in normally developing children for their perception of catchy speeches, and furthermore, such sensitivity facilitates speech development as children's speech-related neural network development benefits from it. Thus, we expected that the neural responses in perceiving SP sentences would somehow associate with children's speech communication abilities.

Studies of sensory loss can provide a model for understanding the mechanism of neural function development. With respect to pre-lingually deaf children, the maturation of the auditory cortex (Knudsen, 2004; Ni et al., 2021) and their speech development (Venail et al., 2010) have been influenced due to hearing deprivation. Although cochlear implantation (CI) has become widely used to restore severe-to-profound deafness for pre-lingually deaf children (Sharma et al., 2015), the cochlear device cannot accurately deliver all kinds of sound information, especially prosodic features, possibly because of its limited number of electrodes (Svirsky, 2017). Behavior studies found CI users had difficulty in perceiving pitch changes (Gfeller et al., 2007; Jung et al., 2010), especially in the higher frequency range and discriminating intonation and tones (Peng et al., 2012), which are all related to the speech prosody. Their impaired ability to perceive low frequency pitch changes was also identified and found to be correlated with the overall speech rehabilitation outcomes. Adult CI users were also found to have lower accuracy in discriminating word stress, vowel length, compound words, or phrases (Morris et al., 2013). Nevertheless, a study found that sentences with stronger prosodic features were easier for adult CI users to understand and repeat, which implies the importance of natural catchy sentences for speech communication abilities (Aarabi et al., 2017).

By so far, little is known about the neural processing characteristics of natural catchy sentences with SP features, especially in pre-lingually deaf children after CI. Given the idea that cortical development plays an important role in speech development in pre-lingually deaf children after CI (Liang et al., 2014; Sharma et al., 2015), detecting the neural dysfunction corresponding to SP speech perception in children with CI can offer a window to identify the developmental characteristics of the neural sensitivity to speech prosody and the possible relevance to speech acquisition and shed light on speech development. We then raised the second hypothesis of this study that CI children had characteristic impairments in perceiving the strong prosodic features in catchy speeches, which probably correlates with their speech development.

This study made use of both the child-friendly and CI-safe neuroimaging technique of fNIRS to explore the neural functional characteristics in perceiving SP sentences, and the relation between neural responses and children's speech communication ability. We expected that SP sentences would induce significantly stronger activation in the right temporal area compared to a weak-prosodic (WP) sentence in normal hearing (NH) children. We also identified CI children's idiosyncratic deficits in the neural sensitivity to SP features by comparing them with the NH group. Last but not the least, we anticipated correlations between neural responses to SP features and speech communication ability for the participants, which might motivate future studies on the significance of catchy sentences in children's speech functional neural development.

MATERIALS AND METHODS

Participants

Twenty-five Chinese NH children belonging to the NH group and twenty-two Chinese children with unilateral CI in the right ear were enrolled in this study, whose native language is Mandarin Chinese. In this study, the age of the NH group ranged from 5 years and 1 month (indicated henceforth as 5; 1) to 7; 8, with an average age of 6 years, while the age of the CI group ranged from 5; 0 to 10; 9, with an average age of 6; 8 years old. We intended to explore the possible link between speech prosody perception and speech communication development. It is reported that by 5 years old, children obtain basic oral speech communication ability and begin to develop social communication skills (Eisenberg et al., 1993; Kostelnik et al., 2014). Therefore, we included children older than 5 years in this study. As for children with CI, it was reported that speech improves rapidly through the first 12–18 months after CI (Martines et al., 2013). As a result, we chose children with CI who were older than 5 years and had CI experience of more than 12 months. As children with CI might have delayed speech communication development compared to their peers, we extended the upper limit but kept the average age roughly the same.

All children with CI are pre-lingually deaf. One child with CI was excluded due to an insufficient number of completed trials. The remaining twenty-one CI subjects and all NH participants were right-handed as confirmed by the Edinburgh Handedness test (Oldfield, 1971) and had no history of neurological illness. All participants were native Chinese speakers with no neurocognitive or motor impairments and had a normal or corrected-to-normal vision. All NH participants had no known hearing problems and passed a pure tone audiometry air-conduction hearing screen performed at 0.5, 1, 2, and 4 kHz at 20 dB HL in both ears [test adapted from the British Society of Audiology (2011)]. The Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II) (Wechsler, 2011; McCrimmon and Smith, 2013) was administered to assess intelligence, and all the participants were determined to have normal intelligence. More details of participants are shown in **Table 1** and **Supplementary Appendix I**.

Parents of all participants provided written informed consent prior to the experiment. The experiments were approved by the Ethical Committee of Sun Yat-sen Memorial Hospital, Sun Yat-sen University.

Experimental Procedure

Before the start of the neural experiment, all participants were asked to respond to a set of questionnaires evaluating handedness, health state, and intelligence. CI participants additionally answered a questionnaire consisting of CI-related questions, such as the duration of deafness and the duration of cochlear device usage.

Neural Experiment

The experiments were carried out in a quiet, shielded room. The participants were instructed to sit in front of the computer screen. After wearing the fNIRS cap, the signal quality of the channel formed between optodes was tested. During the experiment, the subjects were asked to keep still, relax, and listen to the auditory stimuli carefully. They were also informed that no response was required. The experimental materials were randomly played at 75 dB. After each stimulus, there was a 15-s resting period to allow for the hemodynamic response to return to baseline (**Figure 1**). The duration of the whole experiment was approximately 15 min.

Participants were exposed to an auditory task in an event-related format, which included two types of natural utterance stimuli (see **Figure 2**). The first condition contained natural SP sentences frequently used as proverbs in daily life. The proverb sentences have an identical rhyming scheme, and each one was formed by two clauses with parallel meters (indicated by the "/" in the following examples), and stress, as well as an identical number of syllables. The ending syllable in each clause rhymes with each other. With such acoustic and prosodic features integrated into one sentence, the SP sentences all have a strong sense of rhythmic harmony. Counterpart examples in English, though not as common as in Chinese, are "No pain, no gain," "A friend in need is a friend indeed," etc. Proverbs have fixed interpretations and are commonly used in daily life to express some sort of principles, knowledge, beliefs, etc. Here is an example of the usage of the proverb in Example (1): "You must practice more for this race, since 'one minute on stage takes ten years of practice.'"

(1)	*tai*	*shang/*	*san/*	*fen zhong,*
	t^{h}ai^{35}	ʂɑŋ^{51}	san^{55}	fən^{55} tʂuŋ^{55},
	on	*stage/*	*three/*	*minutes,*
	tai xia/	*shi/*	*nian gong*	
	t^{h}ai^{35} ɕiA^{51}	ʂʅ^{35}	niæn^{35} kuŋ^{55}	
	off the stage/	*ten/*	*years of practice*	
	'One minute on stage takes ten years of practice.'			

(2)	*ta/*	*mei tian/*	*dou/hui/*	*qu/*
	t^{h}A^{55}	mei^{214} t^{h}iæn^{55}	tou^{55} xuei^{51}	tɕ^{h}y^{51}
	he/	*every day/*	*always/*	*go/*
	da	*cai shi/*	*mai*	*cai*
	tA^{51}	ʦ^{h}ai^{51}ʂʅ^{51}	mai^{214} ʦ^{h}ai^{51}	
	big	*market/*	*get groceries*	
	'He goes to the big market to get groceries every day.'			

The second condition contained WP sentences, which are commonly used in daily conversations. The well-designed prosodic features existing in the SP were avoided in the WP condition, for instance, parallel meters and stress [refer to an example (2)]. Prosodic features in two types of stimuli were illustrated in the following figure of the waveform. Acoustic records of all the materials are available on https://pan.baidu.com/s/1vgiYoU_XqVd4vWRLSjA0VA?pwd=nirs.

No significant difference was found between the two groups of stimuli for the word frequency (mean: 1,820.43 vs. 2,094.60

TABLE 1 | Demographic information of participants.

Participants	Gender		Age (Range)	Implantation age (Range)	Duration of CI (Range)
	Male	Female			
NH	12	13	6; 0 (5; 1–7; 8)	\	\
CI	10	11	6; 8 (5; 0–10; 9)	3; 1 (1; 5–9; 7)	3; 7 (1; 2–5; 6)

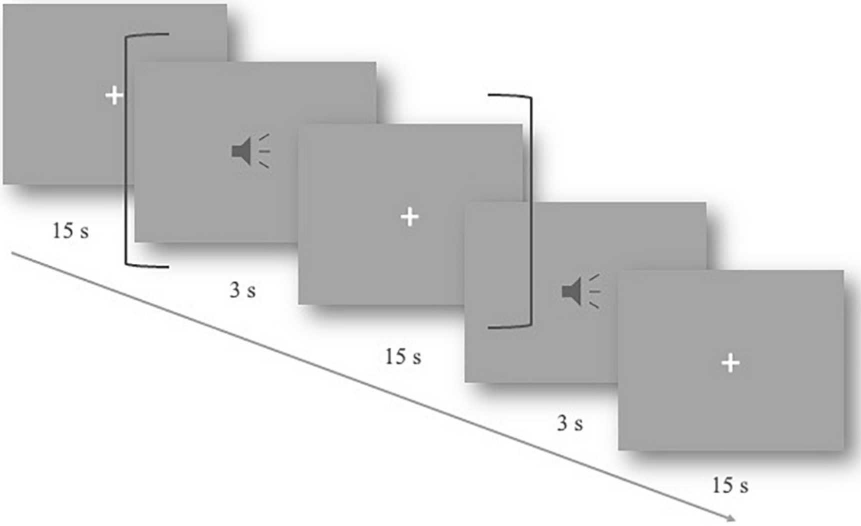

FIGURE 1 | Slow event-related experimental design. One trial lasted 18 s, including 3 s of the auditory stimulus and a 15-s resting period (shown by dark gray brackets). The total procedure includes 40 trials.

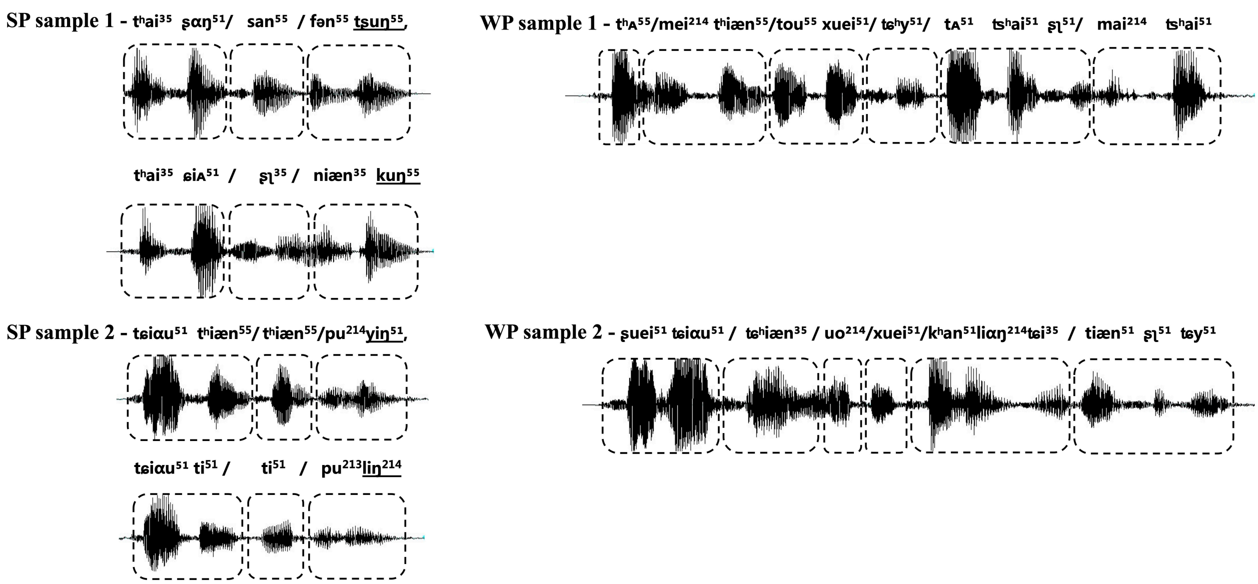

FIGURE 2 | Waveform of samples from two types of experimental stimuli, with strong-prosodic (SP) in the left column and weak-prosodic (WP) in the right one. The boxes with dashed lines indicate the meters of the sentence. The shapes of sound waves show the parallel construction of rhythm and stress within each SP, together with parallel meters as indicated by the boxes with dashed lines; the ending syllables in the two parallel clauses rhyme with each other, as shown by the underlined international phonetic signs in SP samples.

pre-million, $p = 0.20$, based on Chinese Word Frequency Corpus, BLCU) to make sure every character occurs frequently in the daily life. Furthermore, 10 Mandarin Chinese-speaking university students and 10 children at the age of six were invited to grade the frequency of utterances heard and said in daily life by using a 5-point scale for this familiarity rating. The average scores of stimuli with and without SP features rated by university students were 4.90 and 4.80, respectively, while the average scores of children were 3.20 and 3.50, respectively. We also controlled the grammar of the sentences. The two groups of sentences shared five commonly used syntactic composition rules in parallel, including subject-predicate, verb-object, modifier-head, passiveness, and coordination. In addition, the semantics of sentences in both conditions were easy to retrieve due to their common usage.

Both conditions contained 20 trials of sentences. The experimental materials were recorded by a professional male announcer in standard Mandarin. Due to the slow event-related experimental design (for detailed denotations and explanations,

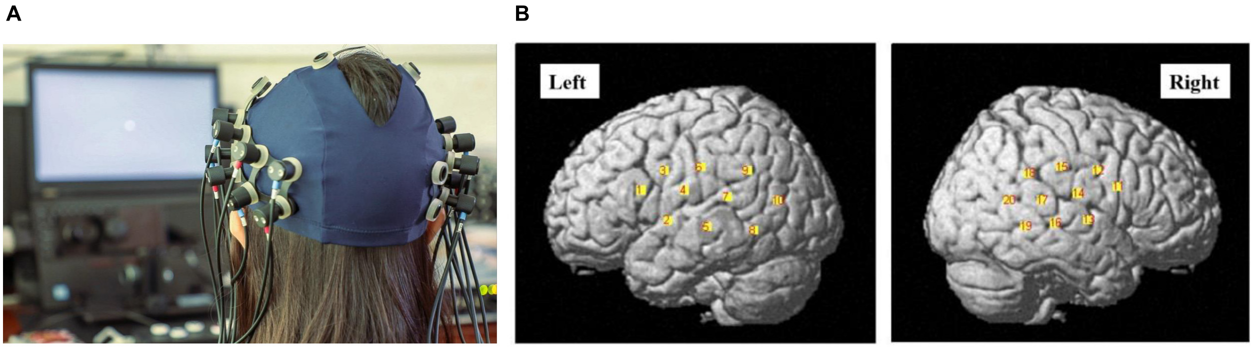

FIGURE 3 | (A) The photograph of the optode array placed on the head of one of the participants (with consent). Red and blue labels on the optodes indicate optical sources and detectors, respectively. **(B)** The channel positions following registration to a brain atlas with different numbers (channels 1–10 left-hemisphere, channels 11–20 right-hemisphere).

refer to Aarabi et al., 2017) with longer intervals (15 s) between stimuli, there was no strict requirement for the duration of stimulation materials to be accurate to milliseconds. The audio was edited by Audacity to cut the blank segments, ensuring that the duration of the stimuli was 2–3 s, about 2.5 s on average. Only 100–500 ms of blanks in long pauses of the utterances caused by the announcer were cut so that the naturality of the speech sentences was well kept.

Functional Near-Infrared Spectroscopy Measurements

Measurements were carried out with a total of 16 optodes arranged in two 2 × 4 arrays (each containing four sources and four detectors). The distance between the source and detector was set at 3 cm, and the optodes were positioned crosswise from each other. Hemodynamic responses were measured at the midpoints between the source and detectors, which were called "NIRS channels," by the fNIRS imaging system (LIGHTNIRS; Shimadzu Co., Ltd., Tokyo, Japan). The arrays were placed on both sides of the head (**Figure 3A**), aiming primarily to measure cortical activation in the bilateral temporal cortex, inferior frontal cortex, and inferior parietal cortex.

Three different wavelengths (780, 805, and 830 nm), each with a pulse width of 5 ms, were used to calculate hemodynamic responses. The details of the head cap and the systems were described previously (Takeuchi et al., 2009; Takamoto et al., 2010; Takakura et al., 2011). After the recording, the three-dimensional (3D) locations of the optodes were measured by a 3D Digitizer (Nirtrack; Shimadzu Co., Ltd.) in reference to the bilateral tragus, nasion, and Cz. The measurement was performed by the "Spatial registration of NIRS channel locations" function of the NIRS-SPM Version 3 software, which is a SPM8- and MATLAB-based software package for the statistical analysis of NIRS signals (Ye et al., 2009). The channel positions are shown in **Figure 3B**.

Speech Communication Ability Evaluation

The speech communication ability of participants was tested following Zhang et al. (2020)'s speech elicitation procedure. The procedure included three types of tasks: picture description, video content statement, and free conversation. In the 15-min test, children were asked to describe the content of two pictures, two videos, and freely talk about familiar personal experiences (e.g., family situations, favorite games, or cartoons).

Each participant's test performance was scored by a specifically designed deep-learning framework (Zhang et al., 2020), which gave an evaluation of five major linguistic aspects of speech and language: pronunciation, expression efficiency, fluency, grammar, and semantics (**Table 2**). The overall speech communication ability score was calculated by averaging the five scores of their different linguistic aspects. The detailed scoring method is described in **Supplementary Appendix II**.

Data Analysis

As Beta weight was reported to be a better parameter reflecting the activation level of certain cortex area to specific stimuli (Plichta et al., 2007; Ye et al., 2009) while oxygen-hemoglobin concentrations (HbO) provided information about how the

TABLE 2 | Speech communication ability (SCA) evaluation scores on three levels.

Overall SCA	Scores on major linguistic aspects	Scores on the finest level of observation
SCA total score		
	PRONUNCIATION	Initial consonants accuracy
		Vowel accuracy
		Tone accuracy
	EXPRESSION EFFICIENCY	Syllable count
		Speech speed
		Pronunciation duration
	FLUENCY	Content restatement/replication
		Redundant articles
		Pause count
		Pause duration
	GRAMMAR	The wrong usage of grammar
	SEMANTICS	Key words missing

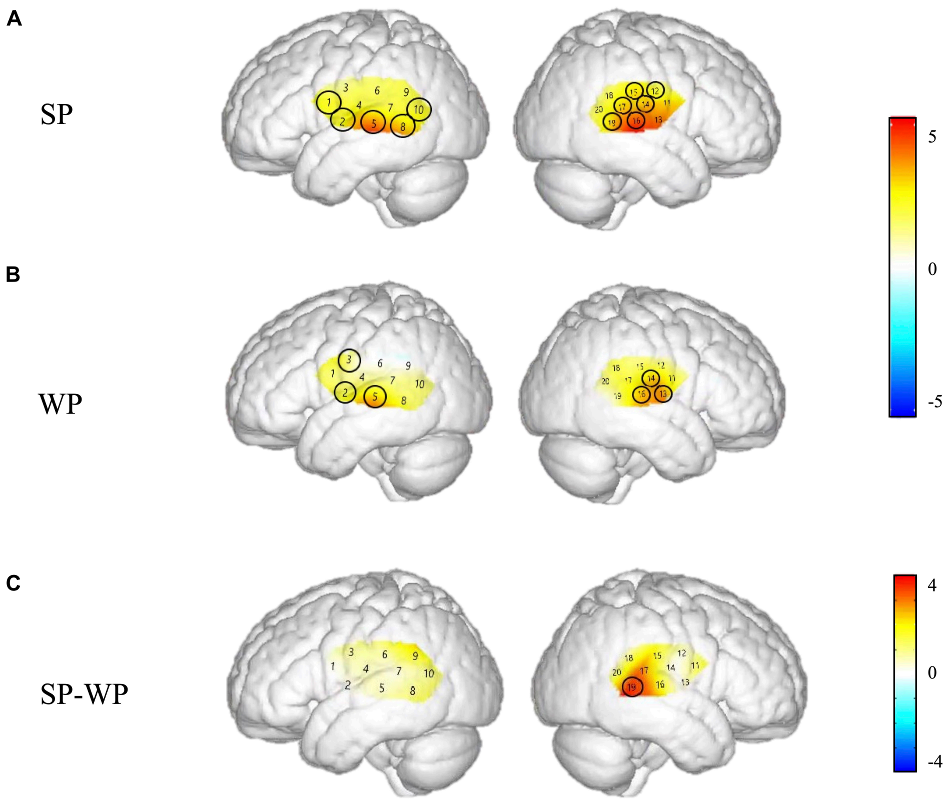

FIGURE 4 | Cortical activation maps of the normal hearing (NH) group on the normalized brain surface. The color scales represent the *T* values, with statistically significant activated channels circled in black [$p_{FRD\text{-}cor} < 0.05$]. **(A,B)** Lateral views of significantly activated channels under strong-prosodic (SP) and weak-prosodic (WP) conditions contrasted against silence, respectively. **(C)** Lateral views of significantly activated channels by contrasting between conditions, indicating the stronger activated neural areas driven by the SP features in SP sentences.

activation level changed with time, we used both parameters to investigate activation patterns of SP stimuli. To explore how activation patterns differ for SP vs. WP stimuli in children with NH (the first research question), we compared the Beta weights, HbO concentrations, and functional connectivities between SP and WP conditions in the NH group. To verify our second hypothesis related to the abnormal pattern of prosodic perception in the CI group, comparisons between groups were conducted using Beta weights, HbO concentrations, functional connectivities, and laterality in response to SP and WP conditions. Furthermore, to investigate the relationship between neural responses to SP features and SCA (the third research question), we calculated the correlations between Beta weights of both conditions and SCA scores based on Pearson's correlation.

Beta Weights Analysis Based on the General Linear Model

The present study focused on the changes in the oxygen–hemoglobin concentration, which has been reported to be sensitive to neuro-hemodynamic relationships (Hoshi et al., 2001; Strangman et al., 2002; Yamamoto and Kato, 2002).

Based on the general linear model (GLM), NIRS-SPM version 3 was used for analysis along with SPM 8 (Plichta et al., 2007; Ye et al., 2009). The time course of HbO was correlated with the design matrix using a boxcar function hemodynamic response function (hrf) to explain the data as a linear combination of an explanatory variable (Beta weight) plus an error term. After preprocessing with DCT detrending, hrf low-pass filtering, 128 Hz high-pass filtering, and Beta weights were obtained using an ordinary least square fit. More details were described by Ye et al. (2009).

Spatial activation maps of two groups under each condition were created using Beta weights and three contrast-model matrices in NIRS-SPM. The three contrast-model matrices were as follows: (1) SP condition [1 0 0] to examine HbO concentrations related to SP perception contrasting to the baseline, (2) WP condition [0 1 0] to examine HbO concentrations related to WP perception contrasting to the baseline, and (3) SP vs. WP condition [1 –1 0] to remove HbO concentrations common to both conditions and the non-task-related activation, which can be interpreted as the background of the brain activities.

To explore how activation patterns differ for SP vs. WP stimuli in NH children, one sample *t*-test was used to create spatial maps of significant activation at the group level for each of these three conditions in the NH group as shown in **Figure 4** (SP and WP conditions) and one-way ANOVA was calculated as shown in **Figure 5** in two groups. We conduct the same analysis for the

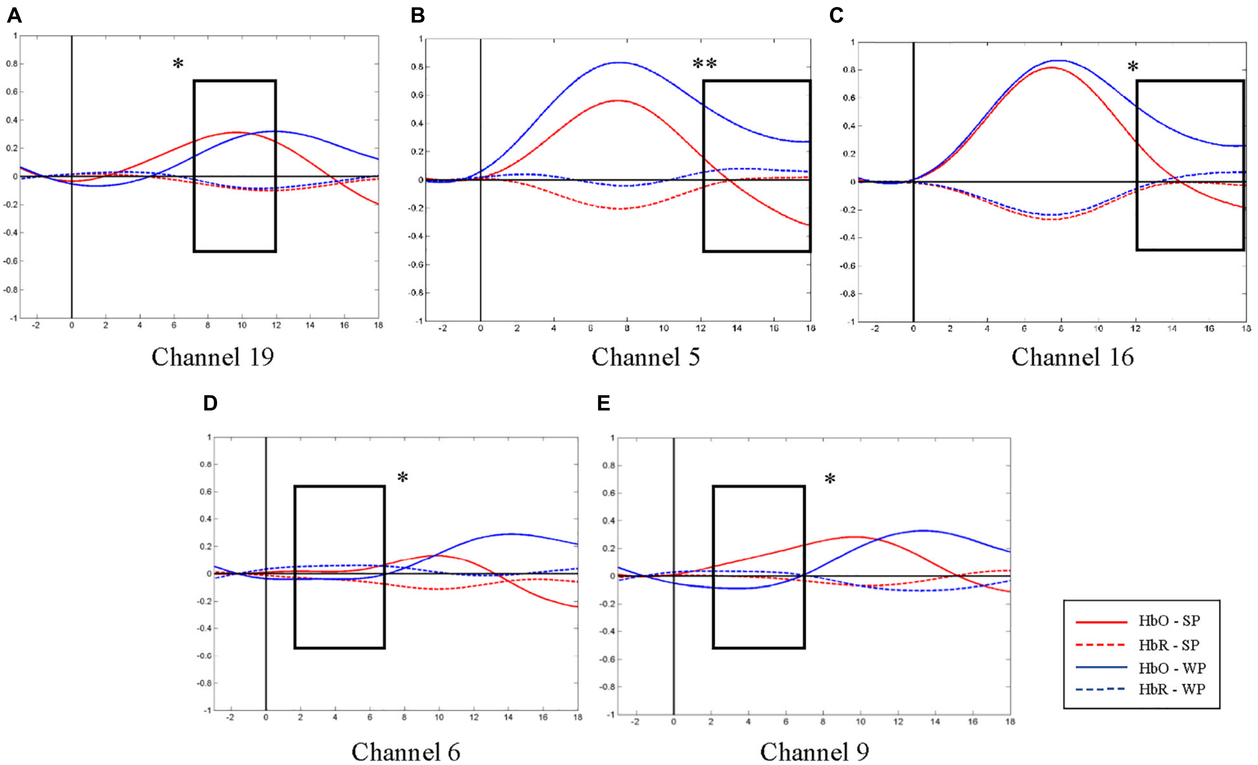

FIGURE 5 | Average oxygen-hemoglobin concentrations (HbO) envelopes of the normal hearing (NH) group in the chosen time windows. **(A)** Channel 19 (the right middle temporal gyrus) had more significant activation in 7–12 s under strong-prosodic (SP) condition [$F(1,48) = 4.086$, $p_{FRD\text{-}cor} = 0.049$]. **(B,C)** Weak-prosodic (WP) stimuli evoked more significant activation in channels 5 [$F(1,48) = 9.543$, $p_{FRD\text{-}cor} = 0.003$] and 16 [$F(1,48) = 4.398$, $p_{FRD\text{-}cor} = 0.041$] in 12–18 s. **(D,E)** Channels 6 [$F(1,48) = 4.358$, $p_{FRD\text{-}cor} = 0.042$] and 9 [$F(1,48) = 4.636$, $p_{FRD\text{-}cor} = 0.036$] showed significant stronger responses to SP. $^*p < 0.05$ and $^{**}p < 0.01$.

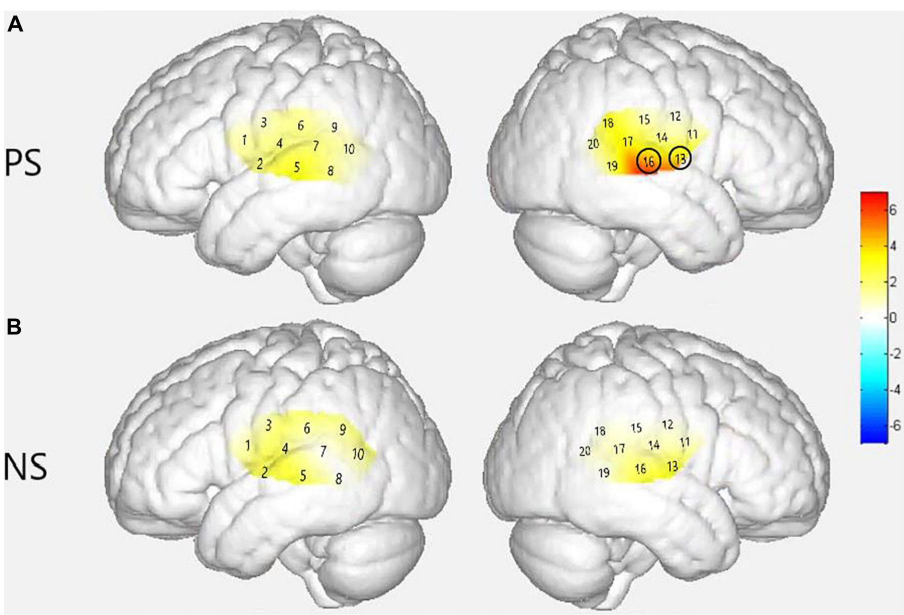

FIGURE 6 | Group level cortical activation maps by contrasting between groups in the views of the normalized brain surface. The color scales represent the *F* values, with statistically significant activated channels circled in black [$p_{FRD\text{-}cor} < 0.05$]. **(A)** Lateral views of significantly activated channels in normal hearing (NH) group under the strong-prosodic (SP) condition than CI. **(B)** Lateral views of activating difference under the weak-prosodic (WP) condition by contrasting two groups of participants.

CI group and contrast activation map of both groups to give a general impress of the abnormal pattern of prosodic perception in the CI group. To further explore the activating difference between groups, one-way ANOVA was conducted under SP and WP conditions, respectively (**Figure 6**). Threshold images for the resulting group data were created using an alpha of 0.05. The value of q specifying the maximum false discovery rate (FDR) was set at 0.05. Areas of significant task-based activity are described in terms of cortical regions as opposed to cortical structures. The Beta values and the statistical results were visualized on spatial maps of the standard head model to indicate channels with significant results.

Event-Related Oxygen-Hemoglobin Concentrations

Event-related HbO concentrations provided information about how the activation level changed with time right after the stimulus began. Analysis of the data was performed in MATLAB (MATHWORKS, NATICK, and MA) using functions in the HOMER2 package (Huppert et al., 2009) together with custom scripts. The analysis included the following steps:

1. Exclusion of channels influenced by the external CI device. The hemodynamic responses of related channels could not be observed appropriately because the position of optodes overlapped with that of several participants' external CI devices. Refer to **Table 3** below for specific exclusions.

However, other effective data of the channel gained from other subjects would still be retained. The data retention rate of all the channels was up to 98.6%, which did not affect the effectiveness of the overall data.

2. Conversion to optical density. The measured light intensity levels were converted to optical density using the HOMER2 *hmrIntensity2OD* function, a standard step in fNIRS data analysis (Huppert et al., 2009).
3. Motion-artifact correction. Motion artifacts were suppressed using the HOMER2 *hmrMotionArtifact* function.
4. Bandpass filtering. The optical density signals were bandpass filtered between 0.01 and 0.1 Hz by the *hmrBandpassFilt* function to attenuate low-frequency drift and cardiac oscillation.
5. Conversion to estimated changes in hemoglobin concentrations. Optical density was converted to estimated changes in the concentrations of HbO and concentration changes of deoxyhemoglobin (HbR) through the application of the modified Beer-Lambert Law (Huppert et al., 2009). A default value of 6.0 was used for the differential path-length factor at both wavelengths.

To gain the fine-grained observation, data were averaged further within four-time windows to make sure that the duration of every segment was the same as 5–6 s due to the slow change of hemodynamic responses, namely, −3–2, 2–7, 7–12, and 12–18 s. The starting point was set before the stimuli to examine the possible top-down control of attention and the anticipating effect. Among them, the first interval was set before and through the trial onset. This time window was regarded as a relatively stable phase right before any neural activity starts to respond according to hemodynamics (cf. Plichta et al., 2007).

To explore how activation patterns differ for SP vs. WP stimuli in NH children, HbO concentrations of NH group in different time windows were analyzed by one-way ANOVA with IBM SPSS Statistic Version 22.0 (IBM Corporation, Armonk, NY, United States). The Greenhouse–Geisser adjustment to the degrees of freedom was applied to all ANOVAs to correct for the violation of the assumption of sphericity. When significant interactions were found, *post hoc* tests were performed using tests for the simple effect of one-factorial ANOVA and/or Fisher's protected least significant difference test. The level of statistical significance was set at $p < 0.05$. The value of q specifying the maximum FDR was set at 0.05 to make sure the false positive rate was no more than 5% on average in processing the HbO from multiple channels.

To explore our second hypothesis related to the abnormal pattern of prosodic perception in the CI group, we conduct the same analysis in the CI group. As we assumed that the CI group had some deficit in perceiving prosody features, we expected to find that significant differences between SP and WP conditions found in the NH group would not appear in the CI group. Then, we used one-way ANOVA to explore between group differences in different time windows under both conditions.

Furthermore, neural processing of speech and language has been widely found to have effects of lateralization (Belin et al., 1998; Gandour et al., 2004). To explore if the SP perception function is allocated to a specific hemisphere (related to the first research question), we calculated the laterality between the symmetrically matched brain regions in the left and right hemispheres (viewed as a channel pair) for the same condition between the two groups. The laterality index was calculated by the formula (Coito et al., 2015) below. Then, one-way ANOVA was employed to investigate between condition differences in different time windows in two groups.

$$Laterality = \frac{HbO\ (Left) - HbO\ (Right)}{HbO\ (Whole)}$$

Analysis of Functional Connectivity

Functional connectivity (FC) was analyzed to evaluate participants' cortical network activity while listening to utterances. The brain regions do not process information in an isolated way but work as a network to accomplish certain functions. Hence, FC can indicate the network patterns that possibly differ between conditions and groups.

TABLE 3 | The excluded channels.

Participants	Excluded channels
Subj_03	18
Subj_10	16, 17, 19
Subj_17	19, 20

All channels were included in the FC analysis. A bandpass Fourier filter (0.01–0.1 Hz) in the time series of HbO signals was used, and then the time series was further separated into segments (3 s before the stimuli and 18 s after the stimuli). Each condition (SP and WP) contained 20 segments. Each trial was separated into two time windows (0–9 s and 9–18 s) and the correlation was calculated for each window. We calculated mean coefficients for each pair of channels in each time window among all subjects. Group mean coefficients greater than 0.6 were mapped as significant FCs in each condition, considering that the correlation coefficient smaller than 0.4 increases false-positive rates and the coefficient greater than 0.7 results in FC density maps with lower sensitivity because of reduced dynamic range (Tomasi and Volkow, 2010; Wang et al., 2018).

For statistical comparison of the FCs among the regions with group mean coefficients greater than 0.6 in each condition, three-way repeated measures ANOVA (condition × region pair × group, hereafter RM-ANOVA) was performed in the 0–9 s time window and 9–18 s time window.

Statistical Analysis Between Hemodynamic Responses and Speech Behaviors

To explore the relationship between neural responses to SP/WP features and SCA development, we calculated the correlations between Beta weights of both conditions and speech communication ability scores based on the Pearson correlation coefficient. To determine whether there was a significant relationship in general, we examined the data of both groups as a continuum. This is due to the fact that all children, whether or not they have CIs, have different speech abilities.

RESULTS

Neural Activities Based on Functional Near-Infrared Spectroscopy

Sensitive Neural Responses to Strong-Prosodic Stimuli

To explore if there is any sensitivity or selectivity of neural responses to SP speech, we compared the Beta weights, HbO concentrations, and functional connectivities between SP and WP conditions in the NH group.

The Beta weights were calculated to identify the brain regions that were engaged in processing SP and WP sentences. Channels 1, 2, 5, 8, 10, 11, 13, 14, 15, 16, 17, and 19 responded to the SP condition significantly while the WP condition yielded activation in channels 2, 3, 5, 13, 14, and 16 (refer to detailed statistical results in **Supplementary Appendix III**), which showed that SP activated a broader scope of the brain areas than WP in both hemispheres.

Figure 4 demonstrates the cortical activation for each stimulation condition at the group level. In the NH group, areas including the left superior temporal gyrus, middle temporal gyrus, fusiform gyrus, angular gyrus and right superior temporal gyrus, middle temporal gyrus, supramarginal gyrus, and subcentral area, responded to SP (**Figure 4A**), namely, channels 1, 8, 10, 11, 15, 17, and 19.

Under WP, the NH group had more responsive areas in the left primary association cortex (channel 13), right primary auditory cortex, superior temporal gyrus, and middle temporal gyrus (channels 14 and 16), while the two groups shared the same activating areas of the bilateral middle temporal gyrus (**Figures 4B, 7B**).

As shown in **Figure 4C**, no significant main effect of conditions was found in any channel in the NH group (see details in **Supplementary Table 3**), but a marginally significant stronger response existed in channel 19 (right middle temporal gyrus) [$F(1,48) = 3.917$, $p_{FRD\text{-}cor} = 0.054$, $\eta^2 = 0.075$], highlighting the sensitivity of the right middle temporal gyrus on processing SP sentences.

Furthermore, to characterize the HbO concentration in the time course under two conditions, we examined the difference within four time windows and the laterality in the NH group. A higher HbO concentration was observed under the SP condition than under the WP condition in earlier time windows: 2–7 s (channels 6 and 9, **Figures 5D,E**) and 7–12 s (channel 19, **Figure 5A**). Among these, the right middle temporal gyrus (channel 19) once again demonstrated its specialization in SP processing. In the later time window of 12–18 s, a higher activation under the WP condition was observed than under the SP condition in channels 5 and 16 (**Figures 5B,C**).

Laterality results showed that the cerebral activation was significantly right-lateralized when processing SP sentences in channel pair 9–18 in −3–2 s [$F(1,48) = 1.324$, $p_{FRD\text{-}cor} = 0.04$] and in channel pair 8–19 in 7–12 s [$F(1,48) = 0.994$, $p_{FRD\text{-}cor} = 0.001$], providing more evidence for the right-lateralization of SP processing in NH group.

Additionally, to test the cross-brain cooperation between different regions, we looked at the FC of two conditions in the NH group. **Figure 7** illustrates pairs of regions with FC values greater than a threshold ($r > 0.6$) in each condition. For the NH group, there were high FCs throughout the trials among channels 17, 18, 19, and 20, which represented cortical regions around the right superior temporal gyrus, the middle temporal gyrus, the supramarginal gyrus, and the angular gyrus. Among them, the right middle temporal gyrus (channel 19) was involved in the broadest FCs. More connectivity results were found under the SP condition than under the WP condition (6 pairs vs. 1 pair), suggesting that SP processing required stronger FC between cortical regions which mainly occurred in the right hemisphere. Only one connection between the LH and RH was found under the SP condition but was not found in the WP condition during the first half of the trials.

Abnormal Neural Response Patterns of Cochlear Implantation Group

To verify our second hypothesis related to the abnormal pattern of prosodic perception in the CI group, we first conducted the same analysis in the Section “Sensitive Neural Responses to Strong-Prosodic Stimuli” for the CI group to see if what we found in the NH group would also be found in the CI group. We conducted direct comparisons between groups using Beta weights and HbO concentrations in response to SP and WP conditions.

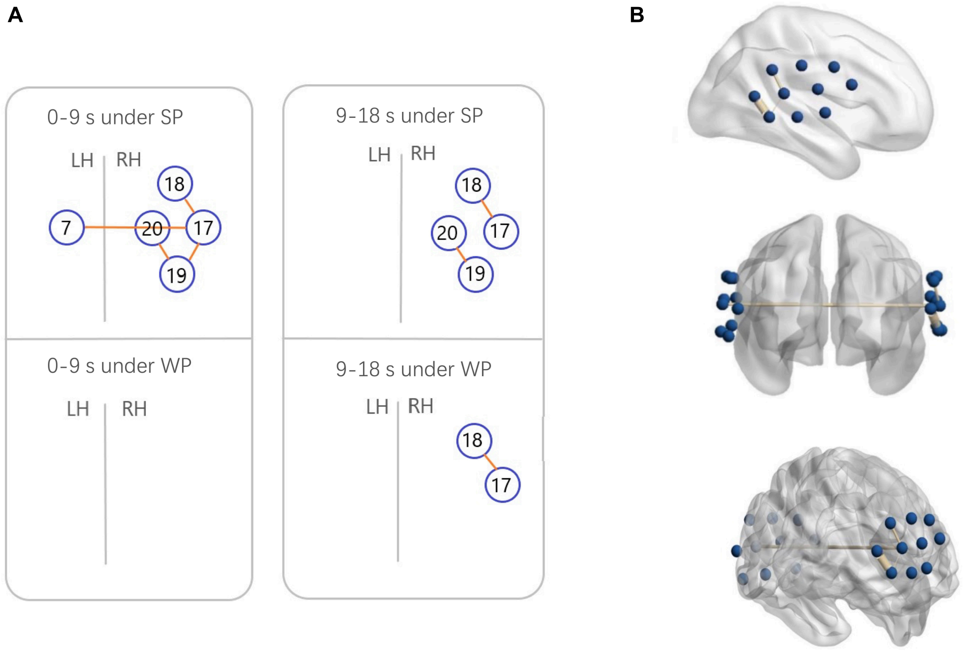

FIGURE 7 | Functional connectivity (FC) maps among the activated cortical regions in the normal hearing (NH) group. **(A)** Blue circles indicate channels, and the orange lines show correlations greater than 0.6 (average across all participants). A significant correlation was found in Ch7-17 [r = 0.6231], Ch17-18 [r = 0.6222], Ch17-19 [r = 0.6194], and Ch19-20 [r = 0.6304] under strong-prosodic (SP) condition in time-window of 0–9 s. Under SP condition in time-windows of 9–18 s, significant correlation was found in Ch17-18 [r = 0.6561], Ch19-20 [r = 0.6088]. Correlation under weak-prosodic (WP) condition in time windows of 9–18 s was found in Ch17-18 [r = 0.6215]. **(B)** FC in the NH group under SP condition on the standard brain in time-windows of 0–9 s in different views. The blue dots indicate 20 channels, and the yellow lines show correlations greater than 0.6 (average across all participants).

As shown in **Figure 8**, between-condition contrast of Beta weights was also found in the CI group, with larger areas responding to SP than WP in channels 4, 8, 16, and 18 (refer to detailed statistical results in **Supplementary Appendix III**). In comparison, the CI group had a limited number of activated channels under the SP condition located only in a small area of the left superior temporal gyrus, the middle temporal gyrus, the subcentral area, the fusiform gyrus and the right superior temporal gyrus, the middle temporal gyrus, and the supramarginal gyrus (**Figure 8A**). As for between condition comparison in the CI group, no significant difference was observed (**Figure 8C**, refer to **Supplementary Appendix III**). Contrary to the different activation for the two conditions in the NH group, no significant main effect of condition was found within the CI group for HbO concentration. These results suggested that the CI group lacks neural sensitivity to process speech prosodic contrast.

Figure 6 demonstrated that responses between group comparison were significantly larger in channels 13 [$F(1,44) = 5.259$, $p_{FRD\text{-}cor} = 0.027$] and 16 [$F(1,44) = 6.861$, $p_{FRD\text{-}cor} = 0.012$] in the NH group compared with the CI group under the SP condition. There was no significant difference between the two groups under WP condition.

A similar between-group contrast was also found in HbO concentration. Under SP condition, the NH group generally had higher HbO concentration than the CI group throughout the time window of 2–12 s in channels 5, 14, and 16 (the left superior temporal gyrus, left middle temporal gyrus, right primary and association cortex, right superior temporal gyrus, and the right middle temporal gyrus) and the latest time window of 12–18 s in channel 8 (left middle temporal gyrus and fusiform gyrus) as shown in **Figure 9A**.

In addition, under WP condition, HbO concentration in the NH group was generally higher during the 2–12 s window (**Figure 9B**). However, there was no significant difference in −3–2 and 12–18 s time windows, but there were significantly different HbO concentrations in channels where the NH group always elicited stronger activation, including the left superior temporal gyrus, middle temporal gyrus, fusiform gyrus, supramarginal gyrus, and the right middle temporal gyrus (channels 5, 7, 8, and 13). It may be interpreted that children with CI had close-to-normal cortical processing for WP perception in the late phase of speech processing.

To analyze the statistical significance of the FC among the two conditions and the pairs of regions greater than the threshold, three-way RM-ANOVA (2 condition × 4 region pair × 2 group) was performed in the 0–9 s time window. A significant group effect was found [$F(1,44) = 39.891$, $p < 0.001$, $\eta^2 = 0.476$]. However, no significant main effect {condition [$F(1,44) = 2.552$, $p = 0.117$, $\eta^2 = 0.055$], region [$F(3,42) = 0.770$, $p = 0.517$, $\eta^2 = 0.052$]} or interaction effect {group × condition [$F(3,42) = 2.258$, $p = 0.140$, $\eta^2 = 0.049$], group × region [$F(3,42) = 0.170$, $p = 0.916$, $\eta^2 = 0.012$], condition × region [$F(3,42) = 0.558$, $p = 0.646$, $\eta^2 = 0.038$],

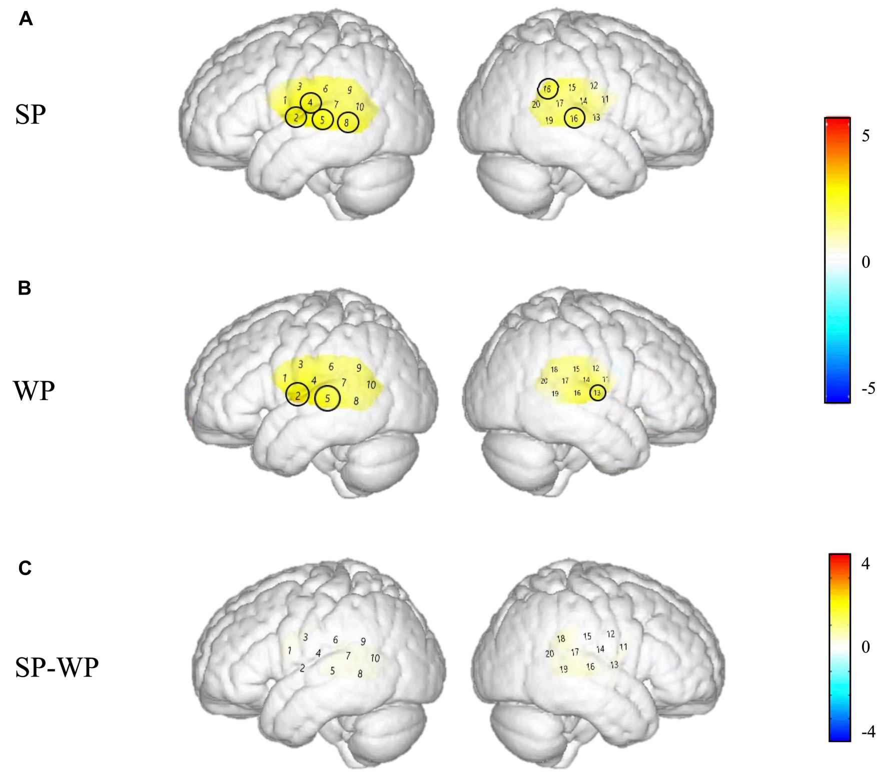

FIGURE 8 | Cortical activation maps of the cochlear implantation (CI) group on the normalized brain surface. The color scales represent the *T* values, with statistically significant activated channels circled in black [$p_{FRD\text{-}cor} < 0.05$]. **(A,B)** Lateral views of significantly activated channels under strong-prosodic (SP) and weak-prosodic (WP) conditions contrasted against silence, respectively. **(C)** Lateral views of significantly activated channels by contrasting between conditions, indicating the stronger activated neural areas driven by the SP features in SP sentences.

and group × condition × region [$F(3,42) = 1.906$, $p = 0.143$, $\eta^2 = 0.120$]} was found.

In addition, three-way RM-ANOVA (2 condition × 4 region pair × 2 group) was performed for the 9–18 s time window. A significant group effect was found [$F(1,44) = 36.77$, $p < 0.001$, $\eta^2 = 0.455$]. No significant main effect {condition [$F(1,44) = 1.716$, $p = 0.197$, $\eta^2 = 0.038$], region [$F(3,42) = 1.697$, $p = 0.186$, $\eta^2 = 0.107$]} or interaction effect {group × condition [$F(1,44) = 0.505$, $p = 0.481$, $\eta^2 = 0.011$], group × region [$F(3,42) = 0.180$, $p = 0.909$, $\eta^2 = 0.013$], condition × region [$F(3,42) = 0.960$, $p = 0.421$, $\eta^2 = 0.046$], and group × condition × region [$F(3,42) = 0.033$, $p = 0.962$, $\eta^2 = 0.007$]} was found.

Neuro-Behavioral Correlation

Speech Behavior Assessment

To address the hypothesis that neural responses to SP features are closely related to children's speech communication ability, we probed all the participants' SCA development by a behavioral test. Since abnormal neural activities were found in children with CI, we expected deficits to be found in their SCA assessment results.

All child participants in this study completed the SCA assessment, and significant differences were found at multiple observation levels of the evaluation (**Figure 10**). Children with CI were severely impaired in their speech development, with a significantly lower SCA total score, compared to their peers [$F(1,38) = 0.895$, $p < 0.001$]. The aspects of pronunciation [$F(1,38) = 10.013$, $p < 0.001$], semantics [$F(1,38) = 5.706$, $p < 0.001$], and expression efficiency [$F(1,38) = 0.758$, $p < 0.001$] had the most significant differences among the five levels of linguistic aspects. Similar deficits were also found in previous studies (Peng et al., 2004; Wu et al., 2008; Chen et al., 2017).

Neuro-Behavioral Correlation

We associated the scores of participants' comprehensive speech communication ability with their hemodynamic performance represented by the Beta weights to further explore the connection between children's neural activation of speech perception and their SCA. To determine whether there was a significant relationship in general, we examined the data of both groups as a continuum. The results of significant correlations under SP and WP conditions are shown in **Tables 4**, **5**, respectively.

More correlations were found under the SP condition than under the WP condition (14 vs. 4), suggesting that the perception of SP was more closely connected to SCA than the perception of WP. The overall level of SCA was predicted by cortical activation in channel 5 and channel 16 under both conditions. Activation in channel 19 (the right middle temporal gyrus) was sensitive to the largest range of evaluation scores (including the

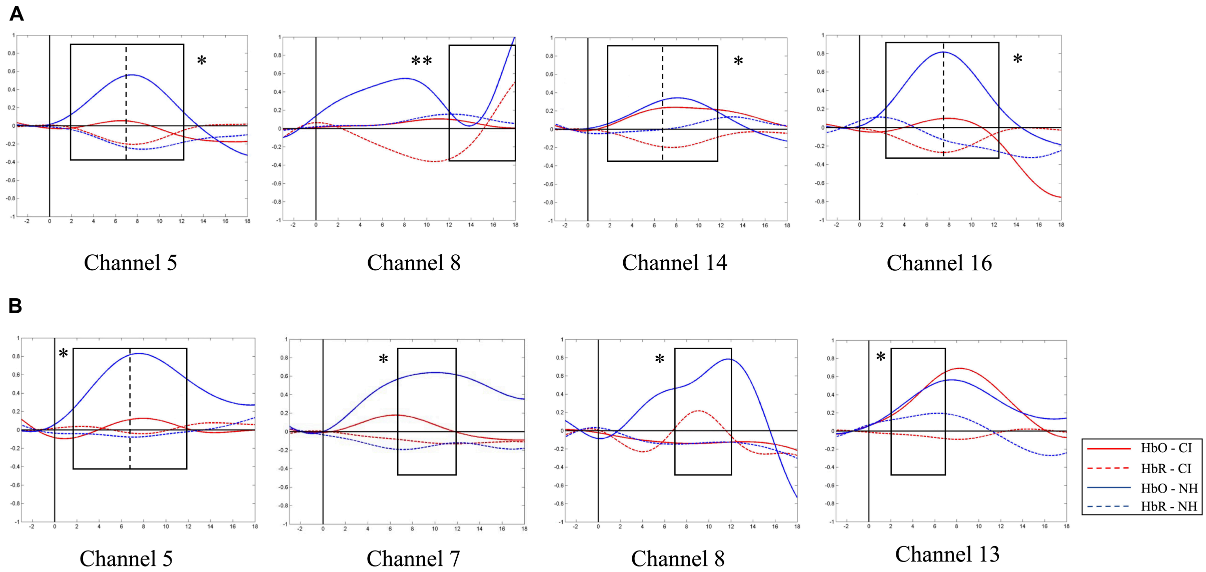

FIGURE 9 | Between-group contrast of average oxygen-hemoglobin concentrations (HbO) envelopes under strong-prosodic (SP) and weak-prosodic (WP) conditions. **(A)** Hemodynamic responses to SP by both groups in channels 5, 8, 14, and 16. Higher activation by normal hearing (NH) group was found in time window 2–7 s in Ch5 [$F(1,44) = 5.186$, $p_{FRD\text{-}cor} = 0.028$], Ch14 [$F(1,44) = 9.367$, $p_{FRD\text{-}cor} = 0.004$], and Ch16 [$F(1,44) = 10.162$, $p_{FRD\text{-}cor} = 0.003$]; time window 7–12 s in Ch5 [$F(1,44) = 5.084$, $p_{FRD\text{-}cor} = 0.029$], Ch14 [$F(1,44) = 9.736$, $p_{FRD\text{-}cor} = 0.003$], and Ch16 [$F(1,44) = 11.506$, $p_{FRD\text{-}cor} = 0.001$]; and time window 12–18 s in Ch8 [$F(1,44) = 8.994$, $p_{FRD\text{-}cor} = 0.004$]. **(B)** Under WP, higher activation by NH group was found in time window 2–7 s in Ch5 [$F(1,44) = 5.743$, $p_{FRD\text{-}cor} = 0.021$], Ch13 [$F(1,44) = 5.185$, $p_{FRD\text{-}cor} = 0.028$]; and time window 7–12 s in Ch5 [$F(1,44) = 4.483$, $p_{FRD\text{-}cor} = 0.04$], Ch7 [$F(1,44) = 5.932$, $p_{FRD\text{-}cor} = 0.019$], and Ch8 [$F(1,44) = 6.606$, $p_{FRD\text{-}cor} = 0.014$].

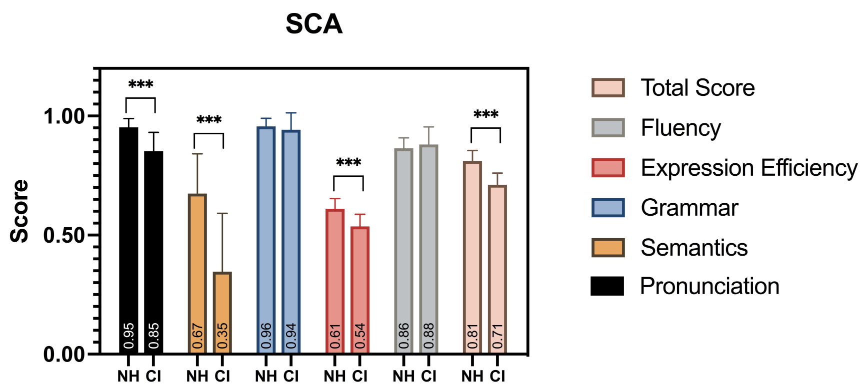

FIGURE 10 | SCA assessment results. Between groups, significant differences in scores were found in total score, pronunciation, semantics, and expression efficiency. ***$p < 0.001$.

total score under the SP condition), which further verified its sensitivity to children's speech development. The fluency score was related to the cortical activation data of the three channels under the SP condition, but no significant result was found under the WP condition.

DISCUSSION

We characterized the cortical responses to the perception of speech sentences with contrasting prosodic features in NH children and children with CI. In NH children, SP sentences evoked broader and right-lateralized cortical responses compared to WP stimuli. To our knowledge, this is the first time that the neural mechanism underlying the perception of natural catchy speech was investigated among children.

In addition, we identified deficits in cortical activation of SP speech perception in children with CI and examined the relationship between their speech behavioral performance and cortical responses. Consistent with previous findings (Luo and Fu, 2007), we observed significantly weaker activation in response to speech perception in the CI group. Moreover, more inhibited neural activities were found under SP than WP conditions for the CI group, suggesting that children with CI had a

TABLE 4 | Correlation between neural activation to strong-prosodic (SP) and evaluation scores of speech communication ability (SCA) test.

Channel	SCA total score	Pronunciation	Semantics	Grammar	Expression efficiency	Fluency
2						+ [*r* = 0.359, *p* = 0.023]
5	+ [*r* = 0.349, *p* = 0.029]		+ [*r* = 0.373, *p* = 0.018]		+ [*r* = 0.365, *p* = 0.021]	
7	+ [*r* = 0.385, *p* = 0.014]		+ [*r* = 0.392, *p* = 0.012]			
9						+ [*r* = 0.411, *p* = 0.008]
16	+ [*r* = 0.321, *p* = 0.033]	+ [*r* = 0.420, *p* = 0.007]				
17					+ [*r* = 0.336, *p* = 0.034]	
19	+ [*r* = 0.440, *p* = 0.016]		+ [*r* = 0.393, *p* = 0.012]		+ [*r* = 0.402, *p* = 0.01]	+ [*r* = 0.383, *p* = 0.015]

+ symbol indicates positive correlation and − symbol refers to negative correlation.

TABLE 5 | Correlation between neural activation to weak-prosodic (WP) and evaluation scores of speech communication ability (SCA) test.

Channel	SCA total score	Pronunciation	Semantics	Grammar	Expression efficiency	Fluency
5	+ [*r* = 0.399, *p* = 0.011]		+ [*r* = 0.401, *p* = 0.01]			
16	+ [*r* = 0.338, *p* = 0.033]					
19					+ [*r* = 0.334, *p* = 0.035]	

+ symbol indicates positive correlation and − symbol refers to negative correlation.

significantly impaired sensitivity to strong prosodic features of speech. Last, the neural responses to the SP sentences were found to correlate with the speech communication abilities of all the child participants.

Strong-Prosodic Information Evoked Right-Lateralized Cortical Responses in the Early Processing Phase

Based on the direct comparison of Beta weights and FC results between SP and WP stimuli in the NH group, sentences with SP features evoked a larger brain network than speech with WP features (see **Figure 8**). Although previous studies on prosodic perception have found broad activation in both hemispheres (Gandour et al., 2004; Tong et al., 2005; Witteman et al., 2011), this is the first time that cortical responses to SP sentences are characterized in direct contrast to utterances with WP features.

As to the lateralization patterns, the results converge on the finding that SP sentences evoked right-lateralized activation and WP sentences activated left lateralized responses. For NH children, a higher HbO concentration was found during the middle time course in the right middle superior gyrus than that in the contralateral area (indicated by the laterality index in channel pair 8–19). An opposite result was observed between the identical pairs of the brain areas under the WP condition, which indicated the left lateralization of WP perception. The laterality index calculated on clustered data of each hemisphere also showed right-lateralized activation in the 7–12 s window in the SP condition. Furthermore, the connectivity results under the SP condition showed strong right laterality with five interhemispheric connectivities on the right and one bilateral connectivity. The different laterality tendency suggests the functional sensitivity of the right hemisphere for the auditory processing of slowly changing acoustic cues (i.e., slow F0 movements along the pronunciation of the whole sentences) that finally give a strong sense of utterance “catchyness.” This finding is in line with previous studies which found the specialization of RH on slow-changing prosodic signals (Meyer et al., 2002; Geiser et al., 2008).

Furthermore, the timewise HbO dynamics revealed selective responses to SP information in the early processing phases. In the first (2–7 s) and middle (7–12 s) time windows, higher activation to SP was found in the LH than in the RH (i.e., channel 19, **Figure 7**), which was consistent with the results of the Beta weights. In contrast, there was a higher HbO concentration in response to the WP condition than in the SP condition in both hemispheres in the late phase (12–18 s), suggesting a late selective response sensitivity to the speech with WP features. The FC results along the time course further proved this observation, with more significant pairs of connectivities observed under the SP condition in the early time window than in the late window (4 pairs vs. 2 pairs), and connectivity under the WP condition was only found in the late time window.

One possible reason for such time-wise imbalanced selectivity was that the rhythmic features in SP speech could be integrated in a short time and facilitate semantic retrieval. In EEG studies that had a high temporal resolution, prosodic information, such as rhyme congruity (Chen et al., 2016; Hahn et al., 2021) was shown to be processed as early as 200 ms after stimuli onset. Its temporal priority in the auditory speech processing was much higher than that of syntactic and semantic information which have been claimed to be processed in the late time window of 400–1,000 ms, i.e., widely recognized ERP components of N400 and P600 (Zhang et al., 2013; Delogu et al., 2019). Considering that the SP stimuli used in our study contained multiple features of prosody, with rhythm, meter, and stress being concordantly combined, such neural enhancement in the early phase of SP speech suggests that the integrated SP features in speech are processed quite fast.

The late higher cortical response in the WP condition was possibly caused by syntactic and semantic integration for sentence interpretation. Although we controlled the semantics of sentences being used in both conditions to include words and phrases commonly used in daily life, the SP sentences had relatively fixed syntactic structures (i.e., the paralleled phrases in each SP sentence use identical syntactic structure) in order to achieve a high level of prosodic harmony, which could facilitate the retrieval of the meaning of the sentence with an alleviative workload of integrating the syntactic and semantic information of the words inside (Martin, 1972; Bubic et al., 2010; Cason and Schön, 2012). Similar observations were also found in studies on formulaic speech perception (Wray, 2002; Jiang and Nekrasova, 2007; Millar, 2011). Extensive studies found that interpretation integration occurred later than phonetic perception (Friederici, 2002; Hickok and Poeppel, 2004; Zhang et al., 2019); hence, higher activation of WP was observed in the late window in trials.

Children With Cochlear Implantation Are More Impaired in Their Neural Sensitivity to Strong-Prosodic Than Weak-Prosodic Sentences

There is well-established (human and animal) literature indicating that early exposure to sound/speech is vital for the proper development of the auditory system (Sharma et al., 2002; Kral and Eggermont, 2007; Kral, 2013), and children with severe-to-profound sensorineural deafness generally lack related experience. With a natural speech perception task in the present study, we found generally decreased activation in the CI group to both SP and WP than in the NH group, which is in line with previous findings. More importantly, characteristic deficits were found in processing SP stimuli, especially in that they showed a general absence of neural distinctiveness for SP vs. WP contrasts.

Inhibited cortical responses in the CI group were found in various ways when processing SP sentences. First, the CI group had particularly weaker activation in the right middle temporal lobe (channels 13 and 16), and no active connection was found in the CI group compared to the normal-hearing children showing as many as 7 pairs of activated FCs. The lower connectivities could result from the immaturity of these areas, as shown by the low activation level generally found in the CI group. Due to hearing deprivation, the function of the language-associated cortex was impaired and thus connection among these areas was still not formed, which might cause a deficiency in processing prosodic information.

Activation of auditory and auditory association brain regions by auditory stimuli after CI was found to be important to speech communication ability (Coez et al., 2008; Kral, 2013). Abnormal neural responses to speech (Lee et al., 2007; Feng et al., 2018b) and lower FC in the auditory tasks (Chen and Wong, 2017; Wang et al., 2021; Yusuf et al., 2021) were found in post-lingually deaf adults with CI. To our best knowledge, this is the first time to identify neural functional deficits in pre-lingually deaf children while perceiving natural SP sentences. Taking into consideration that preservation of neuroanatomical networks in auditory and auditory association brain regions in pre-lingually deaf children was related to better performance after CI (Lee et al., 2007; Feng et al., 2018b), the prohibited neural sensitivity to SP sentences identified here is possibly an important obstacle for the speech development of CI children. Functional recovery of these areas after CI is worth exploring with longitudinal studies for efficient speech rehabilitation.

Moreover, in the direct contrast of neural responses to SP vs. WP perception, different patterns regarding the laterality and time-wise HbO activation were found in NH children, but no between-condition difference was shown by the CI group, suggesting that children with CI had limited sensitivity to contrasting speech prosodic features. This finding was in line with previous studies that found that patients with CI had poorer F0 discrimination and showed a strong deficit in speech prosody perception (Marx et al., 2015; Jiam et al., 2017). Scholars also found that discrimination of lexical stress patterns in infants with CI was reduced compared with that of infants without CI (Segal et al., 2016); they also found that discrimination of lexical stress patterns in infants with CI was one of the prosodic cues that they could utilize in their first steps of speech acquisition. For the first time, our study provided neurological observations and the underlying neural deficits of such reduced prosodic feature sensitivity in pre-lingually deaf children with CI.

It is worth noting that although time-wise comparisons of HbO revealed that children with CI had generally decreased activation throughout the whole processing phase, no difference was found in the late time window of 12–18 s in WP condition, suggesting that in the late phase of speech processing, CI children had close-to-normal cortical responses to perceiving sentences with WP features. These results possibly indicate that syntactic/semantic integration abilities of children with CI were relatively well reserved when excluding the prosody processing requirement, and also further confirmed that the neural auditory development of pre-lingually deaf children was more impaired for perceiving speech prosodic features.

Additionally, we found no laterality effect on children with CI under two conditions, which was different from the right-lateralization under SP, left-lateralization under WP in the NH group. Considering all the children with CI were right-sided implanted, there is a possibility that the laterality of neural activities was influenced by the implantation laterality. However, there was a study suggesting that cortical processing of speech showed no influence on the implantation side in children (Wang et al., 2022). Thus, we postulated that the laterality of CI was not an essential factor for this result.

Considering the broad significant correlations to various aspects of SCA evaluation results (overall evaluation, pronunciation, semantics, expression efficiency, and fluency) found in the right superior and middle temporal gyrus, the neural sensitivity of this area to SP sentences plays an important role in SCA. It is promising to focus on this area in speech development for pre-lingual deaf people. The neural abnormalities regarding catchy utterance perception found in our study may be used to offer an objective assessment technique for young individuals with CI without speech foundations to evaluate speech-related neural development status and predict their rehabilitation outcomes. As speech perception abilities are the foundation for the development of speech expression (Curtin et al., 2017); it was hence worthy to explore if prosodic materials have a positive effect on speech training in the future.

Children's Neural Responses to the Strong-Prosodic Sentences Are Closely Related to Their Speech Communication Development

A striking finding in this study is that the neural responses to the SP sentences were widely correlated with the speech communication abilities of all the child participants. Compared to the WP condition, more activation regions in the SP perception were found to have a positive correlation with the SCA evaluation scores (**Table 5**). We speculate that children's neural perception sensitivity to SP features in sentences is predictive of their speech development. Similar findings were found in a few behavioral studies (Falk, 2004; Wells et al., 2004; Nygaard et al., 2009; Herold et al., 2011). SP utterances (i.e., catchy sentences, sung speech, etc.) were found to enhance children's working memory performance (Yuzawa, 2002; Roy and Chiat, 2004; Yuzawa and Saito, 2006), promote speech production (Adams and Gathercole, 1995; Baddeley, 2003), and facilitate speech acquisition (Mehler et al., 1988; Leong et al., 2017; Amichetti et al., 2021). The findings in our study highlight the significance of catchy utterance materials in the speech development of children. Nevertheless, the current results alone are unable to inform claims about the causal relationship between children's SCA and their neural responses to prosodic perception. The wide range of correlations could plausibly stem from assortative processes involved in prosodic features in speech. The link may lie in that brain activities induced by catchy utterances help neural functional development, such as cortex maturation and FC formation. Future work could test this possibility by combining the forms of data collected in the current study with longitudinal studies to test the possible causality mechanism.

The right middle temporal gyrus (channel 19) was particularly sensitive to SP information as indicated by multiple results, such as significantly stronger activation measured both by Beta weights and HbO to SP than WP. HbO concentrations to SP were found to be significantly lateralized in this area compared to its contralateral part. Besides, broader connectivities were also found in this area, with significant connections to areas of the right middle superior gyrus, suggesting that the whole region around the right middle superior gyrus was particularly sensitive to SP perception. Most importantly, the response to SP in this area correlated to 4 items of SCA evaluation scores, including the overall score, semantics, expression efficiency, and fluency. As previous studies showed, the middle temporal gyrus was widely recognized as a part of the ventral stream in language processing, involved in mapping sound onto meaning (Hickok and Poeppel, 2004; Saur et al., 2008). The selective activation of this area for SP sentences suggests that natural catchy utterances might enhance the connections between prosodic auditory perception and semantic retrieval in children, which in turn accounts for the fast priming of neural processing for SP stimuli found in this study.

Past studies rarely focused on the integrated processing of natural catchy sentences with various prosodic features, as we focused on in the present work. The stimuli employed in previous studies were at the word level (Gandour et al., 1998, 2000; Hsieh et al., 2001; Li et al., 2003; Feng et al., 2018a), pseudo-sentences without semantic processing (Hurschler et al., 2013), or speech that only carries one aspect of prosody, such as intonation (Gandour et al., 2003; Tong et al., 2005), rhythm (Geiser et al., 2008), and stress (Tong et al., 2005; Sato et al., 2013). Our study provided the first insight into the neural responses responsible for the integrated processing of natural catchy sentences for children. This processing required a broader activation of the brain network in both hemispheres and was prominent in the area of the right middle temporal gyrus. Such processing patterns, as indicated by the close links to the speech behaviors, might be crucial for children's speech development, and deficits in the neural processing would cause speech impairment, as discussed in the following section.

CONCLUSION

We identified the characteristics of cortical responses to the perception of natural sentences with SP features in children. In NH children, SP sentences evoked broader and right-lateralized cortical responses than WP sentences. Stronger activation and functional connectivities were observed in the earlier phase of SP sentence processing, highlighting children's neural sensitivity to integrated prosodic features of sentences. In addition, we identified more inhibited patterns in the perception of SP than WP utterance for pre-lingually deaf children with CI, manifested as less and weaker activation, lack of right-lateralization, as well as late response-onset and the abnormalities centered in the right superior and middle temporal gyrus.

Importantly, the neural activities to SP sentences were highly correlated with the speech communication performance of both normal and CI children, suggesting that neural sensitivity in speech prosody perception may be meaningful for children's speech development.

The idiosyncratic neural responses to SP sentence perception in children with pre-lingual hearing loss shed light on the potential efficacy of SP utterances in their speech development, which is worthy of exploration in longitudinal studies.

Despite these insights, some limitations should also be recognized, aiming to provide opportunities for future work. First of all, the sample size in this study was small, with 25 individuals in the control group and 21 people with effective data in the CI group. Another limitation is the fact that we did not conduct actual longitudinal studies targeting the influence of SP sentences on children's speech development and speech rehabilitation of children with CI. Meaningful results from such studies should be further interpreted in future studies and incorporated into clinical practice.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because of the privacy issues of clinical data. Requests to access the datasets should be directed to SL, lushuo@szu.edu.cn.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethical Committee of Sun Yat-sen Memorial Hospital. Written informed consent to participate in this study was provided by the participants or their legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

YC and QL: experimental design, data collection, data analysis, and manuscript writing. ML and LG: data collection and data analysis. JY and JL: data analysis and manuscript editing. RF and YL: manuscript revising. GQ: experimental design and manuscript editing. YZ and SL: study co-supervising, experimental design, data analysis, and manuscript writing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Social Science Foundation of China to SL (Grant No. YYB023) and the National Natural Science Foundation of China to YZ (Grant No. 81771018) and JL (Grant No. 81900954).

ACKNOWLEDGMENTS

We acknowledge all patients and healthy volunteers who participated in our study. We also appreciate the two reviewers for helping in promoting the manuscript in many ways, and Sihui Chen and Huaiyao Qiu for assisting in data collection and analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.892894/full#supplementary-material

REFERENCES

Aarabi, A., Osharina, V., and Wallois, F. (2017). Effect of confounding variables on hemodynamic response function estimation using averaging and deconvolution analysis: an event-related NIRS study. *Neuroimage* 155, 25–49. doi: 10.1016/j.neuroimage.2017.04.048

Adams, A.-M., and Gathercole, S. E. (1995). Phonological working memory and speech production in preschool children. *J. Speech Lang. Hear. Res.* 38, 403–414.

Amichetti, N. M., Neukam, J., Kinney, A. J., Capach, N., March, S. U., Svirsky, M. A., et al. (2021). Adults with cochlear implants can use prosody to determine the clausal structure of spoken sentences. *J. Acoust. Soc. Am.* 150, 4315–4328. doi: 10.1121/10.0008899

Arimitsu, T., Uchida-Ota, M., Yagihashi, T., Kojima, S., Watanabe, S., Hokuto, I., et al. (2011). Functional hemispheric specialization in processing phonemic and prosodic auditory changes in neonates. *Front. Psychol.* 2:202. doi: 10.3389/fpsyg.2011.00202

Baddeley, A. (2003). Working memory and language: an overview. *J. Commun. Disord.* 36, 189–208.

Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M. C. et al. (1998). Lateralization of speech and auditory temporal processing. *J. Cogn. Neurosci.* 10, 536–540.

British Society of Audiology (2011). *Recommended Procedure Pure-Tone Air-Conduction and Bone-Conduction Threshold Audiometry With and Without Masking* (Seafield: British Society of Audiology), 38.

Bubic, A., Von Cramon, D. Y., and Schubotz, R. (2010). Prediction, cognition and the brain [Review]. *Front. Hum Neurosci.* 4:25. doi: 10.3389/fnhum.2010.00025

Cason, N., and Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia* 50, 2652–2658.

Chen, Q., Zhang, J., Xu, X., Scheepers, C., Yang, Y., and Tanenhaus, M. K. (2016). Prosodic expectations in silent reading: ERP evidence from rhyme scheme and semantic congruence in classic Chinese poems. *Cognition* 154, 11–21. doi: 10.1016/j.cognition.2016.05.007

Chen, Y., and Wong, L. (2017). Speech perception in Mandarin-speaking children with cochlear implants: a systematic review. *Int. J. Audiol.* 56, 1–10. doi: 10.1080/14992027.2017.1300694

Chen, Y., Wong, L. L., Zhu, S., and Xi, X. (2017). Vocabulary development in Mandarin-speaking children with cochlear implants and its relationship with speech perception abilities. *Res. Dev. Disabil.* 60, 243–255. doi: 10.1016/j.ridd.2016.10.010

Coez, A., Zilbovicius, M., Ferrary, E., Bouccara, D., Mosnier, I., Ambert-Dahan, E., et al. (2008). Cochlear implant benefits in deafness rehabilitation: PET study of temporal voice activations. *J. Nuclear Med.* 49, 60–67. doi: 10.2967/jnumed.107.044545

Coito, A., Plomp, G., Genetti, M., Abela, E., Wiest, R., Seeck, M., et al. (2015). Dynamic directed interictal connectivity in left and right temporal lobe epilepsy. *Epilepsia* 56, 207–217.

Curtin, S., Hufnagle, D., Mulak, K. E., and Escudero, P. (2017). "Speech perception: development," in *Reference Module in Neuroscience and Biobehavioral Psychology*, ed. J. Stein (Amsterdam: Elsevier), 1–7. doi: 10.1016/B978-0-12-809324-5.01904-0

Delogu, F., Brouwer, H., and Crocker, M. W. (2019). Event-related potentials index lexical retrieval (N400) and integration (P600) during language comprehension. *Brain Cogn.* 135:103569.

Eisenberg, N., Fabes, R. A., Bernzweig, J., Karbon, M., Poulin, R., and Hanish, L. (1993). The relations of emotionality and regulation to preschoolers' social skills and sociometric status. *Child Dev.* 64, 1418–1438.

Everhardt, M. K., Sarampalis, A., Coler, M., Başkent, D., and Lowie, W. (2022). Speech prosody: the musical, magical quality of speech. *Front. Young Minds* 10:698575. doi: 10.3389/frym.2021.698575

Falk, D. (2004). Prelinguistic evolution in early hominins: whence motherese? *Behav. Brain Sci.* 27, 491–503. doi: 10.1017/s0140525x04000111

Feng, G., Gan, Z., Wang, S., Wong, P. C., and Chandrasekaran, B. (2018a). Task-general and acoustic-invariant neural representation of speech categories in the human brain. *Cereb. Cortex* 28, 3241–3254. doi: 10.1093/cercor/bhx195

Feng, G., Ingvalson, E. M., Grieco-Calub, T. M., Roberts, M. Y., Ryan, M. E., Birmingham, P., et al. (2018b). Neural preservation underlies speech improvement from auditory deprivation in young cochlear implant recipients. *Proc. Natl. Acad. Sci.* 115, E1022–E1031. doi: 10.1073/pnas.1717603115

Flom, R., and Bahrick, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: the role of intersensory redundancy. *Dev. Psychol.* 43:238. doi: 10.1037/0012-1649.43.1.238

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* 6, 78–84. doi: 10.1016/S1364-6613(00)01839-8

Gandour, J., Dzemidzic, M., Wong, D., Lowe, M., Tong, Y., Hsieh, L., et al. (2003). Temporal integration of speech prosody is shaped by language experience: an fMRI study. *Brain Lang.* 84, 318–336. doi: 10.1016/s0093-934x(02)00505-9

Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y., et al. (2004). Hemispheric roles in the perception of speech prosody. *Neuroimage* 23, 344–357.

Gandour, J., Wong, D., and Hutchins, G. (1998). Pitch processing in the human brain is influenced by language experience. *Neuroreport* 9, 2115–2119.

Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Lancker, D. V., and Hutchins, G. D. (2000). A crosslinguistic PET study of tone perception. *J. Cogn. Neurosci.* 12, 207–222.

Geiser, E., Zaehle, T., Jancke, L., and Meyer, M. (2008). The neural correlate of speech rhythm as evidenced by metrical speech processing. *J. Cogn. Neurosci.* 20, 541–552. doi: 10.1162/jocn.2008.20029

Gfeller, K., Turner, C., Oleson, J., Zhang, X., Gantz, B., Froman, R., et al. (2007). Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise. *Ear Hear.* 28, 412–423. doi: 10.1097/AUD.0b013e3180479318

Grossmann, T., Oberecker, R., Koch, S. P., and Friederici, A. D. (2010). The developmental origins of voice processing in the human brain. *Neuron* 65, 852–858.

Hahn, L. E., Benders, T., Fikkert, P., and Snijders, T. M. (2021). Infants' implicit rhyme perception in child songs and its relationship with vocabulary. *Front. Psychol.* 12:680882. doi: 10.3389/fpsyg.2021.680882

Herold, D. S., Nygaard, L. C., Chicos, K. A., and Namy, L. L. (2011). The developing role of prosody in novel word interpretation. *J. Exp. Child Psychol.* 108, 229–241. doi: 10.1016/j.jecp.2010.09.005

Hickok, G., and Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99. doi: 10.1016/j.cognition.2003.10.011

Hoshi, Y., Kobayashi, N., and Tamura, M. (2001). Interpretation of near-infrared spectroscopy signals: a study with a newly developed perfused rat brain model. *J. Appl. Physiol.* 90, 1657–1662. doi: 10.1152/jappl.2001.90.5.1657

Hsieh, L., Gandour, J., Wong, D., and Hutchins, G. D. (2001). Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain Lang.* 76, 227–252. doi: 10.1006/brln.2000.2382

Huppert, T. J., Diamond, S. G., Franceschini, M. A., and Boas, D. A. (2009). HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Appl. Optics* 48, D280–D298. doi: 10.1364/ao.48.00d280

Hurschler, M. A., Liem, F., Jäncke, L., and Meyer, M. (2013). Right and left perisylvian cortex and left inferior frontal cortex mediate sentence-level rhyme detection in spoken language as revealed by sparse fMRI. *Hum. Brain Mapp.* 34, 3182–3192. doi: 10.1002/hbm.22134

Jiam, N., Caldwell, M., Deroche, M., Chatterjee, M., and Limb, C. (2017). Voice emotion perception and production in cochlear implant users. *Hear. Res.* 352, 30–39.

Jiang, N. A., and Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *Modern Lang. J.* 91, 433–445.

Johnson, E. K., and Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: when speech cues count more than statistics. *J. Mem. Lang.* 44, 548–567.

Jung, K. H., Cho, Y.-S., Cho, J. K., Park, G. Y., Kim, E. Y., Hong, S. H., et al. (2010). Clinical assessment of music perception in Korean cochlear implant listeners. *Acta Laryngol.* 130, 716–723.

Jusczyk, P. W., Houston, D. M., and Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cogn. Psychol.* 39, 159–207.

Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *J. Cogn. Neurosci.* 16, 1412–1425.

Kostelnik, M., Whiren, A., Soderman, A., Rupiper, M. L., and Gregory, K. (2014). *Guiding Children's Social Development and Learning*. Boston, MA: Cengage Learning.

Kovelman, I., Mascho, K., Millott, L., Mastic, A., Moiseff, B., and Shalinsky, M. H. (2012). At the rhythm of language: brain bases of language-related frequency perception in children. *Neuroimage* 60, 673–682. doi: 10.1016/j.neuroimage.2011.12.066

Kral, A. (2013). Auditory critical periods: a review from system's perspective. *Neuroscience* 247, 117–133. doi: 10.1016/j.neuroscience.2013.05.021

Kral, A., and Eggermont, J. J. (2007). What's to lose and what's to learn: development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Res. Rev.* 56, 259–269. doi: 10.1016/j.brainresrev.2007.07.021

Lee, H.-J., Giraud, A.-L., Kang, E., Oh, S.-H., Kang, H., Kim, C.-S., et al. (2007). Cortical activity at rest predicts cochlear implantation outcome. *Cereb. Cortex* 17, 909–917. doi: 10.1093/cercor/bhl001

Leong, V., Kalashnikova, M., Burnham, D., and Goswami, U. (2017). The temporal modulation structure of infant-directed speech. *Open Mind* 1, 78–90.

Li, X., Gandour, J., Talavage, T., Wong, D., Dzemidzic, M., Lowe, M., et al. (2003). Selective attention to lexical tones recruits left dorsal frontoparietal network. *Neuroreport* 14, 2263–2266. doi: 10.1097/00001756-200312020-00025

Liang, M., Zhang, X., Chen, T., Zheng, Y., Zhao, F., Yang, H., et al. (2014). Evaluation of auditory cortical development in the early stages of post cochlear implantation using mismatch negativity measurement. *Otol. Neurotol.* 35, e7–e14. doi: 10.1097/MAO.0000000000000181

Luo, X., and Fu, Q.-J. (2007). Frequency modulation detection with simultaneous amplitude modulation by cochlear implant users. *J. Acoust. Soc. Am.* 122, 1046–1054. doi: 10.1121/1.2751258

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychol. Rev.* 79, 487–509. doi: 10.1037/h0033467

Martines, F., Martines, E., Ballacchino, A., and Salvago, P. (2013). Speech perception outcomes after cochlear implantation in prelingually deaf infants: the western sicily experience. *Int. J. Pediatr. Otorhinolaryngol.* 77, 707–713. doi: 10.1016/j.ijporl.2013.01.023

Marx, M., James, C., Foxton, J., Capber, A., Fraysse, B., Barone, P., et al. (2015). Speech prosody perception in cochlear implant users with and without residual hearing. *Ear Hear.* 36, 239–248.

McCrimmon, A. W., and Smith, A. D. (2013). *Review of the Wechsler abbreviated scale of Intelligence, (WASI-II)*. Los Angeles, CA: Sage Publications Sage CA.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., and Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition* 29, 143–178.

Meyer, M., Alter, K., Friederici, A. D., Lohmann, G., and von Cramon, D. Y. (2002). FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Hum. Brain Mapp.* 17, 73–88. doi: 10.1002/hbm.10042

Millar, N. (2011). The processing of malformed formulaic language. *Appl. Linguist.* 32, 129–148. doi: 10.1016/s0093-934x(03)00062-2

Morris, D., Magnusson, L., Faulkner, A., Jönsson, R., and Juul, H. (2013). Identification of vowel length, word stress, and compound words and phrases by postlingually deafened cochlear implant listeners. *J. Am. Acad. Audiol.* 24, 879–890. doi: 10.3766/jaaa.24.9.11

Ni, G., Zheng, Q., Liu, Y., Zhao, Y., Yue, T., Han, S., et al. (2021). Objective electroencephalography-based assessment for auditory rehabilitation of pediatric cochlear implant users. *Hear. Res.* 404:108211. doi: 10.1016/j.heares.2021.108211

Nygaard, L. C., Herold, D. S., and Namy, L. L. (2009). The semantics of prosody: acoustic and perceptual evidence of prosodic correlates to word meaning. *Cogn. Sci.* 33, 127–146. doi: 10.1111/j.1551-6709.2008.01007.x

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.

Peng, S.-C., Chatterjee, M., and Lu, N. (2012). Acoustic cue integration in speech intonation recognition with cochlear implants. *Trends Amplif.* 16, 67–82. doi: 10.1177/1084713812451159

Peng, S.-C., Weiss, A. L., Cheung, H., and Lin, Y.-S. (2004). Consonant production and language skills in Mandarin-speaking children with cochlear implants. *Arch. Otolaryngol. Head Neck Surgery* 130, 592–597.

Plichta, M. M., Heinzel, S., Ehlis, A.-C., Pauli, P., and Fallgatter, A. J. (2007). Model-based analysis of rapid event-related functional near-infrared spectroscopy (NIRS) data: a parametric validation study. *Neuroimage* 35, 625–634. doi: 10.1016/j.neuroimage.2006.11.028

Quaresima, V., Bisconti, S., and Ferrari, M. (2012). A brief review on the use of functional near-infrared spectroscopy (fNIRS) for language imaging studies in human newborns and adults. *Brain Lang.* 121, 79–89. doi: 10.1016/j.bandl.2011.03.009

Roy, P., and Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2-to 4-year-olds. *J. Speech Lang. Hear. Res.* 47, 223–234. doi: 10.1044/1092-4388(2004/019)

Saliba, J., Bortfeld, H., Levitin, D. J., and Oghalai, J. S. (2016). Functional near-infrared spectroscopy for neuroimaging in cochlear implant recipients. *Hear. Res.* 338, 64–75.

Sato, Y., Utsugi, A., Yamane, N., Koizumi, M., and Mazuka, R. (2013). Dialectal differences in hemispheric specialization for Japanese lexical pitch accent. *Brain Lang.* 127, 475–483.

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., et al. (2008). Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci.* 105, 18035–18040.

Scheiner, E., Hammerschmidt, K., Jürgens, U., and Zwirner, P. (2006). Vocal expression of emotions in normally hearing and hearing-impaired infants. *J. Voice* 20, 585–604. doi: 10.1016/j.jvoice.2005.09.001

Segal, O., Houston, D., and Kishon-Rabin, L. (2016). Auditory discrimination of lexical stress patterns in hearing-impaired infants with cochlear implants compared with normal hearing: influence of acoustic cues and listening experience to the ambient language. *Ear Hear.* 37, 225–234. doi: 10.1097/AUD.0000000000000243

Sharma, A., Campbell, J., and Cardon, G. (2015). Developmental and cross-modal plasticity in deafness: evidence from the P1 and N1 event related potentials in cochlear implanted children. *Int. J. Psychophysiol.* 95, 135–144. doi: 10.1016/j.ijpsycho.2014.04.007

Sharma, A., Dorman, M. F., and Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear Hear.* 23, 532–539.

Snedeker, J., and Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *J. Mem. Lang.* 58, 574–608. doi: 10.1016/j.jml.2007.08.001

Soltanlou, M., Sitnikova, M. A., Nuerk, H.-C., and Dresler, T. (2018). Applications of functional near-infrared spectroscopy (fNIRS) in studying cognitive development: the case of mathematics and language [Review]. *Front. Psychol.* 9, 1–15. doi: 10.3389/fpsyg.2018.00277

Strangman, G., Culver, J. P., Thompson, J. H., and Boas, D. A. (2002). A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. *Neuroimage* 17, 719–731.

Svirsky, M. (2017). Cochlear implants and electronic hearing. *Phys. Today* 70, 52–58.

Takakura, H., Shojaku, H., Takamoto, K., Urakawa, S., Nishijo, H., and Watanabe, Y. (2011). Cortical hemodynamic responses to intravenous thiamine propyldisulphide administration detected by multichannel near infrared spectroscopy (NIRS) system. *Brain Topogr.* 24, 114–126. doi: 10.1007/s10548-011-0179-9

Takamoto, K., Hori, E., Urakawa, S., Sakai, S., Ishikawa, A., Kohno, S., et al. (2010). Cerebral hemodynamic responses induced by specific acupuncture sensations during needling at trigger points: a near-infrared spectroscopic study. *Brain Topogr.* 23, 279–291. doi: 10.1007/s10548-010-0148-8

Takeuchi, M., Hori, E., Takamoto, K., Tran, A. H., Satoru, K., Ishikawa, A., et al. (2009). Brain cortical mapping by simultaneous recording of functional near infrared spectroscopy and electroencephalograms from the whole brain during right median nerve stimulation. *Brain Topogr.* 22, 197–214. doi: 10.1007/s10548-009-0109-2

Tomasi, D., and Volkow, N. D. (2010). Functional connectivity density mapping. *Proc. Natl. Acad. Sci.* 107, 9885–9890. doi: 10.1073/pnas.1001414107

Tong, Y., Gandour, J., Talavage, T., Wong, D., Dzemidzic, M., Xu, Y., et al. (2005). Neural circuitry underlying sentence-level linguistic prosody. *Neuroimage* 28, 417–428. doi: 10.1016/j.neuroimage.2005.06.002

Venail, F., Vieu, A., Artieres, F., Mondain, M., and Uziel, A. (2010). Educational and employment achievements in prelingually deaf children who receive cochlear implants. *Arch. Otolaryngol. Head Neck Surgery* 136, 366–372. doi: 10.1001/archoto.2010.31

Wang, J., Wei, Q., Yuan, X., Jiang, X., Xu, J., Zhou, X., et al. (2018). Local functional connectivity density is closely associated with the response of electroconvulsive therapy in major depressive disorder. *J. Affect. Disord.* 225, 658–664. doi: 10.1016/j.jad.2017.09.001

Wang, S., Lin, M., Sun, L., Chen, X., Fu, X., Yan, L., et al. (2021). Neural mechanisms of hearing recovery for cochlear-implanted patients: an electroencephalogram follow-up study. *Front. Neurosci.* 14:624484. doi: 10.3389/fnins.2020.624484

Wang, Y., Wu, M., Wu, K., Liu, H., Wu, S., Zhang, Z., et al. (2022). Differential auditory cortical development in left and right cochlear implanted children. *Cereb. Cortex* 1–17. doi: 10.1093/cercor/bhac025 [Epub ahead of print].

Wechsler, D. (2011). *Wechsler Abbreviated scale of Intelligence*, 2nd Edn. San Antonio, TX: NCS Pearson.

Wells, B., Peppé, S., and Goulandris, N. (2004). Intonation development from five to thirteen. *J. Child Lang.* 31, 749–778. doi: 10.1017/s030500090400652x

Werker, J. F., and Hensch, T. K. (2015). Critical periods in speech perception : new directions. *Annu. Rev. Psychol.* 66, 173–196.

Witteman, J., van IJzendoorn, M. H., van de Velde, D., van Heuven, V. J., and Schiller, N. O. (2011). The nature of hemispheric specialization for linguistic and emotional prosodic perception: a meta-analysis of the lesion literature. *Neuropsychologia* 49, 3722–3738. doi: 10.1016/j.neuropsychologia.2011.09.028

Wray, A. (2002). *Formulaic Language and the Lexicon.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511519772

Wu, C.-M., Lee, H.-L., Hwang, J.-H., Sun, Y.-S., and Liu, T.-C. (2008). Intellectual ability of Mandarin-speaking children using cochlear implants. *Audiol. Neurotol.* 13, 302–308. doi: 10.1159/000124278

Yamamoto, T., and Kato, T. (2002). Paradoxical correlation between signal in functional magnetic resonance imaging and deoxygenated haemoglobin content in capillaries: a new theoretical explanation. *Phys. Med. Biol.* 47:1121. doi: 10.1088/0031-9155/47/7/309

Ye, J. C., Tak, S., Jang, K. E., Jung, J., and Jang, J. (2009). NIRS-SPM: statistical parametric mapping for near-infrared spectroscopy. *Neuroimage* 44, 428–447.

Yusuf, P. A., Hubka, P., Tillein, J., Vinck, M., and Kral, A. (2021). Deafness weakens interareal couplings in the auditory cortex. *Front. Neurosci.* 14:625721. doi: 10.3389/fnins.2020.625721

Yuzawa, M. (2002). Effects of pitch accent on short-term retention of words in young children. *Shinrigaku Kenkyu* 73, 258–263. doi: 10.4992/jjpsy.73.258

Yuzawa, M., and Saito, S. (2006). The role of prosody and long-term phonological knowledge in Japanese children's nonword repetition performance. *Cogn. Dev.* 21, 146–157. doi: 10.1016/j.cogdev.2006.01.003

Zhang, G., Si, Y., and Dang, J. (2019). Revealing the dynamic brain connectivity from perception of speech sound to semantic processing by EEG. *Neuroscience* 415, 70–76. doi: 10.1016/j.neuroscience.2019.07.023

Zhang, X., Qin, F., Chen, Z., Gao, L., Qiu, G., and Lu, S. (2020). Fast screening for children's developmental language disorders via comprehensive speech ability evaluation—using a novel deep learning framework. *Ann. Trans. Med.* 8:707. doi: 10.21037/atm-19-3097

Zhang, Y., Li, P., Piao, Q., Liu, Y., Huang, Y., and Shu, H. (2013). Syntax does not necessarily precede semantics in sentence processing: ERP evidence from Chinese. *Brain Lang.* 126, 8–19.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Neuroscience

TYPE Original Research
PUBLISHED 27 July 2022
DOI 10.3389/fnins.2022.903924

OPEN ACCESS

EDITED BY
Xin Luo,
Arizona State University, United States

REVIEWED BY
Brett Anthony Swanson,
Cochlear, Australia
David M. Landsberger,
New York University, United States

*CORRESPONDENCE
Susan Rebekah Subrahmanyam Bissmeyer
ssubrahm@usc.edu

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

RECEIVED 24 March 2022
ACCEPTED 11 July 2022
PUBLISHED 27 July 2022

CITATION
Bissmeyer SRS, Ortiz JR, Gan H and Goldsworthy RL (2022) Computer-based musical interval training program for Cochlear implant users and listeners with no known hearing loss.
Front. Neurosci. 16:903924.
doi: 10.3389/fnins.2022.903924

Computer-based musical interval training program for Cochlear implant users and listeners with no known hearing loss

Susan Rebekah Subrahmanyam Bissmeyer[1,2*†], Jacqueline Rose Ortiz[1†], Helena Gan[1] and Raymond Lee Goldsworthy[1]

[1]Caruso Department of Otolaryngology, Auditory Research Center, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States, [2]Department of Biomedical Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, CA, United States

A musical interval is the difference in pitch between two sounds. The way that musical intervals are used in melodies relative to the tonal center of a key can strongly affect the emotion conveyed by the melody. The present study examines musical interval identification in people with no known hearing loss and in cochlear implant users. Pitch resolution varies widely among cochlear implant users with average resolution an order of magnitude worse than in normal hearing. The present study considers the effect of training on musical interval identification and tests for correlations between low-level psychophysics and higher-level musical abilities. The overarching hypothesis is that cochlear implant users are limited in their ability to identify musical intervals both by low-level access to frequency cues for pitch as well as higher-level mapping of the novel encoding of pitch that implants provide. Participants completed a 2-week, online interval identification training. The benchmark tests considered before and after interval identification training were pure tone detection thresholds, pure tone frequency discrimination, fundamental frequency discrimination, tonal and rhythm comparisons, and interval identification. The results indicate strong correlations between measures of pitch resolution with interval identification; however, only a small effect of training on interval identification was observed for the cochlear implant users. Discussion focuses on improving access to pitch cues for cochlear implant users and on improving auditory training for musical intervals.

KEYWORDS

auditory neuroscience, cochlear implant, hearing loss, learning, music, plasticity

Introduction

Cochlear implants have successfully restored speech perception to people with severe hearing loss. Most cochlear implant users achieve high levels of speech recognition and spoken language skills (Shannon et al., 2004; Wilson and Dorman, 2008). However, cochlear implant users struggle to understand speech in noisy environments and many complain about the sound of music (Fetterman and Domico, 2002; Kong et al., 2004; McDermott, 2004; do Nascimento and Bevilacqua, 2005; Fu and Nogaki, 2005; Nimmons et al., 2008). Studies have shown that current cochlear implant technology is limited in its ability to convey the musical percepts of pitch and timbre (Drennan and Rubinstein, 2008; Limb and Rubinstein, 2012). This has resulted in both pitch resolution and timbre recognition being markedly diminished for cochlear implant users compared to their normal hearing peers (Gfeller et al., 2002b, 2007; McDermott, 2004; Drennan et al., 2008; Goldsworthy et al., 2013; Limb and Roy, 2014; Goldsworthy, 2015; Luo et al., 2019). This loss of resolution and fidelity has several potential causes including limited number of implanted electrodes, electrode array placement, broad current spread, sound processing designed for speech rather than music, poor coding of timing cues for pitch, and poor neural health (Finley et al., 2008; Rebscher, 2008; Crew et al., 2012; Limb and Roy, 2014; van der Marel et al., 2014; Würfel et al., 2014; Zeng et al., 2014; Landsberger et al., 2015; Venail et al., 2015; Nogueira et al., 2016; Caldwell et al., 2017; Dhanasingh and Jolly, 2017; Mangado et al., 2018).

These technological and physiological constraints limit how music is transmitted by the implant and, consequently, limits music enjoyment for cochlear implant users. Studies have assessed adult cochlear implant user's listening habits and music enjoyment through questionnaires (Gfeller et al., 2000; Looi and She, 2010). They found that many were dissatisfied and spent less time listening to music post-implantation. Assessment studies have also shown that cochlear implant users have more difficulty than normal hearing listeners with pitch-based perceptual tasks, including frequency discrimination and melody recognition (Gfeller et al., 2002a, 2005, 2007; Penninger et al., 2013; Goldsworthy, 2015).

Melody is a fundamental aspect of music made up of a sequence of musical intervals which not only relies on the detection and direction of pitch changes, but also their magnitude. Even for those who casually listen to music, identifying the magnitude between pitches is a basic component which allows a listener to readily recognize a melody whether sung in a different register or played in a different key. If a difference in frequency cannot reliably be heard as an equivalent change in pitch, then the intended melody sounds cacophonous and out-of-tune. This has been confirmed by Luo et al. (2014) who found that cochlear implant users perceived melodies as out-of-tune more often than normal hearing listeners. Furthermore, the ability to perceive musical intervals also has implications for the emotion and tension conveyed by music. A single semitone difference between two pitches will determine the tonality of the interval (e.g., major, minor, diminished, perfect, or augmented) which, along with other important cues like timbre and tempo, will affect the listener's emotional response to a melody (Luo and Warner, 2020; Camarena et al., 2022). The ability to reliably distinguish intervals requires listeners to have a resolution of at least a semitone (McDermott et al., 2010), and it is well established that most cochlear implant users have pitch resolution that is worse than a semitone (e.g., Pretorius and Hanekom, 2008; Goldsworthy, 2015). Without accurate perception of a musical interval, it is likely that tonality and emotion intended to be conveyed by music will be lost and this is likely a contributing factor to decreased musical enjoyment in cochlear implant users.

Musical interval labeling is an important skill for musicians and any individual who desires to participate in musical activities such as playing an instrument or singing. It is difficult to master identifying musical intervals, even in normal-hearing listeners and musicians (McDermott et al., 2010). Given the evidence discussed that suggests that musical interval perception is distorted in the context of melody perception for cochlear implant users, it is likely that cochlear implant users struggle to identify musical intervals as well. It is necessary for cochlear implant users to take steps to regain access to interval cues for musical tension and emotion. They must first undergo a period of focused aural rehabilitation to learn how the lower-level pitch cues are provided by electrical stimulation *via* their device (Gfeller, 2001), then develop the higher-level association between specific musical intervals and intent through further musical interval training (Fujioka et al., 2004).

Despite the importance of intervals to melody, there is only a small body of research investigating musical interval perception in cochlear implant users. Existing studies have shown that cochlear implant users have poor interval identification compared to their normal-hearing peers, especially above middle C. Pitch and relative intervals can be conveyed by stimulation timing (i.e., the modulation or stimulation rate) but with much variability in pitch salience and in the upper frequency that can be conveyed by stimulation rate (Pijl and Schwarz, 1995a,b; Pijl, 1997; Todd et al., 2017). Place cues for pitch (i.e., active electrodes and stimulation configuration) provide a strong sense of pitch but one that is compressed compared to normal (Stupak et al., 2021). Stupak et al. (2021) found consistent warping of intervals amongst cochlear implant users, suggesting the ability to perceive intervals is likely not linked to duration of deafness. Spitzer et al. (2021) investigated musical interval distortion in cochlear implant users who had normal hearing in their non-implanted ear (i.e., single-sided deafness). They found that the musical interval needed to create a match in the implanted ear was, on average, 1.7 times greater than the corresponding interval in the acoustic hearing ear.

Given the distorted representation of pitch and the issue of frequency compression in current cochlear implant signal processing, experience and training may be required to improve interval identification and enable access to melody through clinical devices. Interval identification is challenging for normal-hearing people and cochlear implant users alike, which makes it a demanding task for training. Moore and Amitay found that pitch training with a more difficult, or even impossible, task resulted in more robust learning (Moore and Amitay, 2007). Musical interval training in normal hearing listeners has led to improvement in both the trained and untrained tasks (Little et al., 2019). There are currently no studies investigating the effectiveness of musical interval training in cochlear implant users.

In the present study, we use an interval labeling task to evaluate subject's ability to strengthen the association between specific musical intervals and musical intent and to consistently label intervals across an ecologically relevant musical range (i.e., the typical vocal range of humans). We note the connection between musical intervals and musical intent does not require the ability to label intervals, for example, a listener may readily associate a song in a major key as happy or bright and a song in a minor key as sad or dark (Camarena et al., 2022) without being able to label the interval pattern being used. However, given that we are interested in the restoration of a stable interval percept in cochlear implant users, we chose to use a labeling task as an important intermediary to quantify the consistency of interval labeling across musical octaves when those cochlear implant users are provided with training to the interval cues. This training task requires participants to attend to multiple musical interval presentations, associate interval magnitudes with specific labels (e.g., major third, octave), and compare presentations to intervals heard in preceding trials.

The present study has two objectives. First, to examine the performance on the trained task of interval identification and on a battery of untrained musical tasks, including frequency discrimination and tonal and rhythm comparisons before and after a two-week musical interval training program. Second, to characterize the relationship between the dimensions of music perception with low-level psychoacoustics and higher-level rhythm and tonal comparisons, interval identification, and musical sophistication. The overarching hypothesis motivating this study is that both low-level psychophysical access to pitch cues as well as higher-level labeling of intervals limits interval identification accuracy in cochlear implant users, and, to a certain extent, those with no known hearing loss. The results show that the low-level psychophysical tasks probing pitch resolution serve as predictors of higher-level measures of music perception. The results also clarify the extent that interval training improves access to the low-level and higher-level cues necessary for music perception. Discussion focuses on the importance of basic elements of pitch perception for reestablishing musical interval perception for cochlear implant users and on methods for improving training programs for musical interval identification.

Materials and methods

Overview

Participants with no known hearing loss and cochlear implant users completed assessments before and after 2 weeks of interval training. The pre- and post-assessments included measures of pure tone detection, pure tone frequency and fundamental frequency discrimination, tonal and rhythm comparisons, and musical interval identification (the trained task) administered on the Team Hearing website coded in JavaScript. The measures of pure tone detection, pure tone frequency, and fundamental frequency discrimination used synthesized stimuli generated using JavaScript. The measures of tonal and rhythm comparisons used marimba notes rendered using Finale Version 3.5.1 software (Coda Music), and the measures of interval identification for both training and assessment used piano notes rendered using MuseScore 3 software[1]. **Figure 1** shows typical normal hearing neural response patterns (left subpanel) and Cochlear Corporation cochlear implant stimulation patterns (right subpanel) for representative musical notes, highlighting the difference in frequency representation between the two groups. In the left subpanel, the 110 Hz and 220 Hz place cues can be visualized at the fundamental as well as the ascending harmonic frequencies and temporal cues can be observed with a doubling of the rate for 220 Hz. In the right subpanel, the place and temporal cues are not as clearly visualized, with the harmonic structure coarsely represented and the fundamental frequencies conveyed only through weak amplitude modulation. The place and temporal representation in cochlear implant stimulation is poor compared to the cues available for pitch perception in the normal auditory system. This representation reinforces the basis of the first part of the hypothesis, that cochlear implant users are limited in low-level psychophysical access to pitch cues. A permalink for this experiment can be found at: https://www.teamhearing.org/81, after entering the site, press the "Homework" button to enter the experiment.

Participants

Thirteen adult cochlear implant users, with six bilaterally implanted and seven unilaterally implanted, and seven listeners with no known hearing loss took part in this experiment. All participants completed the 2-week interval training

1 https://musescore.org/en

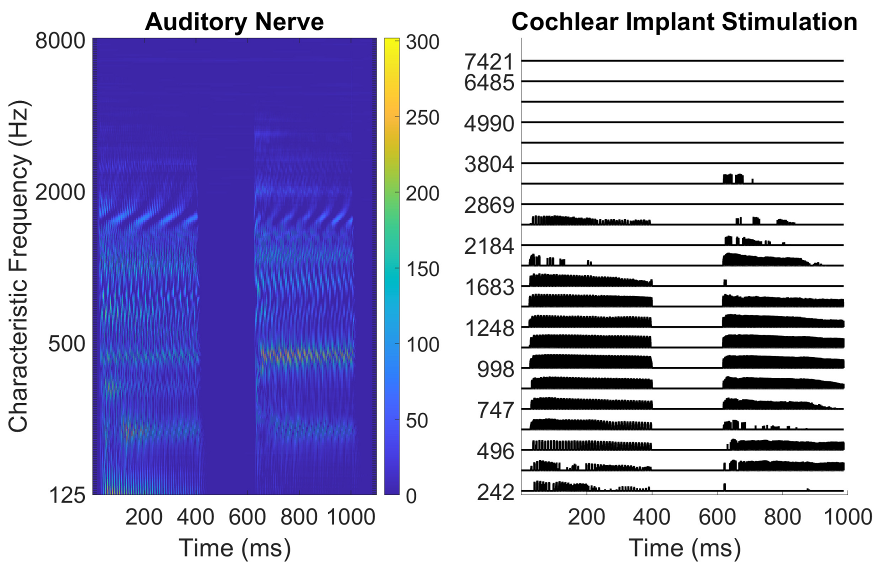

FIGURE 1
Visualizations of musical notes. The left subpanel shows auditory nerve response to musical notes for normal hearing using physiology modeling software (Zilany et al., 2014). The right subpanel shows cochlear implant stimulation patterns emulated using the Nucleus MATLAB Toolbox (Nucleus MATLAB Toolbox version 4.42, Swanson and Mauch, 2006). For both visualizations, the two notes being compared are A2 (110 Hz) and A3 (220 Hz).

protocol. Participant ages ranged from 23 to 77 years old with an average age of 62.9 years in the cochlear implant user group and 42.3 years in listeners with no known hearing loss. Relevant subject information is provided in Table 1. Participants provided informed consent and were paid for their participation. The experimental protocol was approved by the University of Southern California Institutional Review Board.

Training

All assessments and the musical interval training program were completed remotely by participants using a web application. For training, participants complete six listening exercises each day requiring approximately 20 min each day for 2 weeks. Each listening exercise included twenty trials of interval identification for which participants needed to identify 80% of the intervals correctly to proceed to the next difficulty training level. Levels were organized into thirty-six increasingly difficult levels with fewer comparisons and larger interval spacings on lower difficulty training levels.

For each trial, listeners were presented with an ascending musical interval and asked to indicate the interval that they heard. The online interface displayed two to four response buttons on screen depending on the level, with specific musical interval labels provided for selection. In total, training was provided for six different ascending melodic intervals consisting of two sequentially presented piano notes. The intervals presented and the corresponding semitone spacings between notes are listed in Table 2. Practice was provided for intervals with base notes near A2 (110 Hz), A3 (220 Hz), and A4 (440 Hz). These training levels were divided into 6 different interval groupings with 6 base note frequencies within each interval grouping. The interval groupings, described in semitone spacing between notes, were [2,12], [2,7], [7,12], [4,7,12], [2,4,7], and [1,2,3,4]. The base note frequencies within each interval grouping were (1) A2 (110 Hz) no variation, (2) A2 (110 Hz) +/- 6 semitones, (3) A3 (220 Hz) no variation, (4) A3 (220 Hz) +/- 6 semitones, (5) A4 (440 Hz) no variation, and (6) A4 (440 Hz) +/- 6 semitones. See Supplementary Table 1 for more information about the training levels. Feedback was displayed after each response on the response button selected with a green check mark for correct answers and a red "X" for wrong answers. For wrong answers, participants were given the correct answer on screen and the option to replay the interval comparison as needed.

Pre- and post-training assessments

Participants completed pre- and post-training assessments to characterize the effect of training on the trained task and on untrained measures of pitch discrimination and music perception. The assessments included pure tone detection, pure tone frequency discrimination, fundamental frequency discrimination, tonal and rhythm comparisons, and musical interval identification.

TABLE 1 Subject information.

Subject	Age	Gender	Etiology	Ear tested	MSI score	Age at onset	Years implanted	CI company and processor	Implant model	Duration of deafness before implantation	Method of streaming
H1	53	M	No Known Hearing Loss	Both Together	3.61	N/A	N/A	N/A	N/A	N/A	Apple Earbuds
H2	24	F	No Known Hearing Loss	Both Together	5.89	N/A	N/A	N/A	N/A	N/A	Koss UR20 Headphones
H3	66	F	No Known Hearing Loss	Both Together	3.5	N/A	N/A	N/A	N/A	N/A	Apple Earbuds
H4	54	M	No Known Hearing Loss	Both Together	3.83	N/A	N/A	N/A	N/A	N/A	Free Field through Dell Optiplex 3080 Speakers
H5	39	M	No Known Hearing Loss	Both Together	4.33	N/A	N/A	N/A	N/A	N/A	Free Field through Panasonic TV TH-50PX80U speakers
H6	23	F	No Known Hearing Loss	Both Together	5.94	N/A	N/A	N/A	N/A	N/A	Free Field through Yamaha HS5 Powered Studio Monitor Speaker
H7	37	F	No Known Hearing Loss	Both Together	6.39	N/A	N/A	N/A	N/A	N/A	Beyer Dynamic DT 770 Pro Headphones
C2	37	F	Unknown	Both Together	4.78	15	L:9 R:13	Cochlear N7s	L:CI24RE (CA) R:CI24RE (CA)	L:5 R:1	Mini Mic2
C3	76	F	Progressive SNHL	Both Together	2.11	40	L:21 R:17	Cochlear N6s	L:CI24R (CS) R:CI24RE (CA)	L:1 R:5	Cochlear Binaural Cable
C10	46	M	Ototoxic Medicine	Left	3.28	12	33	Cochlear N6	CI22M-, United States	1	Mini Mic
C11	58	F	Sudden SNHL	Right	1.83	55	2	Advanced Bionics Naida CI Q90	HiRes Ultra 3D CI HIFocus SlimJ	1	AB Bluetooth
C13	59	M	Mumps Disease	Right	3.39	14	3	Med-El Sonnet	Sonata 2 Mi1260	42	I-loop streaming
C15	58	M	Ototoxic Medicine	Left	2	54	1	Advanced Bionics Naida	HiRes Ultra 3D CI with HiFocus Mid-Scala Electrode	1	Bluetooth/Compilot
C16	66	M	Ototoxic Medicine	Left	4.11	38	18	Cochlear N5	CI24R (CS)	5	Sony MDR-D150 Headphones
C17	74	F	Unknown	Both Together	1.78	Birth	L:20 R:15	Cochlear N6s	L:CI24R (CS) R:CI24RE (CA)	L:9 R:9	Free Field through HP Computer Speakers
C18	72	F	Measles In Utero	Both Together	2.56	Birth	L:12 R:10	Cochlear N6s	L:CI24RE (CA) R:CI512	L:1 R:1	Free Field through HP Computer Speakers
C20	67	F	Unknown	Both Together	4.11	18	L:4 R:5	L:Cochlear N6 R:Cochlear N7	L:CI522 R:CI522	L:14 R:16	Free Field through iPad Speakers
C22	65	F	Mumps Disease	Left	5.11	5	2	Cochlear N7	CI512	58	Mini Mic
C28	77	M	Unknown	Both Together	3.33	60	L:2 R:1	Med-El Rondo 3s	Synchrony 2 Mi1250	L:1 R:2	Bluetooth streaming using AudioLink
C32	63	F	Progressive SNHL	Left	6.28	20	13	Cochlear N7	CI24RE (CA)	5	Direct Bluetooth streaming from iPad

Age at time of testing and age at onset of hearing loss (when applicable) is given in years. Duration of profound hearing loss prior to implantation (when applicable) is given in years and estimated from subject interviews. SNHL, Sensorineural Hearing Loss.

TABLE 2 Interval notation with the corresponding semitone spacing between notes.

Interval	Semitone spacing
Minor 2nd	1
Major 2nd	2
Minor 3rd	3
Major 3rd	4
Perfect 5th	7
Octave	12

Calibration procedures

Before completing the assessments, participants completed two procedures to characterize relative loudness levels with their devices (computer, audio device, hearing device, etc.) kept how the subject would normally listen. First, participants were asked to use a method of adjustment to set a 1 kHz pure tone to subjective "soft," "medium soft," "medium," and "medium loud" intensity levels in dB relative to the maximum output level of sound card without clipping. Second, pure tone detection thresholds were measured in dB relative to the maximum output level of sound card at 250, 1,000, and 4,000 Hz to provide a comparison of relative detection levels across frequencies. Stimuli were 400 ms sinusoids with 20 ms raised-cosine attack and release ramps. At the beginning of a measurement run, participants set the volume to a "soft but audible" level. The detection thresholds were then measured using a three-alternative, three-interval, forced-choice procedure in which two of the intervals contained silence and one interval contained the gain-adjusted tone. Participants were told *via* on-screen instructions to select the interval that contained the tone. The starting gain value was a threshold level as specified by the participant through method of adjustment. This value was reduced by 2 dB after correct answers and increased by 6 dB after mistakes to obtain the true detection threshold level. A run continued until three mistakes were made and the average of the last four reversals was taken as the detection threshold. This procedure converges to 75% detection accuracy (Kaernbach, 1991). Relative dynamic range could then be calculated by subtracting the detection threshold from the comfortable listening intensity level set at 1,000 Hz. The remainder of the assessments and interval training were conducted at the volume the participant set as "comfortable."

Pure tone frequency discrimination

Pure tone frequency discrimination was measured for pure tones near 250, 1,000, and 4,000 Hz. Stimuli were 400 ms in duration with 20 ms raised-cosine attack and release ramps. Discrimination was measured using a two-alternative, two-interval, forced-choice procedure where the target stimulus had an adaptively higher frequency than the standard. Participants were provided with on-screen instructions to choose the sound that was "higher in pitch." Each measurement run began with a frequency difference of 100% (an octave) between the standard and target stimuli. This frequency difference was reduced by a factor of $\sqrt[3]{2}$ after correct answers and increased by a factor of two after mistakes. For each trial, the precise frequency tested was roved to add perturbations which contribute to the ecological relevance of the stimulus (e.g., vocal pitch fluctuations) while avoiding both artifactual effects (e.g., sidebands outside of the filter, beating) and habituation to the base note frequency. The frequency roving was done within a quarter-octave range uniformly distributed and geometrically centered on the nominal condition frequency. Relative to the roved frequency value, the standard frequency was lowered, and the target raised by $\sqrt{1+\triangle/100}$. The gain of the standard and target were roved by 6 dB based on a uniform distribution centered on the participant's comfortable listening level. A run ended when the participant made four mistakes and the average of the last four reversals was taken as the discrimination threshold.

Fundamental frequency discrimination

Fundamental frequency discrimination was measured for fundamental frequencies near 110, 220, and 440 Hz for low pass filtered harmonic complexes. Stimuli were 400 ms in duration with 20 ms raised-cosine attack and release ramps. These fundamental frequencies were chosen as representative of the fundamental frequencies used in the interval identification assessment and training. A total of nine measurement runs were conducted consisting of three repetitions of the three fundamental frequencies. The condition order was randomized for each repetition. Harmonic complexes were constructed in the frequency domain by summing all non-zero harmonics from the fundamental to 2 kHz with a low pass filtering function. All harmonics were of equal amplitude prior to filtering. The form of the low pass filtering function was:

$$gain = \begin{cases} 1 & iff < f_e \\ \max\left(0,\ 1-(\log_2 f - \log_2 f_e)^2\right) & otherwise \end{cases} \quad (1)$$

where *gain* is the gain expressed as a linear multiplier applied to each harmonic component, f is the frequency of the component, and f_e is the edge frequency of the passband, which was set as 1 kHz for the low pass filter. Note, as thus defined, the low pass filter gain is zero above 2 kHz. Fundamental frequency discrimination was measured using a two-alternative, two-interval, forced-choice procedure where the target had an adaptively higher fundamental frequency compared to the standard. The same adaptive procedure, amplitude and frequency roving, and scoring logic were used as for pure tone frequency discrimination but with adaptive control over fundamental frequency.

Tonal and rhythm comparisons

Participant performance on tonal and rhythm comparisons was measured using a two-alternative, two-interval, forced-choice procedure. The stimuli were the same as those generated and used by Habibi et al. (2016). In each trial, participants were presented with two 2.5 s long pre-rendered melodies rendered with marimba-like timbre, which contained 5 distinct pitches corresponding to the first 5 notes of the C major scale with fundamental frequencies ranging from 261 to 392 Hz (Habibi et al., 2016). The melodies were either the same or differed on a single note in terms of tonality or rhythm, and the listener had to choose between the on-screen options: "Same" or "Different." The tonal and rhythm comparison procedures tested the subjects ability to identify deviations in either tonality or rhythm between pairs of unfamiliar 5-note melodies based on Western classical rules (Habibi et al., 2013, 2014, 2016). Tonal, or pitch, deviations involved the pitch change of a single note in the 5-note melody. The pitch deviations were restricted to the first 5 notes of the C major scale. Rhythm deviations involved the prolongation of a single note creating a delay in the subsequent note, the duration of which was consequently shorter so that the offset time was unchanged. The duration of each note ranged from 125 ms to 1,500 ms to create rhythmic patterns. The standard melody had no deviations in pitch or note duration. This assessment consisted of three repetitions of each set, consisting of twenty-four trials, half of which were tonal comparisons and half of which were rhythm comparisons. Performance was measured as the percentage of correct responses for each comparison domain.

Interval identification

Performance on musical interval identification was assessed with piano notes for three note ranges near A2, A3, and A4 (110, 220, and 440 Hz, respectively). Participants were presented with two sequentially played piano notes separated by 4, 7, or 12 semitones to represent a major 3rd, perfect 5th, or octave interval, respectively. Note, these specific test conditions corresponded to training levels 20, 22, and 24 of the training program. Responses were collected using a three-alternative forced-choice procedure where the participant had to choose between the on-screen options: "major 3rd," perfect 5th," or "octave." Each measurement run consisted of twenty trials and there were three repetitions of each condition (A2, A3, A4) for a total of nine measurement runs. The musical interval chosen on any trial was randomly selected. In total, each participant completed 180 trials during the interval identification assessment and was presented with approximately 60 presentations of each of the three intervals utilized in this assessment. The base note of the comparison was roved within an octave range centered on the nominal condition note.

The goldsmith musical sophistication index

The level of prior musical experience was measured using the Goldsmith Musical Sophistication Index Self-Report Inventory (MSI), a 39-item psychometric instrument used to quantify the amount of musical engagement, skill, and behavior of an individual (Müllensiefen et al., 2014). The questions on this assessment are grouped into five subscales: active engagement, perceptual abilities, musical training, singing abilities, and emotion. Questions under the active engagement category consider instances of deliberate interaction with music (i.e., "I listen attentively to music for *X* hours per day"). The perceptual abilities category includes questions about music listening skills (e.g., "I can tell when people sing or play out of tune"). Musical training questions inquire about individuals' formal and non-formal music practice experiences ("I engaged in regular daily practice of a musical instrument including voice for *X* years"). Singing abilities questions inquire about individuals' singing skills and activities (e.g., "After hearing a new song two or three times I can usually sing it by myself"). Questions under the emotion category reflect on instances of active emotional responses to music (e.g., "I sometimes choose music that can trigger shivers down my spine"). These topics together consider an individual's holistic musical ability, including instances of formal and non-formal music training and engagement. The composite score of these subscales makes up an individual's general musical sophistication score. All items, except those assessing musical training, are scored on a seven-point Likert scale with choices that range from *completely disagree* to *completely agree* (Müllensiefen et al., 2014).

Results

Data analysis

Results from each procedure were analyzed using a mixed-effect analysis of variance. The analysis factors depended on the procedure, but all analyses included test group (cochlear implant users versus listeners with no known hearing loss) as a between-subject factor and test session (pre- versus post-training) as a within-subject factor. Planned comparisons were made between test group and test session for all assessment tasks to test whether musical interval identification training would improve the performance of the two groups on different musical tasks from pre- to post-training. Effect size was calculated using Cohen's method (Cohen, 1992) and significance levels using multiple comparisons with Bonferroni adjustments. Comparisons between individual results across measures were performed using Pearson's correlation coefficients.

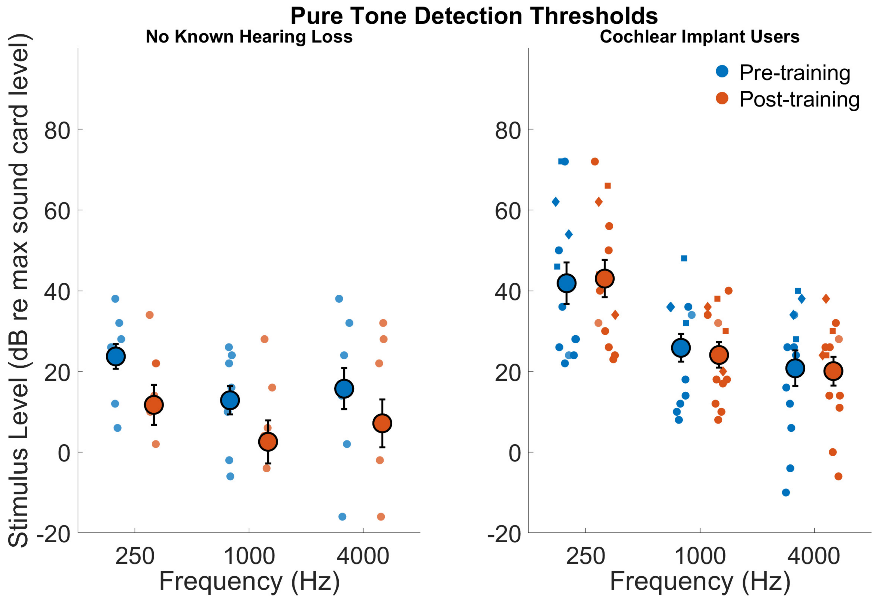

FIGURE 2
Stimulus level associated with detection threshold for 250, 1,000, and 4,000 Hz for those with no known hearing loss (left subpanel) and for cochlear implant users (right subpanel). The gain is in decibels with a gain of 100 dB corresponding to the maximum gain of the listening device. Smaller symbols indicate individual thresholds. Individual thresholds for CI users with implants from Cochlear Corporation are represented with a circle, Advanced Bionics with a diamond, and MED-EL with a square. Larger circles indicate group averages for each session with error bars indicating standard errors of the means.

Pure tone detection thresholds

Figure 2 shows the pure tone detection thresholds measured as a calibration procedure in dB relative to soundcard at 250, 1,000, and 4,000 Hz for those with no known hearing loss and for cochlear implant users. The difference in average detection thresholds between groups was significant exhibiting a large effect size ($F_{1,18} = 10.1$, $p = 0.005$, $d_{Cohen} = 1.5$) with cochlear implant users setting the average software volume higher (29.2 = 11.8) than those with no known hearing loss (12.3 = 10.4). Importantly, these thresholds are measured relative to the system volume that participants adjust their computers for at-home listening. These results are not indicative of absolute detection, but they do indicate that when participants adjust their computers and listening devices to be comfortable, cochlear implant users had elevated detection thresholds. It is important to note as well that relative to the self-selected comfortable listening level at 1,000 Hz, cochlear implant users had elevated detection thresholds, or a smaller relative dynamic range ($F_{1,18} = 3, 14$, $p = 0.09$, $d_{Cohen} = 0.7$). For relative detection thresholds, the effect of frequency was significant ($F_{2,36} = 17.3$, $p = 0.001$) as was the interaction between frequency and participant group ($F_{2,36} = 4.1$, $p = 0.024$). The interaction effect is evident in the particularly elevated thresholds at 250 Hz for the cochlear implant users. The effect of session (pre- versus post-training) was not significant ($F_{1,18} = 2.4$, $p = 0.14$) nor was the interaction between session and participant group ($F_{1,18} = 2.0$, $p = 0.17$). The interaction effect of frequency and session (pre- versus post-training) was not significant ($F_{2,36} = 0.09$, $p = 0.91$) nor was the interaction for frequency, session, and participant group ($F_{2,36} = 0.4$, $p = 0.68$).

Pure tone frequency discrimination

Figure 3 shows pure tone frequency discrimination for all participants before and after training. The cochlear implant users had poorer discrimination compared to those with no known hearing loss ($F_{1,18} = 12.84$, $p = 0.002$). Average discrimination thresholds across frequencies and sessions was 7.04% (or 1.18 semitones) for cochlear implant users and 1.05% (or 0.18 semitones) for those with no known hearing loss ($d_{Cohen} = 1.6$). There was a small effect of frequency ($F_{2,36} = 1.95$, $p = 0.09$) as well as a small effect for the interaction between frequency and participant group ($F_{2,36} = 2.15$, $p = 0.074$). The interaction effect can be seen in that discrimination improved with increasing frequency for those with no known hearing loss, but cochlear implant users had best discrimination near 1 kHz. The effect of test session was not significant ($F_{1,18} = 0.03$, $p = 0.87$) nor was the interaction between session and participant group ($F_{1,18} = 0.006$, $p = 0.94$). The interaction effect of frequency

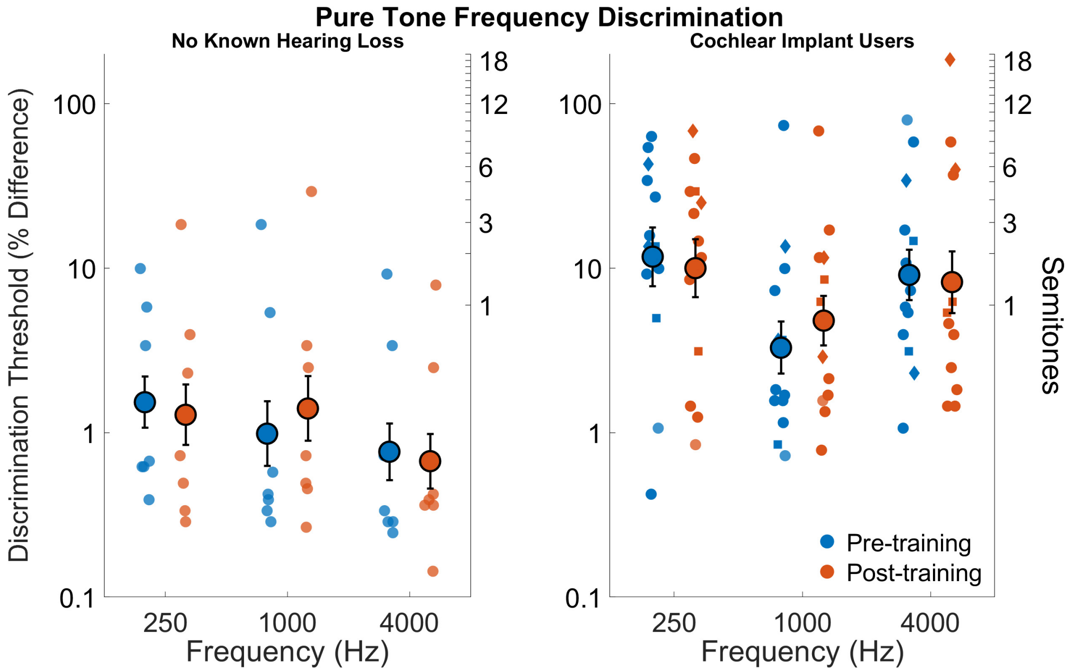

FIGURE 3
Pure tone frequency discrimination thresholds as percent difference on logarithmic scale (left y axis) and semitones (right y axis) for frequencies 250, 1,000, and 4,000 Hz for participants with no known hearing loss (left subpanel) and for cochlear implant users (right subpanel). Smaller symbols indicate individual thresholds. Individual thresholds for cochlear implant users with implants from Cochlear Corporation are represented with a circle, Advanced Bionics with a diamond, and MED-EL with a square. Larger circles indicate group averages for each session with error bars indicating standard errors of the means.

and session (pre- versus post-training) was not significant ($F_{2,36} = 1.63, p = 0.21$) nor was the interaction for frequency, session, and participant group ($F_{2,36} = 0.0003,\ p = 0.99$).

Fundamental frequency discrimination

Figure 4 shows fundamental frequency discrimination thresholds for all participants before and after training. The cochlear implant users had poorer discrimination compared to those with no known hearing loss ($F_{1,18} = 19.3, p = 0.001$). Average discrimination thresholds across frequencies and sessions was 11.8% (or 1.93 semitones) for cochlear implant users and 0.9% (or 0.16 semitones) for those with no known hearing loss ($d_{Cohen} = 2.1$). The effect of fundamental frequency was significant ($F_{2,36} = 8.6, p = 0.001$) as well the interaction between fundamental frequency and group ($F_{2,36} = 4.5, p = 0.017$). The effect of fundamental frequency is evident in that discrimination generally worsened with increasing fundamental frequency, which is more pronounced in the cochlear implant users. The effect of test session was not significant ($F_{1,18} = 2.0, p = 0.18$) nor was the interaction between session and participant group ($F_{1,18} = 0.33, p = 0.57$). Averaged across groups and conditions, the effect of training on discrimination was small but positive ($d_{Cohen} = 0.13$). The interaction effect of frequency and session (pre- versus post-training) was not significant ($F_{2,36} = 0.49, p = 0.62$) nor was the interaction for frequency, session, and participant group ($F_{2,36} = 0.07,\ p = 0.93$).

Tonal and rhythm comparisons

Figure 5 shows performance on tonal and rhythm comparisons for all participants before and after training. Cochlear implant users had poorer performance on tonal comparisons compared to those with no known hearing loss ($F_{1,18} = 13.2, p = 0.0019$). Average performance across sessions was 69.1% correct for cochlear implant users and 91.3% correct for those with no known hearing loss ($d_{Cohen} = 1.85$). The effect of test session was not significant ($F_{1,18} = 0.35, p = 0.56$) nor was the interaction between session and participant group ($F_{1,18} = 0.01, p = 0.92$). Neither group significantly improved on tonal comparisons across sessions.

Cochlear implant users also had poorer performance on rhythm comparisons compared to those with no known hearing loss ($F_{1,18} = 21.5, p = 0.001$). Average performance across sessions was 76.8% correct for cochlear implant users

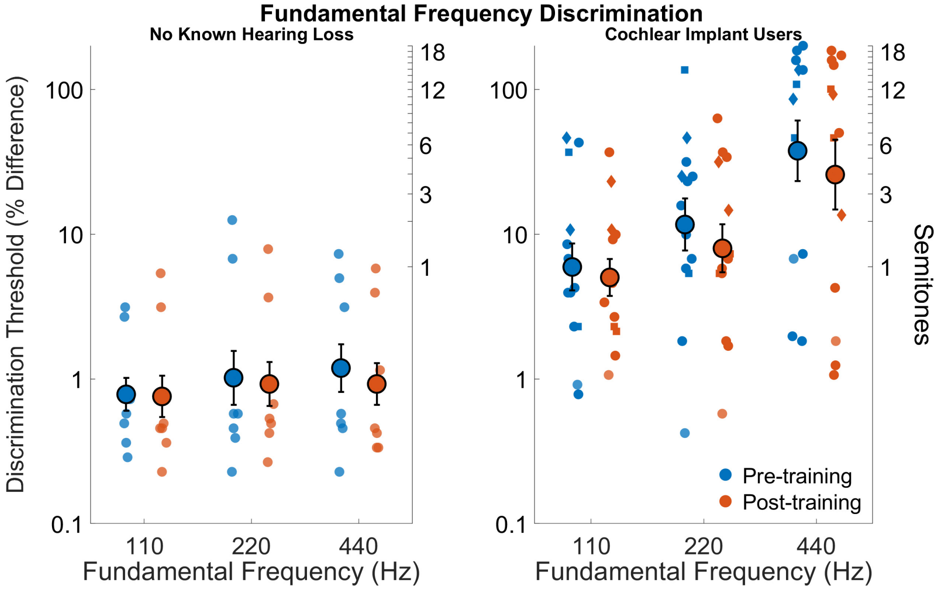

FIGURE 4
Fundamental frequency discrimination thresholds as percent difference on a logarithmic scale (left y axis) and semitones (right y axis) for fundamental frequencies 110, 220, and 440 Hz for participants with no known hearing loss (left subpanel) and for cochlear implant users (right subpanel). Smaller symbols indicate individual thresholds. Individual thresholds for cochlear implant users with implants from Cochlear Corporation are represented with a circle, Advanced Bionics with a diamond, and MED-EL with a square. Larger circles indicate group averages for each session with error bars indicating standard errors of the means.

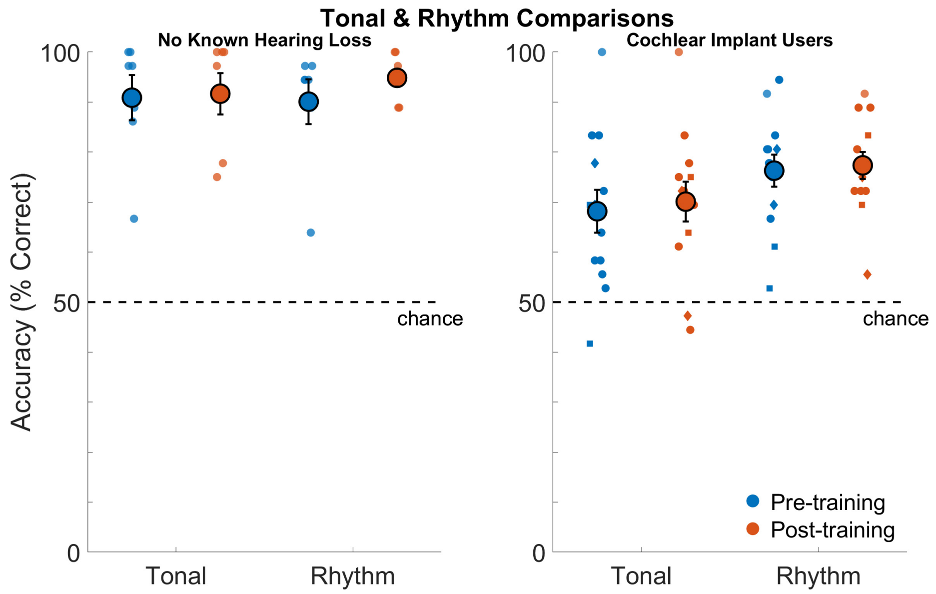

FIGURE 5
Tonal and rhythm comparisons as percentage of correct responses for listeners with no known hearing loss (left subpanel) and for cochlear implant users (right subpanel). Smaller symbols indicate individual thresholds. Individual thresholds for cochlear implant users with implants from Cochlear Corporation are represented with a circle, Advanced Bionics with a diamond, and MED-EL with a square. Larger circles indicate group averages for each session with error bars indicating standard errors of the means.

and 92.5% correct for those with no known hearing loss ($d_{Cohen} = 1.9$). The effect of test session was not significant ($F_{1,18} = 1.75, p = 0.2$) nor was the interaction between session and participant group ($F_{1,18} = 1.1, p = 0.31$). Neither group significantly improved on rhythm comparisons across sessions.

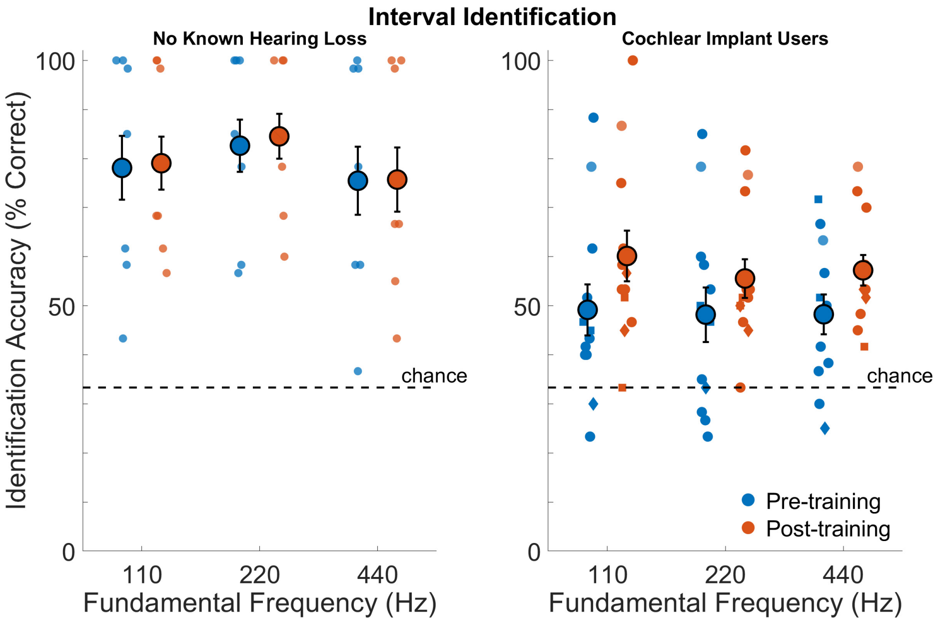

FIGURE 6
Interval identification as percentage of correct responses for participants with no known hearing loss (left subpanel) and for cochlear implant users (right subpanel) at 110, 220, and 440 Hz. Smaller symbols indicate individual thresholds. Individual thresholds for cochlear implant users with implants from Cochlear Corporation are represented with a circle, Advanced Bionics with a diamond, and MED-EL with a square. Larger circles indicate group averages for each session with error bars indicating standard errors of the means.

Interval identification

Figure 6 shows performance on interval identification for all participants before and after training. Cochlear implant users had poorer interval identification compared to those with no known hearing loss ($F_{1,18} = 9.0, p = 0.009$). Average performance across sessions was 52.4% correct for cochlear implant users and 79.2% correct for those with no known hearing loss ($d_{Cohe} = 1.5$). There was a no effect of frequency ($F_{2,30} = 2.05, p = 0.15$) but a small effect for the interaction between frequency and participant group ($F_{2,30} = 2.94, p = 0.068$). The effect of test session was not significant ($F_{1,18} = 3.6, p = 0.076$) nor was the interaction between session and participant group ($F_{1,18} = 2.0, p = 0.17$). Planned comparisons of the performance before and after training indicated that, on average, the cochlear implant users improved from 48.6 to 58.2% correct ($d_{Cohen} = 0.63$). The interaction effect of frequency and session (pre- versus post-training) was not significant ($F_{2,30} = 0.2, p = 0.82$) nor was the interaction for frequency, session, and participant group ($F_{2,30} = 0.6,\ p = 0.56$).

Correlation analysis

Correlations were calculated between results from different procedures based on averages across conditions. Correlations were calculated for all participants (Table 3) and for the two participant groups separately [Table 4 (no known hearing loss) and Table 5 (cochlear implant)]. While the current measures of statistical significance for these tables are $p = 0.05$ (*), $p = 0.01$ (**), $p = 0.0024$ (^{x}) and $p = 0.001$ (***), only the correlations with $p = 0.0024$ (^{x}) or $p = 0.001$ (***) were statistically significant for the stringent Bonferroni-adjusted criteria which adjusts alpha from

TABLE 3 Correlations between results from different procedures averaged across conditions.

	FDT	F0DT	TC	RC	II	MSI
DT	0.52*	0.55*	0.41	0.59**	0.46*	0.52*
FDT		0.94***	0.81***	0.82***	0.92***	0.72***
F0DT			0.88***	0.88***	0.93***	0.73***
TC				0.90***	0.86***	0.75***
RC					0.82***	0.70***
II						0.75***

For clarity, only the correlation magnitudes are displayed, but all comparisons were congruent in that better performance on one measure corresponded with better performance on another. Correlation coefficients and p-values associated with p-values less than 0.05 are emboldened. DT, detection thresholds; FDT, frequency discrimination thresholds; F0DT, fundamental frequency discrimination thresholds; TC, tonal comparisons; RC, rhythm comparisons; II, interval identification; MSI, musical sophistication index; $p = 0.05$ (*), $p = 0.01$ (**), $p = 0.0024$ (^{x}) and $p = 0.001$ (***). Note that only the correlations with $p = 0.0024$ (^{x}) or $p = 0.001$ (***) were statistically significant for the stringent Bonferroni-adjusted criteria which adjusts alpha from 0.05 to 0.05/21 or 0.0024.

TABLE 4 As for Table 3 but only including those with no known hearing loss.

	FDT	F0DT	TC	RC	II	MSI
DT	0.14	0.03	0.03	0.21	0.14	0.35
FDT		0.97***	0.88**	0.83*	0.91**	0.79*
F0DT			0.83*	0.81*	0.86*	0.69
TC				0.94^{x}	0.86*	0.68
RC					0.83*	0.61
II						0.93^{x}

TABLE 5 As for Table 3 but only including cochlear implant users.

	FDT	F0DT	TC	RC	II	MSI
DT	0.26	0.28	0.04	0.51	0.15	0.35
FDT		0.85***	0.58*	0.62*	0.84***	0.55
F0DT			0.77^{x}	0.76**	0.94***	0.63*
TC				0.77^{x}	0.74**	0.67*
RC					0.63*	0.61*
II						0.55

0.05 to 0.05/21 or 0.0024. Considering the correlations for all participants in Table 3, all correlations, except between detection thresholds and tonal comparisons ($p = 0.07$), were significant indicating the general trend that the best performing participants were consistent across procedures. While detection thresholds were correlated with other measures, the explained variance was not as high as for the other comparisons. These correlations with detection thresholds were likely driven by group effects with cochlear implant users having elevated detection thresholds and consistently poorer performance on other measures. This notion is supported by the fact that none of the within-group correlations were significant for comparisons with detection thresholds. The low-level measures of pure tone and fundamental frequency discrimination were highly correlated with the higher-level measures of tonal and rhythm comparisons and interval identification. The strength of these correlations generally held when considering correlations within each participant group. For cochlear implant users, both pure tone and fundamental frequency discrimination were particularly well correlated to interval identification. The strong relationship between frequency discrimination and interval identification suggests that training on one of these dimensions could strengthen the other, although it is important to note that no training effects were found in this study. While fundamental frequency discrimination produced the highest correlation with interval identification, the other assessments were all significantly correlated with interval identification as well. Multiple regression analyses were calculated to determine which pairs of assessments including an interaction term provided the highest joint correlation with interval identification. The highest correlation was observed between interval identification with a multiple regression analysis of fundamental frequency discrimination and MSI scores, which produced a correlation coefficient of 0.97 when the interaction between measures was included and 0.94 when the interaction was not modeled. In general, the correlation between assessments were strongly interdependent (additional variance was not well explained by combining measures), with the most notable exception that jointly modeling MSI scores and fundamental frequency discrimination produced the largest correlation.

As an example of specific relationships, Figure 7 compares performance on pure tone frequency discrimination, tonal and rhythm comparisons, and interval identification with fundamental frequency discrimination. Participants who had better fundamental frequency discrimination for complex tones tended to have better performance on all other measures. As a second example of specific relationships, Figure 8 compares performance on detection thresholds, pure tone frequency discrimination, fundamental frequency discrimination, tonal comparisons, rhythm comparisons, and interval identification with MSI scores. Participants who had higher MSI scores—in particular, those with normal hearing—tended to have better performance on all other measures.

Details of the training program

Figure 9 shows the number of cumulative failed runs during training across difficulty levels for individual participants. The purpose of reporting training progress in terms of cumulative fails was to highlight which subjects had the most difficulty completing the training task at specific difficulty training levels and overall. Subjects H2, H6, and H7 had perfect performance on all difficulty training levels, so their data points overlap and only H7 is visible. Subject H5 had impressive performance as well. These four subjects were all accomplished musicians, which is reflected in their exceptional performance. The other three subjects (H1, H3, and H4) were non-musicians, with H4 struggling the most in this group. Subject C2 was a musician from a young age before getting the cochlear implant, which may have contributed to the great performance. Subjects C16 and C32 were both avid musicians who passed most difficulty training levels with ease. Subject C22 was an accomplished musician but also a bimodal listener who has not had much focused rehabilitation of the cochlear implant alone, which may have contributed to the difficulty getting past even the first level.

Discussion

The primary aim of this study was to characterize performance on assessment tasks for cochlear implant users

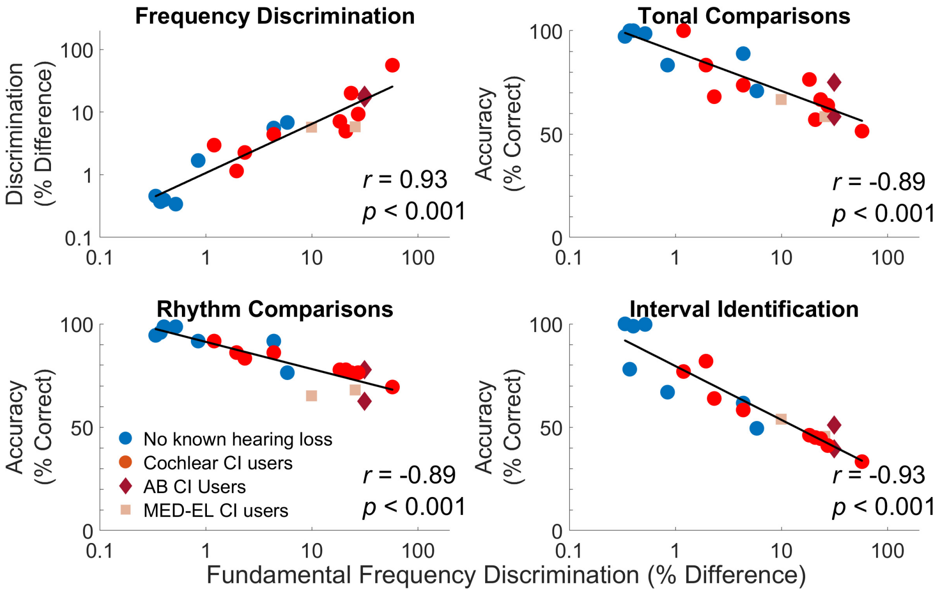

FIGURE 7
Comparisons of individual results from different procedures based on averages across conditions. For each comparison, each symbol represents the average measure for each individual participant averaged across conditions and repetitions. Individual thresholds for cochlear implant users with implants from Cochlear Corporation are represented with a circle (red), Advanced Bionics with a diamond (dark red), and MED-EL with a square (light red).

and listeners with no known hearing loss before and after 2 weeks of online musical interval training. Pre-training and post-training assessments measured pure tone and fundamental frequency discrimination, tonal and rhythm comparisons, and interval identification. The overarching hypothesis motivating this study is that both low-level psychophysical access to pitch cues as well as higher-level labeling of intervals limits identification accuracy in cochlear implant users, and, to a certain extent, those with no known hearing loss. Strong correlations were found between low-level measures of frequency and fundamental frequency resolution with higher-level rhythm and tonal comparisons, interval identification, and musical sophistication, thus supporting the first part of the overarching hypothesis. Furthermore, dedicated training on interval identification during this study provided cochlear implant participants opportunity to build (or rebuild) the association between interval and naming convention, along with experience with assessment tasks requiring pitch judgments.

The strength of the relationship between interval identification and frequency discrimination is well explained by separating the skills needed to perform interval identification into two components. The listener must first, hear the difference in pitch between two successive notes and second, label the magnitude of the pitch difference with the corresponding interval. Challenged in this way, participants use increasingly fine distinctions between interval magnitudes to determine the interval label. It was surprising then that a few listeners with no known hearing loss had pitch resolution at or worse than 1 semitone (note, one semitone is approximately a 6% difference in fundamental frequency). This could have been a function of age ($p < 0.02$ for correlations with PT, F0, II, and MSI), experience, unknown hearing loss, or even attention. Most cochlear implant users had pitch resolution worse than two semitones, and although age was not a factor. This poor resolution makes it difficult to form magnitude judgments, except for stark interval comparisons such as a major 2nd versus an octave. One cochlear implant user, who had pitch resolution better than a semitone, was able to correctly label 80% of intervals on the assessment task. While it was not guaranteed that the higher-level task of interval labeling would directly influence performance on lower-level psychoacoustic tasks in this brief training, given the strength of the relationship between interval identification and frequency resolution, it is possible that more extensive practice at interval labeling may transfer to simpler tasks such as pitch ranking and melodic contour identification, although this study did not find any evidence for this claim. It has also been shown that incidental listening to musical materials can improve resolution of those materials (Little et al., 2019).

The absence of significant learning in both participants groups should be taken into consideration when evaluating the effectiveness of training strategies. It has been proposed that auditory perceptual learning requires both stimulus exposure

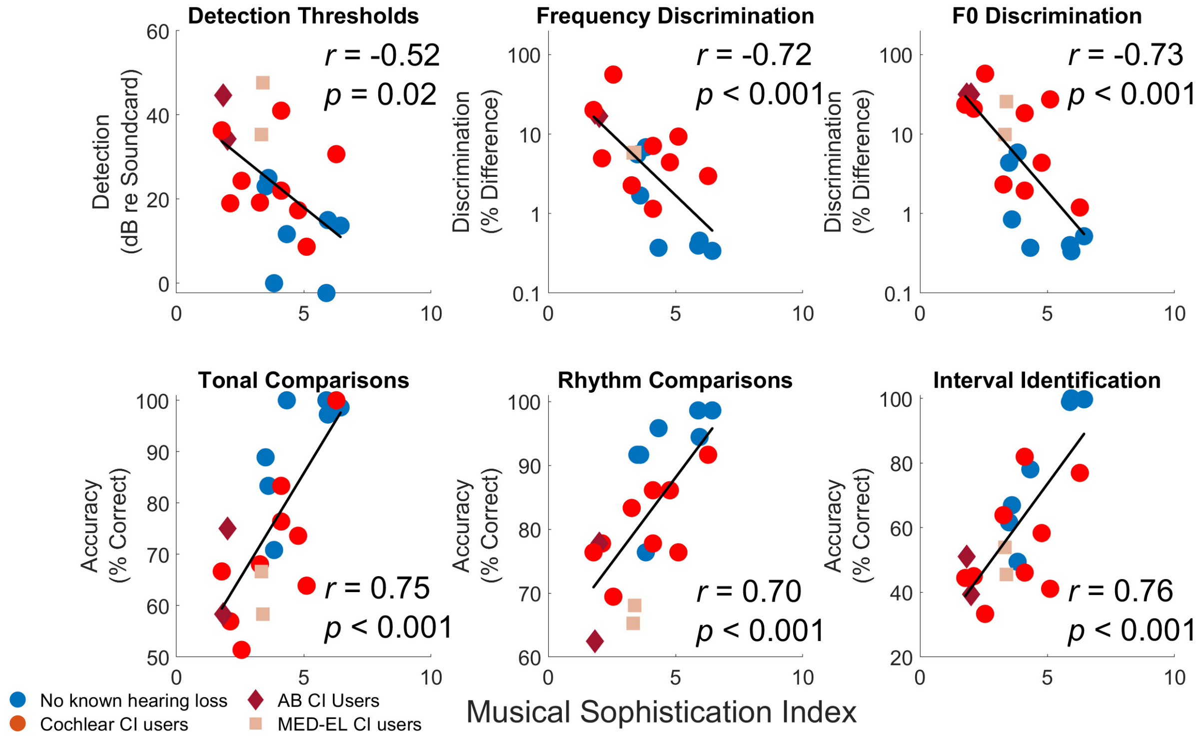

FIGURE 8

Comparisons of Musical Sophistication Index (MSI) with individual results from different procedures based on averages across conditions. For each comparison, each symbols represents the average measure for each individual participant averaged across conditions and repetitions. Individual thresholds for cochlear implant users with implants from Cochlear Corporation are represented with a circle (red), Advanced Bionics with a diamond (dark red), and MED-EL with a square (light red).

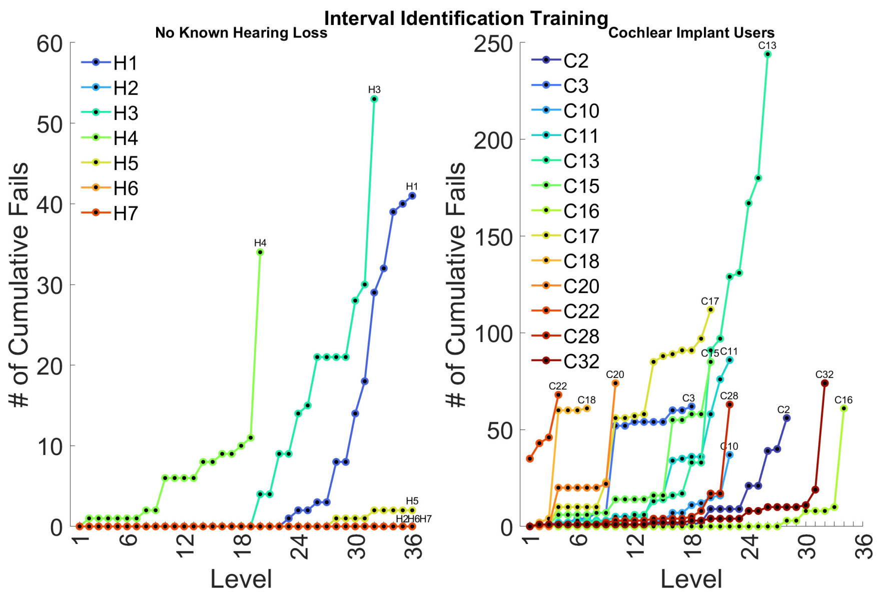

FIGURE 9

Number of cumulative failed runs across levels for individual participants. Each line ends when the participant completed 2 weeks of training or reached the final level of the training program. Note, the ordinate has a different scale for the two participant groups. The assessment conditions used for interval identification correspond to training levels of 20, 22, and 24.

and execution of the task to be learned—provided in the current study by task practice—and a sufficient amount of practice per day (Wright, 2013). These requirements for learning must be balanced with common barriers to training paradigm success—fatigue and attrition. Studies of computer-based auditory training programs for individuals with hearing loss have varying definitions of retention and many studies do not report their compliance level (Henshaw and Ferguson, 2013). The present study aimed to make musical interval training accessible and convenient by providing an online training program that participants could use at home and by limiting training sessions to 20 min per day for 2 weeks. This is a relatively brief training protocol compared to other training programs for cochlear implant users (Looi et al., 2012). While this brief period of training likely contributed to the 100% retention rate, it may not have provided enough practice needed for learning, leading to the lack of improved performance on the trained task. Further investigation into musical interval identification training in cochlear implant users is necessary to clarify the optimal amount of daily and total training needed for learning.

An additional consideration is the difficulty of using the online interface given the age of participants in the cochlear implant user group. Technological literacy is generally lower among older populations and the mean age of the cochlear implant users was 62.9 years compared to 42.3 years for the listeners with no known hearing loss. Multiple participants reported difficulty using the online interface throughout the study. This may have made learning through the online interface difficult and training sessions may not have progressed as intended. While age did not significantly correlate with performance in the cochlear implant group, it did for the group with no known hearing loss for pure tone frequency discrimination ($p = 0.014$), fundamental frequency discrimination ($p = 0.017$), and interval identification ($p = 0.001$).

The assessment used for interval identification may also have been too difficult for the cochlear implant users. The conditions for the assessment procedure presented the participants with three types of intervals (major 3rd, perfect 5th, and octave) over three root note frequency ranges (octave ranges centered on A2, A3, and A4). These conditions correspond to training levels 20, 22, and 24 of the training program. Many participants in the cochlear implant group did not progress beyond level 22 within the 2-week training period. Therefore, one explanation for the lack of improvement in musical interval identification after training is that some participants may have only been exposed to easier levels of musical interval identification.

Our training protocol required participants to learn a difficult task in a brief amount of time. Musical intervals are a relatively abstract concept and represent the pitch ratio, a concept that is difficult to grasp without prior musical training. Given that it is well known that interval labeling is a skill that cannot be learned without dedicated musical training, a control group of participants who did not train on interval labeling was not included. It is possible that task familiarity had a small impact on participant performance that cannot be assessed without a control group, since tasks in the pre- and post-training assessments were identical. However, task familiarity is unlikely to have contributed significantly in this study given that there were no significant improvements in performance found across sessions. Furthermore, the interval labeling task was chosen due to its challenging nature, requiring participants to attend to multiple musical interval stimuli in order to progress through the difficulty training levels. Studies have suggested that an auditory task must be sufficiently difficult to result in learning since adequate amounts of attention is a requirement of learning, but there is evidence that exceptionally difficult tasks can still facilitate perceptual learning (Amitay et al., 2006; Moore and Amitay, 2007). However, the extent that task difficulty limits the higher-level labeling aspect of interval identification is poorly understood.

Musical interval identification also requires a listener to distinguish between two pitches and many listeners without prior musical training have poor resolution. McDermott et al. (2010) demonstrated that normal hearing non-musicians and even some amateur musicians had pitch interval thresholds greater than a semitone for pure and complex tone conditions. They found that interval resolution was up to 8 times worse than frequency resolution, indicating that the frequency resolution necessary to discriminate between intervals of one semitone difference in width (e.g., minor second vs major second) may need to be better than 1 semitone.

Even poorer pitch resolution is demonstrated in cochlear implant users (e.g., Pretorius and Hanekom, 2008; Goldsworthy, 2015). The ability to distinguish between two pitches is affected by the cues (temporal and place-of-excitation) provided by the processor for different stimuli (see **Figure 1** for representative encoding of musical notes). Cochlear Corporation (9/13 subjects) generally discards temporal fine structure while providing temporal cues through F0 envelope modulation, MED-EL (2/13 subjects) and Advanced Bionics (2/13 subjects) attempt to encode more temporal cues through their processors, especially at lower frequencies (Arnoldner et al., 2007; Gazibegovic et al., 2010; Wouters et al., 2015). Swanson et al. (2019) unpacked the temporal and place-of-excitation cues for pure tones and harmonic complexes for the Cochlear Corporation signal processing strategy. They showed that pure tones provide only place cues to pitch, with the filter bandwidth at different frequencies having a substantial effect on the pitch resolution. For pure tones near 1,000 Hz, variation in pure tone frequency will produce variation in the relative amplitude of two neighboring filters, hence variation in currents on neighboring electrodes. For pure tones near 250 Hz, this mechanism does not work as well because the lowest filter is centered at 250 Hz, so there is no lower neighbor. For pure tones near 4,000 Hz, the filters are much wider, and if two tones are both within one filter passband, then there may be little difference in the two corresponding stimulation patterns. This may explain the general pattern of results in **Figure 3** with poor resolution at both lower (<250 Hz) and higher (>4 kHz) frequencies and better resolution between 250 and 4 kHz (Pretorius and Hanekom, 2008). Our rationale for using pure tones of 250, 1,000, and 4,000 Hz is to broadly characterize spectral resolution as conveyed by place-pitch cues across the electrode array. Pure tones primarily provide place-pitch cues with the exception of strategies that attempt to provide timing cues for pure tones. While MED-EL and Advanced Bionics would attempt to provide temporal cues for 250 Hz pure tones, it does not appear to have broad effect on individual performance in **Figure 3**. Harmonic complexes below 220 Hz would have good temporal cues provided by amplitude modulation at the fundamental because the individual harmonics are not resolved by the ACE filter bank, between 220 and 440 would have a mixture of the two cues, and above 440 Hz would have only place cues because the individual harmonics are resolved (Swanson et al., 2019). The results of **Figure 4** suggest that subjects may have been more sensitive to temporal pitch cues than place pitch cues. Interval identification was done with musical piano notes to provide the richest encoding of musical tonality (von Helmholtz, 1885; Siedenburg et al., 2019). The cues provided by these notes varied based on frequency, with trials presenting place, temporal, or a mixture of cues. Although the higher-level assessments could have been designed with stimuli to isolate a single pitch cue, as was done by Vandali et al. (2015), the present study focused on providing musical notes with the most potential cues for pitch and interval judgments, leaving the cues chosen up to each subject's clinical signal processing strategy.

Considering the relationships between pitch resolution and music perception, this study demonstrates that pure tone frequency and fundamental frequency discrimination are both highly correlated with musical interval identification. This correlation is anticipated given that a musical interval is comprised of two different pitches. These correlations suggest that improving access to low-level cues for pure tone frequency and/or fundamental frequency perception could improve higher-level musical abilities. To improve perception of complex listening situations and musical perception for cochlear implant users, future signal processing strategies should improve access to stimulation cues that support pitch perception, whether that be through better coding of place-of-excitation cues, better coding of temporal modulation cues, or a synergy of these two (Fu and Shannon, 1999; Leigh et al., 2004; Laneau et al., 2006; Arnoldner et al., 2007; Riss et al., 2008, 2014, 2016; Stohl et al., 2008; Firszt et al., 2009; Lorens et al., 2010; Vermeire et al., 2010; Luo et al., 2012; Müller et al., 2012; Grasmeder et al., 2014; Francart et al., 2015; Rader et al., 2016; Erfanian Saeedi et al., 2017). Concomitant with better signal processing, structured aural rehabilitation programs should be designed to reintroduce cochlear implant users to the subtle stimulation cues for pitch perception. Given the correlations of low-level pitch perception with higher-level musical perception tasks, improvement in signal processing and dedicated aural rehabilitation will likely improve musical enjoyment and appreciation for cochlear implant users.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the University of Southern California Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JO, HG, and RG: conceptualization and visualization. SB, JO, HG, and RG: methodology, software, and validation. SB: formal analysis and data curation. SB and JO: investigation, resources, project administration, and writing—original draft preparation. SB, JO, and RG: writing—review and editing. RG: supervision and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding

This study was sponsored by the National Institutes of Health through the National Institute on Deafness and Other Communication Disorders (R01 DC018701).

Acknowledgments

We thank our Cochlear implant and normal hearing subjects who worked tirelessly on testing and training.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.903924/full#supplementary-material

References

Amitay, S., Irwin, A., and Moore, D. R. (2006). Discrimination learning induced by training with identical stimuli. *Nat. Neurosci.* 9, 1446–1448. doi: 10.1038/nn1787

Arnoldner, C., Riss, D., Brunner, M., Durisin, M., Baumgartner, W.-D., and Hamzavi, J.-S. (2007). Speech and music perception with the new fine structure speech coding strategy: preliminary results. *Acta Otolaryngol.* 127, 1298–1303. doi: 10.1080/00016480701275261

Caldwell, M. T., Jiam, N. T., and Limb, C. J. (2017). Assessment and improvement of sound quality in cochlear implant users. *Laryngoscope Investig. Otolaryngol.* 2, 119–124. doi: 10.1002/lio2.71

Camarena, A., Manchala, G., Papadopoulos, J., O'Connell, S. R., and Goldsworthy, R. L. (2022). Pleasantness ratings of musical dyads in cochlear implant users. *Brain Sci.* 12:33. doi: 10.3390/brainsci12010033

Cohen, J. (1992). Statistical power analysis. *Curr. Dir. Psychol. Sci.* 1, 98–101. doi: 10.1111/1467-8721.ep10768783

Crew, J. D., Galvin, J. J., and Fu, Q.-J. (2012). Channel interaction limits melodic pitch perception in simulated cochlear implants. *J. Acoust. Soc. Am.* 132, EL429–EL435. doi: 10.1121/1.4758770

Dhanasingh, A., and Jolly, C. (2017). An overview of cochlear implant electrode array designs. *Hear. Res.* 356, 93–103. doi: 10.1016/j.heares.2017.10.005

do Nascimento, L. T., and Bevilacqua, M. C. (2005). Evaluation of speech perception in noise in cochlear implanted adults. *Braz. J. Otorhinolaryngol.* 71, 432–438. doi: 10.1016/S1808-8694(15)31195-2

Drennan, W. R., and Rubinstein, J. T. (2008). Music perception in cochlear implant users and its relationship with psychophysical capabilities. *J. Rehabil. Res. Dev. Wash.* 45, 779–789. doi: 10.1682/jrrd.2007.08.0118

Drennan, W. R., Longnion, J. K., Ruffin, C., and Rubinstein, J. T. (2008). Discrimination of schroeder-phase harmonic complexes by normal-hearing and cochlear-implant listeners. *J. Assoc. Res. Otolaryngol.* 9, 138–149. doi: 10.1007/s10162-007-0107-6

Erfanian Saeedi, N., Blamey, P. J., Burkitt, A. N., and Grayden, D. B. (2017). An integrated model of pitch perception incorporating place and temporal pitch codes with application to cochlear implant research. *Hear. Res.* 344, 135–147. doi: 10.1016/j.heares.2016.11.005

Fetterman, B. L., and Domico, E. H. (2002). Speech recognition in background noise of cochlear implant patients. *Otolaryngol. Head Neck Surg.* 126, 257–263. doi: 10.1067/mhn.2002.123044

Finley, C. C., Holden, T. A., Holden, L. K., Whiting, B. R., Chole, R. A., Neely, G. J., et al. (2008). Role of electrode placement as a contributor to variability in cochlear implant outcomes. *Otol. Neurotol.* 29, 920–928. doi: 10.1097/MAO.0b013e318184f492

Firszt, J. B., Holden, L. K., Reeder, R. M., and Skinner, M. W. (2009). Speech recognition in cochlear implant recipients: comparison of standard HiRes and HiRes 120 sound processing. *Otol. Neurotol.* 30:146. doi: 10.1097/MAO.0b013e3181924ff8

Francart, T., Osses, A., and Wouters, J. (2015). Speech perception with F0mod, a cochlear implant pitch coding strategy. *Int. J. Audiol.* 54, 424–432. doi: 10.3109/14992027.2014.989455

Fu, Q. J., and Shannon, R. V. (1999). Effects of electrode configuration and frequency allocation on vowel recognition with the Nucleus-22 cochlear implant. *Ear Hear.* 20, 332–344. doi: 10.1097/00003446-199908000-00006

Fu, Q.-J., and Nogaki, G. (2005). Noise susceptibility of cochlear implant users: the role of spectral resolution and smearing. *J. Assoc. Res. Otolaryngol.* 6, 19–27. doi: 10.1007/s10162-004-5024-3

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., and Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *J. Cogn. Neurosci.* 16, 1010–1021. doi: 10.1162/0898929041502706

Gazibegovic, D., Arnold, L., Rocca, C., and Boyle, P. (2010). Evaluation of music perception in adult users of HiRes® 120 and previous generations of advanced Bionics® Sound coding strategies. *Cochlear Implants Int.* 11, 296–301. doi: 10.1179/146701010X12671177989354

Gfeller, K. (2001). Aural rehabilitation of music listening for adult cochlear implant recipients: addressing learner characteristics. *Music Ther. Perspect.* 19, 88–95. doi: 10.1093/mtp/19.2.88

Gfeller, K., Christ, A., Knutson, J. F., Witt, S., Murray, K. T., and Tyler, R. S. (2000). Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *J. Am. Acad. Audiol.* 11, 390–406.

Gfeller, K., Olszewski, C., Rychener, M., Sena, K., Knutson, J. F., Witt, S., et al. (2005). Recognition of "real-world" musical excerpts by cochlear implant recipients and normal-hearing adults. *Ear Hear.* 26, 237–250. doi: 10.1097/00003446-200506000-00001

Gfeller, K., Turner, C., Mehr, M., Woodworth, G., Fearn, R., Knutson, J. F., et al. (2002a). Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implants Int.* 3, 29–53. doi: 10.1002/cii.50

Gfeller, K., Turner, C., Oleson, J., Zhang, X., Gantz, B., Froman, R., et al. (2007). Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise. *Ear Hear.* 28, 412–423. doi: 10.1097/AUD.0b013e3180479318

Gfeller, K., Witt, S., Adamek, M., Mehr, M., Rogers, J., Stordahl, J., et al. (2002b). Effects of training on timbre recognition and appraisal by postlingually deafened cochlear implant recipients. *J. Am. Acad. Audiol.* 13, 132–145.

Goldsworthy, R. L. (2015). Correlations between pitch and phoneme perception in cochlear implant users and their normal hearing peers. *J. Assoc. Res. Otolaryngol.* 16, 797–809. doi: 10.1007/s10162-015-0541-9

Goldsworthy, R. L., Delhorne, L. A., Braida, L. D., and Reed, C. M. (2013). Psychoacoustic and phoneme identification measures in cochlear-implant and normal-hearing listeners. *Trends Amplif.* 17, 27–44. doi: 10.1177/1084713813477244

Grasmeder, M. L., Verschuur, C. A., and Batty, V. B. (2014). Optimizing frequency-to-electrode allocation for individual cochlear implant users. *J. Acoust. Soc. Am.* 136, 3313–3324. doi: 10.1121/1.4900831

Habibi, A., Cahn, B. R., Damasio, A., and Damasio, H. (2016). Neural correlates of accelerated auditory processing in children engaged in music training. *Dev. Cogn. Neurosci.* 21, 1–14. doi: 10.1016/j.dcn.2016.04.003

Habibi, A., Wirantana, V., and Starr, A. (2013). Cortical activity during perception of musical pitch: comparing musicians and nonmusicians. *Music Percept. Interdiscip. J.* 30, 463–479. doi: 10.1525/mp.2013.30.5.463

Habibi, A., Wirantana, V., and Starr, A. (2014). Cortical activity during perception of musical rhythm; comparing musicians and non-musicians. *Psychomusicology* 24, 125–135. doi: 10.1037/pmu0000046

Henshaw, H., and Ferguson, M. A. (2013). Efficacy of individual computer-based auditory training for people with hearing loss: a systematic review of the evidence. *PLoS One* 8:e62836. doi: 10.1371/journal.pone.0062836

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Percept. Psychophys.* 49, 227–229. doi: 10.3758/BF03214307

Kong, Y.-Y., Cruz, R., Jones, J. A., and Zeng, F.-G. (2004). Music perception with temporal cues in acoustic and electric hearing. *Ear Hear.* 25, 173–185. doi: 10.1097/01.aud.0000120365.97792.2f

Landsberger, D. M., Svrakic, M., Roland, J. T. J., and Svirsky, M. (2015). The relationship between insertion angles, default frequency allocations, and spiral ganglion place pitch in cochlear implants. *Ear Hear.* 36, e207–13. doi: 10.1097/AUD.0000000000000163

Laneau, J., Wouters, J., and Moonen, M. (2006). Improved music perception with explicit pitch coding in cochlear implants. *Audiol. Neurotol.* 11, 38–52. doi: 10.1159/000088853

Leigh, J. R., Henshall, K. R., and McKay, C. M. (2004). Optimizing frequency-to-electrode allocation in cochlear implants. *J. Am. Acad. Audiol.* 15, 574–584. doi: 10.3766/jaaa.15.8.5

Limb, C. J., and Roy, A. T. (2014). Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hear. Res.* 308, 13–26. doi: 10.1016/j.heares.2013.04.009

Limb, C. J., and Rubinstein, J. T. (2012). Current research on music perception in cochlear implant users. *Otolaryngol. Clin. North Am.* 45, 129–140. doi: 10.1016/j.otc.2011.08.021

Little, D. F., Cheng, H. H., and Wright, B. A. (2019). Inducing musical-interval learning by combining task practice with periods of stimulus exposure alone. *Atten. Percept. Psychophys.* 81, 344–357. doi: 10.3758/s13414-018-1584-x

Looi, V., and She, J. (2010). Music perception of cochlear implant users: a questionnaire, and its implications for a music training program. *Int. J. Audiol.* 49, 116–128. doi: 10.3109/14992020903405987

Looi, V., Gfeller, K., and Driscoll, V. (2012). MUsic appreciation and training for cochlear implant recipients: a review. *Semin. Hear.* 33, 307–334. doi: 10.1055/s-0032-1329222

Lorens, A., Zgoda, M., Obrycka, A., and Skarżynski, H. (2010). Fine structure processing improves speech perception as well as objective and subjective benefits in pediatric MED-EL COMBI 40+ users. *Int. J. Pediatr. Otorhinolaryngol.* 74, 1372–1378. doi: 10.1016/j.ijporl.2010.09.005

Luo, X., and Warner, B. (2020). Effect of instrument timbre on musical emotion recognition in normal-hearing listeners and cochlear implant users. *J. Acoust. Soc. Am.* 147:EL535. doi: 10.1121/10.0001475

Luo, X., Masterson, M. E., and Wu, C.-C. (2014). Melodic interval perception by normal-hearing listeners and cochlear implant users. *J. Acoust. Soc. Am.* 136, 1831–1844. doi: 10.1121/1.4894738

Luo, X., Padilla, M., and Landsberger, D. M. (2012). Pitch contour identification with combined place and temporal cues using cochlear implants. *J. Acoust. Soc. Am.* 131, 1325–1336. doi: 10.1121/1.3672708

Luo, X., Soslowsky, S., and Pulling, K. R. (2019). Interaction between pitch and timbre perception in normal-hearing listeners and cochlear implant users. *J. Assoc. Res. Otolaryngol.* 20, 57–72. doi: 10.1007/s10162-018-00701-3

Mangado, N., Pons-Prats, J., Coma, M., Mistrík, P., Piella, G., Ceresa, M., et al. (2018). Computational evaluation of cochlear implant surgery outcomes accounting for uncertainty and parameter variability. *Front. Physiol.* 9:498. doi: 10.3389/fphys.2018.00498

McDermott, H. J. (2004). Music perception with cochlear implants: a review. *Trends Amplif.* 8, 49–82. doi: 10.1177/108471380400800203

McDermott, J. H., Keebler, M. V., Micheyl, C., and Oxenham, A. J. (2010). Musical intervals and relative pitch: frequency resolution, not interval resolution, is special. *J. Acoust. Soc. Am.* 128, 1943–1951. doi: 10.1121/1.3478785

Moore, D., and Amitay, S. (2007). Auditory training: rules and applications. *Semin. Hear.* 28, 099–109. doi: 10.1055/s-2007-973436

Müllensiefen, D., Gingras, B., Musil, J., and Stewart, L. (2014). The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PLoS One* 9:e89642. doi: 10.1371/journal.pone.0089642

Müller, J., Brill, S., Hagen, R., Moeltner, A., Brockmeier, S., Stark, T., et al. (2012). Clinical trial results with the MED-EL fine structure processing coding strategy in experienced cochlear implant users. *ORL J. Otorhinolaryngol. Relat. Spec.* 74, 185–198. doi: 10.1159/000337089

Nimmons, G. L., Kang, R. S., Drennan, W. R., Longnion, J., Ruffin, C., Worman, T., et al. (2008). Clinical assessment of music perception in cochlear implant listeners. *Otol. Neurotol.* 29, 149–155. doi: 10.1097/mao.0b013e31812f7244

Nogueira, W., Schurzig, D., Büchner, A., Penninger, R. T., and Würfel, W. (2016). Validation of a cochlear implant patient-specific model of the voltage distribution in a clinical setting. *Front. Bioeng. Biotechnol.* 4:84. doi: 10.3389/fbioe.2016.00084

Penninger, R. T., Chien, W. W., Jiradejvong, P., Boeke, E., Carver, C. L., and Limb, C. J. (2013). Perception of pure tones and iterated rippled noise for normal hearing and cochlear implant users. *Trends Amplif.* 17, 45–53. doi: 10.1177/1084713813482759

Pijl, S. (1997). Labeling of musical interval size by cochlear implant patients and normally hearing subjects. *Ear Hear.* 18, 364–372. doi: 10.1097/00003446-199710000-00002

Pijl, S., and Schwarz, D. W. F. (1995a). Intonation of musical intervals by musical intervals by deaf subjects stimulated with single bipolar cochlear implant electrodes. *Hear. Res.* 89, 203–211. doi: 10.1016/0378-5955(95)00138-9

Pijl, S., and Schwarz, D. W. F. (1995b). Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes. *J. Acoust. Soc. Am.* 98, 886–895. doi: 10.1121/1.413514

Pretorius, L. L., and Hanekom, J. J. (2008). Free field frequency discrimination abilities of cochlear implant users. *Hear. Res.* 244, 77–84. doi: 10.1016/j.heares.2008.07.005

Rader, T., Döge, J., Adel, Y., Weissgerber, T., and Baumann, U. (2016). Place dependent stimulation rates improve pitch perception in cochlear implantees with single-sided deafness. *Hear. Res.* 339, 94–103. doi: 10.1016/j.heares.2016.06.013

Rebscher, S. J. (2008). Considerations for design of future cochlear implant electrode arrays: electrode array stiffness, size. *J. Rehabil. Res. Dev.* 45, 731–748. doi: 10.1682/JRRD.2007.08.0119

Riss, D., Arnoldner, C., Baumgartner, W.-D., Kaider, A., and Hamzavi, J.-S. (2008). A new fine structure speech coding strategy: speech perception at a reduced number of channels. *Otol. Neurotol.* 29, 784–788. doi: 10.1097/MAO.0b013e31817fe00f

Riss, D., Hamzavi, J.-S., Blineder, M., Flak, S., Baumgartner, W.-D., Kaider, A., et al. (2016). Effects of stimulation rate with the FS4 and HDCIS coding strategies in cochlear implant recipients. *Otol. Neurotol.* 37, 882–888. doi: 10.1097/MAO.0000000000001107

Riss, D., Hamzavi, J.-S., Blineder, M., Honeder, C., Ehrenreich, I., Kaider, A., et al. (2014). FS4, FS4-p, and FSP: a 4-month crossover study of 3 fine structure sound-coding strategies. *Ear Hear.* 35:e272. doi: 10.1097/AUD.0000000000000063

Shannon, R. V., Fu, Q.-J., Galvin, J., and Friesen, L. (2004). "Speech perception with cochlear implants," in *Cochlear Implants: Auditory Prostheses and Electric Hearing Springer Handbook of Auditory Research*, eds F. G. Zeng, A. N. Popper, and R. R. Fay (New York, NY: Springer), 334–376. doi: 10.1007/978-0-387-22585-2_8

Siedenburg, K., Saitis, C., and McAdams, S. (2019). "The present, past, and future of timbre research," in *Timbre: Acoustics, Perception, and Cognition Springer Handbook of Auditory Research*, eds K. Siedenburg, C. Saitis, S. McAdams, A. N. Popper, and R. R. Fay (Cham: Springer International Publishing), 1–19. doi: 10.1007/978-3-030-14832-4_1

Spitzer, E. R., Galvin, J. J., Friedmann, D. R., and Landsberger, D. M. (2021). Melodic interval perception with acoustic and electric hearing in bimodal and single-sided deaf cochlear implant listeners. *Hear. Res.* 400:108136. doi: 10.1016/j.heares.2020.108136

Stohl, J. S., Throckmorton, C. S., and Collins, L. M. (2008). Assessing the pitch structure associated with multiple rates and places for cochlear implant users. *J. Acoust. Soc. Am.* 123, 1043–1053. doi: 10.1121/1.2821980

Stupak, N., Todd, A. E., and Landsberger, D. M. (2021). Place-pitch interval perception with a cochlear implant. *Ear Hear.* 42, 301–312. doi: 10.1097/AUD.0000000000000922

Swanson, B. A., Marimuthu, V. M. R., and Mannell, R. H. (2019). Place and temporal cues in cochlear implant pitch and melody perception. *Front. Neurosci.* 13:1266. doi: 10.3389/fnins.2019.01266

Swanson, B., and Mauch, H. (2006). *Nucleus Matlab Toolbox. 420 Software User Manual*. Lane Cove, NSW: Cochlear Ltd.

Todd, A. E., Mertens, G., Van de Heyning, P., and Landsberger, D. M. (2017). Encoding a melody using only temporal information for cochlear-implant and normal-hearing listeners. *Trends Hear.* 21:2331216517739745. doi: 10.1177/2331216517739745

van der Marel, K. S., Briaire, J. J., Wolterbeek, R., Snel-Bongers, J., Verbist, B. M., and Frijns, J. H. M. (2014). Diversity in cochlear morphology and its influence on cochlear implant electrode position. *Ear Hear.* 35:e9. doi: 10.1097/01.aud.0000436256.06395.63

Vandali, A., Sly, D., Cowan, R., and van Hoesel, R. (2015). Training of cochlear implant users to improve pitch perception in the presence of competing place cues. *Ear Hear.* 36, e1–e13. doi: 10.1097/AUD.0000000000000109

Venail, F., Mura, T., Akkari, M., Mathiolon, C., Menjot de Champfleur, S., Piron, J. P., et al. (2015). Modeling of auditory neuron response thresholds with cochlear implants. *Biomed. Res. Int.* 2015:e394687. doi: 10.1155/2015/394687

Vermeire, K., Kleine Punte, A., and Van de Heyning, P. (2010). Better speech recognition in noise with the fine structure processing coding strategy. *ORL J. Otorhinolaryngol. Relat. Spec.* 72, 305–311. doi: 10.1159/000319748

von Helmholtz, H. (1885). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. London: Longmans.

Wilson, B. S., and Dorman, M. F. (2008). Cochlear implants: a remarkable past and a brilliant future. *Hear. Res.* 242, 3–21. doi: 10.1016/j.heares.2008.06.005

Wouters, J., McDermott, H. J., and Francart, T. (2015). Sound coding in cochlear implants: from electric pulses to hearing. *IEEE Signal Process. Mag.* 32, 67–80. doi: 10.1109/MSP.2014.2371671

Wright, B. A. (2013). Induction of auditory perceptual learning. *Proc. Int. Symp. Audit. Audiol. Res.* 4, 1–11.

Würfel, W., Lanfermann, H., Lenarz, T., and Majdani, O. (2014). Cochlear length determination using Cone Beam Computed Tomography in a clinical setting. *Hear. Res.* 316, 65–72. doi: 10.1016/j.heares.2014.07.013

Zeng, F.-G., Tang, Q., Lu, T., and Bensmaia, S. J. (2014). Abnormal pitch perception produced by cochlear implant stimulation. *PLoS One* 9:e88662. doi: 10.1371/journal.pone.0088662

Zilany, M. S. A., Bruce, I. C., and Carney, L. H. (2014). Updated parameters and expanded simulation options for a model of the auditory periphery. *J. Acoust. Soc. Am.* 135, 283–286. doi: 10.1121/1.4837815

frontiers | Frontiers in Neuroscience

TYPE Original Research
PUBLISHED 09 August 2022
DOI 10.3389/fnins.2022.927872

Cortical entrainment to speech produced by cochlear implant talkers and normal-hearing talkers

Shruthi Raghavendra[1]*, Sungmin Lee[2], Hyungi Chun[3], Brett A. Martin[3] and Chin-Tuan Tan[1]

[1]Department of Electrical and Computer Engineering, University of Texas at Dallas, Richardson, TX, United States, [2]Department of Speech-Language Pathology and Audiology, Tongmyong University, Busan, South Korea, [3]Graduate Center, City University of New York, New York City, NY, United States

Check for updates

OPEN ACCESS

EDITED BY
Jing Chen,
Peking University, China

REVIEWED BY
James W. Dias,
Medical University of South Carolina, United States
Yi-Wen Liu,
National Tsing Hua University, Taiwan

*CORRESPONDENCE
Shruthi Raghavendra
shruthi.raghavendra@utdallas.edu

SPECIALTY SECTION
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

RECEIVED 25 April 2022
ACCEPTED 01 July 2022
PUBLISHED 09 August 2022

CITATION
Raghavendra S, Lee S, Chun H, Martin BA and Tan C-T (2022) Cortical entrainment to speech produced by cochlear implant talkers and normal-hearing talkers. *Front. Neurosci.* 16:927872. doi: 10.3389/fnins.2022.927872

Cochlear implants (CIs) are commonly used to restore the ability to hear in those with severe or profound hearing loss. CIs provide the necessary auditory feedback for them to monitor and control speech production. However, the speech produced by CI users may not be fully restored to achieve similar perceived sound quality to that produced by normal-hearing talkers and this difference is easily noticeable in their daily conversation. In this study, we attempt to address this difference as perceived by normal-hearing listeners, when listening to continuous speech produced by CI talkers and normal-hearing talkers. We used a regenerative model to decode and reconstruct the speech envelope from the single-trial electroencephalogram (EEG) recorded on the scalp of the normal-hearing listeners. Bootstrap Spearman correlation between the actual speech envelope and the envelope reconstructed from the EEG was computed as a metric to quantify the difference in response to the speech produced by the two talker groups. The same listeners were asked to rate the perceived sound quality of the speech produced by the two talker groups as a behavioral sound quality assessment. The results show that both the perceived sound quality ratings and the computed metric, which can be seen as the degree of cortical entrainment to the actual speech envelope across the normal-hearing listeners, were higher in value for speech produced by normal hearing talkers than that for CI talkers. The first purpose of the study was to determine how well the envelope of speech is represented neurophysiologically via its similarity to the envelope reconstructed from EEG. The second purpose was to show how well this representation of speech for both CI and normal hearing talker groups differentiates in term of perceived sound quality.

KEYWORDS
cortical entrainment, electroencephalogram, cochlear implant, perceived sound quality, speech envelope

Introduction

Sound quality is classically estimated from the physical difference of an utterance produced by a talker from its standard reference and is used as a metric to quantify "how well" the talker has spoken (Loizou, 2011), which may not align well with outcomes obtained perceptually. Perceived sound quality-based listener judgments may be a more direct way to determine "how well" talker has spoken, however, the outcome can vary greatly from one listener to another. Adding physiological data measurement (i.e., cortical activity) to the behavioral sound quality judgments may facilitate the needed consistency and consensus across listeners.

A cochlear implant (CI) is a common device used to restore the ability to hear and provide the necessary auditory feedback to produce and monitor speech for hard-of-hearing individuals. However, there remains a large variability in speech production proficiency among implant recipients, which could be attributed to the age of implantation, duration of hearing loss, duration of device used and remaining residual hearing (Ruff et al., 2017; Kim et al., 2018; Gautam et al., 2019). In this study, we obtained perceived sound quality ratings of speech produced by both a CI talker group and a normal-hearing (NH) talker group and captured their cortical entrainment to the speech as indicated by associated cortical activities of the normal-listening listeners. The first purpose of the study was to determine how well the actual envelope of speech is represented neurophysiologically via its similarity to the envelope reconstructed from the co-fluctuating electroencephalogram (EEG) activities using a re-generative model (Crosse et al., 2016). The second purpose was to show how well the speech envelope was represented in response to both CI and NH talker groups and differentiated the groups in terms of their perceived sound quality. The goal is to achieve a metric to assess "how well" hard-of-hearing talkers have spoken and the auditory feedback they received in their current aural compensation.

Neurophysiological processing of sound is usually examined using event-related potentials (ERPs). The P3 component has been recently used to evaluate the effects of perceived quality changes in speech (Uhrig et al., 2019a,b) and it was reported that the peak amplitude and latency were modulated by sound quality. Likewise, others studies (Martin et al., 1997; Martin and Stapells, 2005; Antons et al., 2010, 2012; Porbadnigk et al., 2013) also showed that, as the level of degradation decreases when compared to the standard stimulus, which means it is harder for listeners to discriminate the deviant stimulus from the standard stimulus, the amplitude and latency of the P3 component become lower and longer, respectively. These ERP techniques commonly utilize auditory stimuli of short duration, which are not optimal to use to make perceptual sound quality judgments.

Cortical entrainment to the envelope of speech may serve as a useful alternative to investigate the neurophysiologic processing of continuous speech, as evidence has shown that the dynamic cortical activity tracks the envelope of continuous, natural speech (Aiken and Picton, 2008; Lalor and Foxe, 2010). This phenomenon reflects the activity of distinct neural populations that implement different functional roles including encoding acoustic features (for a review, see Ding and Simon, 2014) and can be captured by EEG (Aiken and Picton, 2008). The temporal envelope, a slow variation of the amplitude of speech is considered to be one of the most important cues for speech intelligibility (Peelle and Davis, 2012) and speech perception (Shannon et al., 1995). Particularly in delta (1–4 Hz) and theta (4–8 Hz) frequency bands, neural activity is known to track the amplitude envelope of speech (Ding and Simon, 2014).

This cortical tracking of the speech envelope can be inferred from the correlation between the actual speech envelope and the speech envelope predicted/decoded from the EEG/magnetoencephalography (MEG). Many studies have demonstrated that the speech envelope can be decoded from single-trial EEG/MEG recordings obtained by presenting the stimulus only once (Ding and Simon, 2012, 2013; Di Liberto et al., 2015; O'Sullivan et al., 2015). A multivariate linear model (Crosse et al., 2016) was developed to map the multi-channel EEG signal into a single-channel speech envelope with the intent of minimizing the mean-squared error between the actual speech envelope and the reconstructed envelope. To accomplish this, the time-shifted version of the EEG channels is first obtained by applying a range of delays also known as the temporal integration window (e.g., 0 and 500 ms) to each channel, then all of the delayed channels are weighted, to linearly reconstruct the envelope of speech. The actual speech envelope and the reconstructed envelope are then correlated with each other, which yields a measure of envelope entrainment. Using this technique, previous studies have examined the cortical entrainment to the envelope of speech and correlated the degree of entrainment to behavioral speech intelligibility (Ding and Simon, 2013; Kong et al., 2015; Vanthornhout et al., 2018). It has been shown that higher speech intelligibility coincides with improved cortical entrainment to the speech envelope. This technique was also used as an EEG-based measure of attention decoding in a cocktail party environment (O'Sullivan et al., 2015).

Likewise, we used this technique to study the cortical entrainment to the speech envelope in relation to sound quality as perceived by normal-hearing listeners, and developed a metric to differentiate speech spoken by CI and NH talker groups. Bootstrapped Spearman correlation between the actual speech envelope and the envelope reconstructed from the EEG was computed to quantify the cortical entrainment to the speech envelope, and compared to the sound quality as perceived by the listener. We hypothesized there would be closer cortical tracking of the speech envelope (higher correlation) when speech is of higher perceived sound quality. We therefore anticipated that closer cortical tracking of speech envelope would be obtained with speech passages spoken by NH talkers than with those spoken by CI talkers.

Materials

Participants

Eleven normal-hearing listeners were recruited from the University of Texas at Dallas student population for this study. Their age ranged from 19 to 29 years (mean age = 21.5 years; 5 female, 6 male). All normal-hearing listeners were screened by presenting pure tones at 20 dB HL from 250 HZ to 8 kHz at octave frequencies and had normal hearing thresholds < 20 dB HL. This study was approved by the Institutional Review Board at the University of Texas at Dallas. All participants signed the informed consent forms prior to participation in the experiment and were paid for their participation.

Talkers

Two groups of 8 talkers each were selected from the "Corpus of deaf speech for acoustic and speech production research" database collected at the University of Memphis (Mendel et al., 2017). This corpus is a pool of speech recordings digitally sampled at 44,100 Hz, spoken by NH talkers and hearing impaired (HI) talkers. The entire "Rainbow Passage" spoken by each talker was recorded by the authors using a Shure SM93 prolog dynamic microphone. We selected speech passages read by 8 CI talkers (4 female; 4 male) and 8 NH talkers (6 female; 2 male) for our present study. The two groups of talkers are, respectively, referred to as CI talker group and NH talker group. **Table 1** shows the duration of speech passage recorded by each talker in both talker groups with mean and standard deviation across each talker group.

Table 2 presents the demographic details of the chosen CI talkers from the database (Mendel et al., 2017). The CI talkers are aged between 16 and 77 years with an average age of 47.5 years. Other than CI talker #3, the rest of the CI talkers were post-lingually deaf. This should not be confused with the participants in our study who are normal-hearing listeners listening to the speech passages produced by these CI talkers and NH talkers.

Methods

Figure 1 shows the overall setup for both behavioral and electrophysiological experiments. Normal-hearing listeners listened to the speech passages produced by both CI talkers and NH talkers. The behavioral sound quality assessment experiment was always conducted first, followed by the EEG experiment, which were performed on the same day. For each listener, the total duration of the behavioral and EEG experiments varied between 2 and 3 h. In the behavioral experiment, each listener was presented with speech passages spoken by 8 CI talkers and 8 NH talkers in a randomized order and was asked to rate the perceived sound quality of each speech passage. In the electrophysiological experiment, the single-trial EEG responses of the same listeners were recorded while they were presented with the same speech passages they heard in the behavioral experiment. The stimulus presentation order was randomized across the two experiments.

From each speech passage, the envelope was extracted and referred as the actual speech envelope. Then the envelope of the speech was reconstructed from its associated EEG signal using a decoder referred to as the reconstructed/predicted envelope. The bootstrapped Spearman correlation between the actual speech envelope and the reconstructed envelope is employed as a metric that measures the cortical entrainment to the actual speech envelope in each normal-hearing listener. Then the sound quality and the cortical entrainment to the speech envelope were compared for the speech passages produced by two groups of talkers.

Behavioral experiment

For the sound quality assessment, listeners were seated in a soundproof booth in front of a touch screen computer monitor. They were seated 1 m from a loudspeaker at 0°azimuth at their ear height. The speech passages were presented via the loudspeaker at 65 dB SPL. Each normal-hearing listener was asked to perform two trials of behavioral sound quality assessments. In one trial, the speech passages spoken by 8 CI talkers and 8 NH talkers (total of 16 passages) were presented one at a time in a randomized order to a listener. Each speech passage was presented only one time. Listeners were instructed to listen to the speech passages and to rate the sound quality of each passage on a Likert 10 point scale (Sangthong, 2020), with 1 being the most distorted and 10 being the most undistorted using a touch screen monitor. In the second trial, the same speech passages were randomly presented and listeners again rated the sound quality. The perceived sound quality rating of each speech passage was computed as the average of the ratings obtained across the two trials.

Electroencephalogram experiment

EEG recording was performed after the behavioral sound quality rating assessment. A 64-channel actiCHamp amplifier EEG setup (Brain Products GmbH, Munich, Germany) was used

TABLE 1 Duration of the stimuli.

CI talker	#1	#2	#3	#4	#5	#6	#7	#8	$\bar{x}$	s.d
Duration of spoken passage (s)	115	105	115	136	127	108	156	141	125.4	17.8
NH talker	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	$\bar{x}$	**s.d**
Duration of spoken passage (s)	78	94	88	88	87	91	104	83	89.1	7.7

TABLE 2 Demographic details of the CI talkers from the deaf speech corpus.

CI talker	Age (years)	Gender	Age of first amplification use (years)		Onset of hearing loss	Current type of amplification		Communication mode[4]
			Right	Left		Right	Left	
#1	37	Male	5	5	Post-lingual	CI[2]	CI	Oral
#2	38	Female	28	NA[1]	Post-lingual	CI	NA	Oral
#3	16	Female	0.5	2	Pre-lingual	HA[3]	CI	Oral
#4	77	Male	18	73	Post-lingual	HA	CI	Oral and sign
#5	62	Female	38	38	Post-lingual	CI	HA	Oral and sign
#6	60	Male	52	52	Post-lingual	HA	CI	Oral and sign
#7	57	Female	3	58	Post-lingual	HA	CI	Oral
#8	33	Male	NA	3	Post-lingual	NA	CI	Sign only

[1]NA, not applicable. [2]CI, cochlear implant. [3]HA, hearing aid. [4]Oral indicates that the talker used oral speech and language. Sign indicates that the talker used sign language.

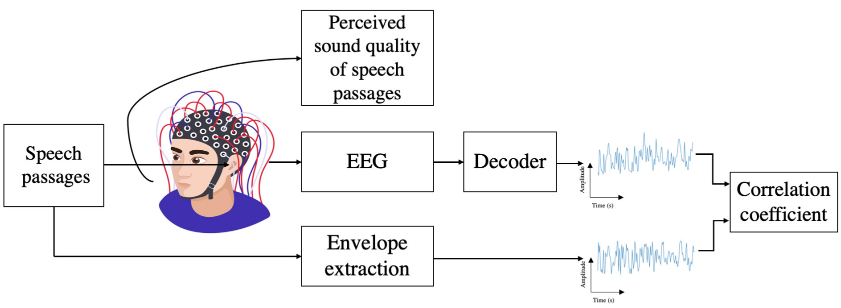

FIGURE 1
Behavioral and EEG experiment setup.

to record the ongoing EEG in response to the same passages produced by two groups of talkers used in the behavioral test. The EEG signals were recorded using an electrode cap (actiCAP, Brain Products GmbH, Munich, Germany) placed in accord to the 10–20 system (Oostenveld and Praamstra, 2001). The ground channel and the reference channel were located at FPz and FCz, respectively. To monitor eye-movement artifacts, the HEOG was monitored from electrodes placed at the lateral outer canthi and the VEOG was recorded from electrodes placed above and below the left eye. All electrode impedances were maintained below 10 kOhms. Each listener was asked to minimize body movement and watch a silent, captioned movie while EEG recording was in progress. EEG data were recorded with a sampling rate of 1,000 Hz. The EEG recordings were time aligned with the stimulus based on a trigger event inserted at the onset of each passage.

Electroencephalogram preprocessing

All EEG data were analyzed offline using custom scripts in MATLAB_R2021a (MathWorks, Natick, MA, United States). EEG data were preprocessed using the EEGLAB toolbox (Version14.1.2b; Delorme and Makeig, 2004) in MATLAB to prune unwanted artifacts. The portion of the EEG contaminated with artifacts related to muscle was removed by visual inspection. Artifacts from the eye blink, lateral eye movement, and heart beat were pruned from the EEG using independent component analysis (ICA) in the

EEGLAB toolbox. To prune the independent components reflecting eye blinks and lateral eye movements, the fully automated Eye-Catch approach (Bigdely-Shamlo et al., 2013) was used. After the artifact removal procedure, the data were re-referenced to a common average reference (CAR) and the reference channel FCz was added back to the data.

Signal processing: Electroencephalogram and speech envelope

Electroencephalogram

Cortical tracking of the acoustic features of speech is typically analyzed in specific frequency bands, including delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–14 Hz), beta (14–30 Hz), low-gamma (30–70 Hz), and high-gamma (70–100 Hz). Cortical activities in delta (1–4 Hz) and theta (4–8 Hz) frequency bands are particularly known to track the amplitude envelope of speech (Ding and Simon, 2014) and are the focus here. To extract the EEG activity in the delta and theta frequency bands, the preprocessed EEG signal was band-pass filtered between 0.5–4 Hz (delta) and 4–8 Hz (theta) using a zero-phase Butterworth filter with 80 dB attenuation at 10% outside the passband (Vanthornhout et al., 2018; Lesenfants et al., 2019a,b). The zero-phase filtering was performed using the *filtfilt* command in MATLAB. **Figure 2** presents the magnitude responses of the Butterworth filters in delta (**Figure 2A**) and theta (**Figure 2B**) frequency bands designed at sampling frequency = 1,000 Hz. The computed order of the Butterworth filters was 136 and 118 (68 and 59, 2nd order in cascade), respectively, to realize the delta and theta frequency bands. We chose this filter to have a sharper roll-off at the edge of the pass band as delta and theta are consecutive frequency bands. The performance of the filters was shown in **Figure 2**.

Speech envelope

The Hilbert transformation was first applied to the speech signal, to obtain the complex-valued output. The absolute value of the complex-valued output was computed which provides the instantaneous amplitude of the signal followed by low-pass filtering at 40 Hz to extract the speech envelope. In our study, we also filtered the extracted speech envelope into 0.5–4 Hz (delta) and 4–8 Hz (theta) to match the bandwidth of the EEG signals. The extracted speech envelope was resampled to 1,000 Hz to match the sampling rate of EEG signals before applying the zero-phase Butterworth filters shown in **Figure 2**. All speech envelopes and EEG data were further downsampled to 128 Hz (Crosse et al., 2016) to reduce the computation time.

Speech envelope reconstruction

A linear decoder as proposed in Crosse et al. (2016) was used to predict and reconstruct the speech envelope from the associated EEG activity. The decoder acts as a spatiotemporal filter that linearly maps the EEG to the speech envelope thereby reconstructing the speech envelope estimated from the corresponding EEG response recorded when listening to the speech passages. The time-shifted version of the EEG channels was obtained by applying a range of delays (in general, between 0 and 500 ms) to each channel, then all of the delayed channels were weighted, in order to linearly reconstruct the envelope. The actual speech envelope and the reconstructed envelope were then correlated with each other, which yields a measure of cortical entrainment to the actual speech envelope. The process is explained as follows:

Given a linear decoder $g(\tau, n)$ representing the linear mapping from the EEG response, $r(t, n)$, back to the stimulus envelope $s(t)$, a single estimate of the stimulus envelope $\hat{s}(t)$ was computed as follows:

$$\hat{s}(t) = \Sigma_n \Sigma_\tau \, r(t + \tau, n) g(\tau, n) \quad (1)$$

with t ranges from 0 to T, length of the signal. τ is the integration window length and n is the index of the N EEG channels. The decoder $g(\tau, n)$ was derived by minimizing the mean-squared-error (MSE) between the actual stimulus envelope $s(t)$, and the estimated stimulus envelope $\hat{s}(t)$, i.e.,

$$\min \varepsilon(t) = \Sigma_t [s(t) - \hat{s}(t)]^2 \quad (2)$$

The decoder computation can be expressed using the following matrix operation:

$$\boldsymbol{g} = (\boldsymbol{R}^T \boldsymbol{R} + \lambda \boldsymbol{I})^{-1} \boldsymbol{R}^T \boldsymbol{s} \quad (3)$$

where the superscript T represents the transpose of a matrix, **I** is the identity matrix and λ is the ridge/regularization parameter chosen to make the decoder less prone to overfitting. **R** represents the lagged time series of EEG response matrix r, for N channels, the dimensions of matrix **R** is $T \times N\tau_{window}$, where $\tau_{window} = \tau_{max} - \tau_{min}$ with τ_{min} and τ_{max} represent the minimum and maximum time lags (in samples), respectively. The stimulus envelope, **s**, is a column-wise vector of length T and the resulting decoder, **g**, would be a vector of $N\tau_{window}$ samples.

Figure 3 illustrates the procedure of reconstruction of a speech envelope in a single listener. For each normal-hearing listener, 16 decoders were obtained, corresponding to the 16 speech passages combined from the two groups of talkers. These 16 decoders were, respectively, computed as Eq. 3 using each of the 16 speech envelopes extracted from the 16 spoken speech passages by 8 CI and 8 NH talkers and their associated EEG signals collected from the normal-hearing listeners listening to these speech passages. Later, the reconstructed envelope which corresponds to the speech passage

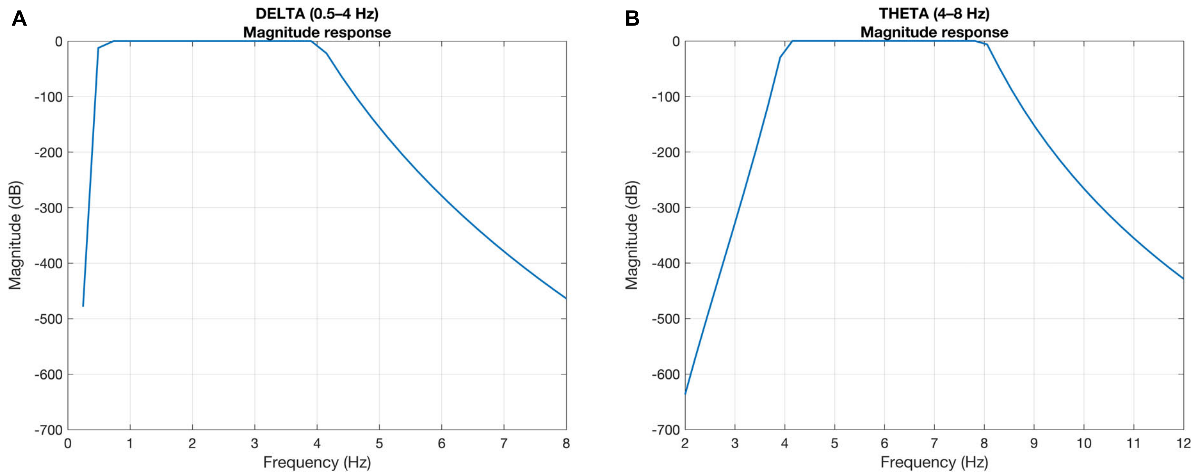

FIGURE 2
Magnitude responses of the Butterworth filter showing sharper roll-off at the edge of the pass bands in **(A)** delta and **(B)** theta frequency bands.

produced by CI talker #1 was obtained by correlating an "Average decoder" with the EEG signal recorded from the listener while listening to spoken passage by CI talker #1 as shown in Eq. 1. This "Average decoder" was computed as the average of the remaining 15 decoders, i.e., decoder 2 to decoder 16. This "leave-one-out" model approach was repeated to reconstruct the envelope corresponding to the remaining 15 speech passages in the same listener. Finally, to compute the bootstrapped Spearman correlation between the actual speech envelope and the reconstructed envelope, we randomly permuted the reconstructed envelope 1,000 times and calculated Spearman's correlation between the result and the actual speech signal for each permutation. The final correlation value was evaluated as the averaged value of the resulted 1,000 correlation values.

The sample size for this study is supported by a power analysis conducted on the data. Cohen's "d" (Cohen, 1988) was computed on the paired samples t-tests for the sound quality and envelope entrainment data. Assuming a significance level, alpha = 0.01 and power of 80%, the sample sizes were, respectively, estimated as "8" and "14" with the behavioral and EEG experimental results and our current study's sample size falls between the estimated sample sizes.

Results

Behavioral results

Figure 4 presents the perceived sound quality ratings as rated by individual normal-hearing listeners for the speech passages produced by the two talker groups. **Figure 4** also shows the mean and standard deviation of the perceived sound quality ratings for the spoken speech passages by CI talker (red curve) and NH talker (black curve) groups across the normal-hearing listeners. The perceived sound quality ratings for the speech passages spoken by CI talkers varied widely along the range. Whereas for those spoken by the NH talkers, there was little difference in the sound quality ratings across the NH talker group as each of the speech passage spoken by NH talkers was rated almost equally high in sound quality. Within the CI talker group, CI talker #2 was rated with highest mean sound quality of 9.2 and CI talker #8 was rated with lowest mean sound quality of 1.2. In addition, there was a larger relative difference in the standard deviation for the speech passages spoken by CI talkers compared to the speech passages spoken by the NH talkers especially in case of CI talker #1, #4, and #7.

Statistical analysis

Statistical analysis was performed using MATLAB_R2021a and R studio (version 3.3.0). First, a paired-samples t-test on perceived sound quality between the CI talker group and NH talker group was performed to investigate whether there is a significant difference in perceived sound quality ratings for these two groups of talkers. For paired-samples t-test, the two dependent samples contained one entry for each listener with a single averaged perceived sound quality across CI talker group and NH talker group. The results showed that the mean sound quality ratings for the NH talker group were significantly higher than that of the CI talker group [$t(10) = 9.8$, $p < 0.001$]. To appropriately model our dataset statistically and to assess the relevant factors such as talker group (CI/NH), talker gender (male/female), and duration of the utterance of spoken passages (referred to as duration; **Table 1**) contributing to perceived sound quality, we used a linear mixed effects regression (lmer) model using the lme4 package (Bates et al., 2015). The above three fixed factors are, respectively, referred

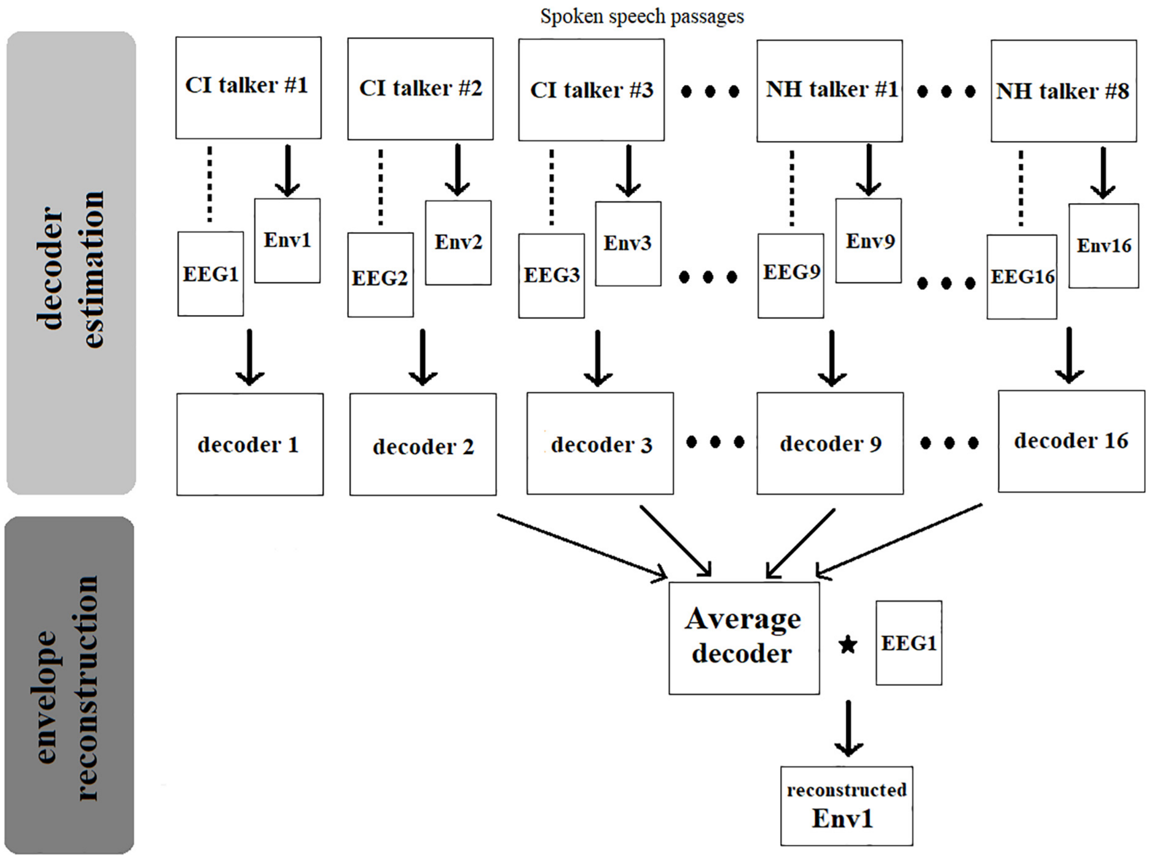

FIGURE 3
Flowchart illustrating the stages involved in the reconstruction of a speech envelope.

to as TALKERGROUP, TALKERGENDER, and DURATION. In the lmer model, the perceived sound quality, as rated by the listeners, was considered as the outcome variable and the listener as the random factor. By constructing two hierarchical regression models, we assessed how the three fixed factors/predictors such as TALKERGROUP, TALKERGENDER, and DURATION modulated the perceived sound quality. The hierarchical regression analysis consisted of two models: Model 1 (m1) used only TALKERGROUP as a predictor, while Model 2 (m2) used TALKERGROUP, TALKERGENDER, and DURATION as predictors. Model comparison between the simpler and the complex model, i.e., m1 and m2, respectively, was obtained with the R-function ANOVA that uses the chi-square test. The improvement in the model fit for adding more predictors was determined by comparing the Akaike Information Criterion (AIC; Akaike, 1974) of the simpler and the complex model. If the AIC of the more complex model was smaller than that of the simpler model with $p < 0.05$ (statistically significant), the more complex model was considered to have a better fit. The best-fitting model was then investigated for the modulation of the outcome variable by various fixed factors. The results showed that the complex model m2 yielded the lower AIC (AIC = 704.10) whereas the simpler model m1's AIC was significantly higher (AIC = 781.78) [chi-square(6) = 81.68, $p < 0.001$].

Speaking of the effects of the fixed factors on the perceived sound quality, the effect of TALKERGROUP was significant (beta = −0.857, $F(1,162) = 6.89$, $p = 0.009$), also the talker-group related differences in perceived sound quality in listeners can be observed in **Figure 4**, which demonstrates that the mean quality rating for the NH talker group ($\bar{x}$: 9.4, s.d: 0.8) was higher than that of the CI talker group ($\bar{x}$: 6.9, s.d: 3). A significant effect of the TALKERGENDER factor was also observed [beta = −0.094, $F(1,162) = 10$, $p = 0.002$] and **Figure 5A** visualizes the talker-gender related differences such that the median value of the perceived sound quality across the speech produced by female talkers (median = 9.5) was higher than that of the speech produced by male talkers (median = 8.6). **Figure 5B** visualizes the duration related difference in perceived sound quality across the listeners, and the effect of DURATION on the perceived sound quality was also statistically significant [beta = −0.973, $F(1,162) = 92.9$, $p < 0.001$]. In general, the perceived sound quality was higher for the speech of shorter duration of utterances.

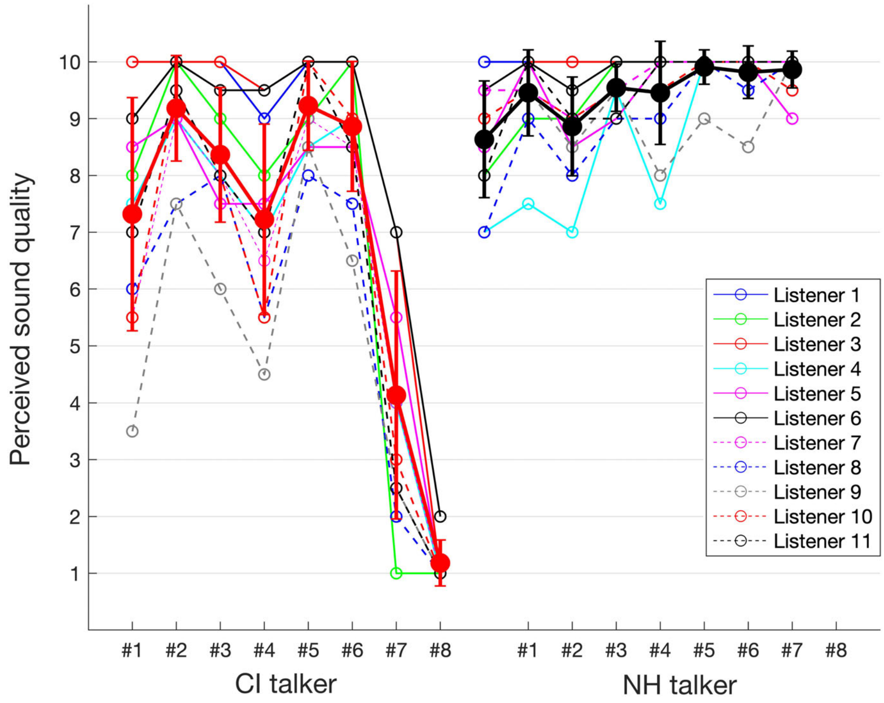

FIGURE 4
Individual listener assessed sound quality data point for the speech passages spoken by CI and NH talkers and also showing the mean and SD of the sound quality ratings for the CI and NH talker groups.

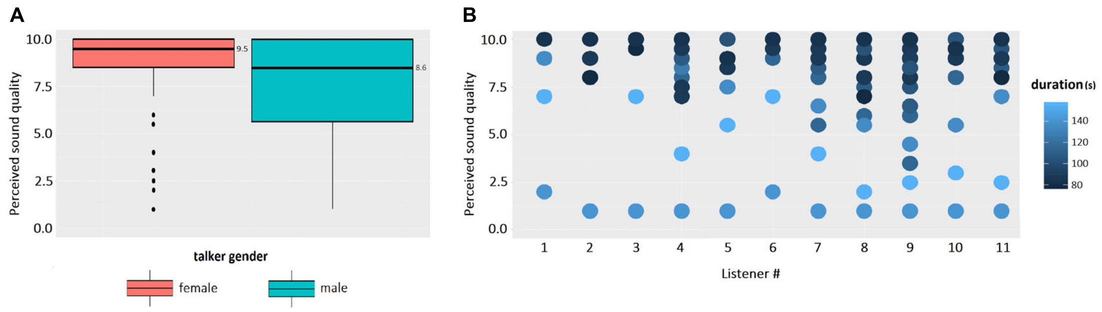

FIGURE 5
(A) Perceived sound quality vs. talker gender. **(B)** Perceived sound quality vs. duration of the spoken passage across two talker groups.

Envelope entrainment in response to speech spoken by cochlear implant talkers and normal-hearing talkers

An optimal regularization parameter λ that minimizes the MSE between the actual speech envelope and the reconstructed envelope was first selected to train the decoder, and followed by choosing the optimal integration window over which the decoder integrates the EEG to reconstruct the speech envelope. With these optimal parameters, the correlation between the actual speech envelope and the reconstructed envelope was computed as the metric to quantify the degree of cortical entrainment to the speech envelope. The computed metric is then compared between the two talker groups in relation to their perceived sound quality.

Entrainment to speech envelope (<40 Hz)

Selection of parameters to train decoders

An optimized λ was chosen from a set of values (10^{-1}, 10^0,..., 10^3, 10^4, 10^5,.., 10^{10}) which minimizes the MSE between the actual speech envelope and the reconstructed envelope as in Eq. 2. In both delta EEG band and theta EEG band, an optimal λ value was evaluated for each listener, and the value, however, turned out to be 10^6 for all the listeners. In general, an integration window is chosen from the range of time lags between 0 and 400/500 ms (Crosse et al., 2016). Likewise, we varied the temporal integration window of the decoder from 0–20 ms to 0–400 ms with a step size of 20 ms and chose an integration window of 0–400 ms for all the listeners, as there was no consistent peaking of correlation vs. time lags was observed across spoken speech passages by CI talkers and NH talkers.

The decoders are trained using these chosen parameters to reconstruct the envelope of speech from delta and theta EEG bands separately. Some instances of the reconstructed envelope and the actual speech envelope, and also the bootstrapped Spearman correlation "*rho*" between those two envelopes are shown in **Figure 6** (Listener 6's data). The correlation values shown at the top of the respective waveforms is considered as a metric reflecting the degree of cortical entrainment to the respective actual speech envelope. In general, a higher correlation value between the actual speech envelope and the reconstructed envelope indicates higher cortical entrainment to the actual speech envelope. Using the delta or theta EEG band and speech envelope $<$ 40 Hz, the linear decoder poorly reconstructed the envelope from the EEG when correlated to the actual speech envelope. The highest value of correlation between the actual (for the speech passage spoken by NH talker #6) and predicted envelope observed was 0.1 using delta EEG band.

Figure 7 presents the cortical entrainment to speech envelope ($<$40 Hz) with EEG in the delta (top panel) and theta (bottom panel) frequency bands for the speech passages spoken by two groups of talkers (CI and NH). The mean sound quality rating as assessed by normal-hearing listeners for each speech passage in the two talker groups is also shown over the entrainment boxplots to show their perceived sound quality. Overall, the range of correlation values computed between the speech envelopes and the reconstructed envelopes from the EEG was higher using the delta EEG band (**Figures 7A,B**) than those obtained using the theta EEG band (**Figures 7C,D**). The above observation was true for the two groups of talkers as well. Comparing the envelope entrainment between the two groups of talkers, a higher variation in the median values of the envelope entrainment among spoken speech passages within the CI talker group was observed as compared to the NH talker group using delta EEG band (**Figures 7A,B**). Similar to the sound quality ratings, the variability in the median value of the envelope entrainment among spoken speech passages is less obvious within the NH talker group using delta EEG band (**Figure 7B**). Also, the range of correlation values were reduced greatly from delta to theta bands in both talker groups and no difference could be observed between the two talker groups using the theta EEG frequency bands.

Statistical analysis

Statistical analyses was conducted to investigate whether there was a significant difference in envelope entrainment (correlation values) in response to the speech passages produced by the two talker groups. A paired-samples t-test was performed on the Spearman correlation values for the 11 listeners, each entry of the listener with a single averaged correlation value across CI talker group and NH talker group as two dependent samples in the delta and theta EEG bands separately. The t-test results showed no significant difference in the envelope entrainment between the two talker groups using the delta EEG band [$t(10)$ = 0.28, p = 0.79] or using theta EEG band [$t(10)$ = 1.9, p = 0.09]. Using delta EEG band, the mean value of the envelope entrainment across the NH talker group 0.17 and the mean value of envelope entrainment across the CI talker group was 0.15. Whereas, using the theta EEG band, the mean value of envelope entrainment across the NH talker group was 0.16 and that of the CI talker group was 0.13.

Entrainment to speech envelope filtered to match bandwidth of delta (0.5–4 Hz) and theta (4–8 Hz) bands

Selection of parameters to train decoders

For each listener, an optimal λ value was evaluated in both delta and theta bands and the values are tabulated in **Table 3**. **Figure 8** presents the Spearman correlation as a function of different time lags in each talker group across 11 normal-hearing listeners (spoken speech passages by 8 talkers in each talker group*11 NH listeners = 8*11 curves). Overall, there was quite a large variability in the Spearman correlation curves across the time lags observed in both **Figures 8A,B** showing the inter-stimulus (talker) and inter-subject (listener) differences except for the earlier integration window of 0–150 ms. The earlier integration window of 0–150 ms resulted in comparatively lower correlation values across the listeners and the correlation values seem to increase as the upper bound of the time lag increases. Hence, the integration window to train a decoder was chosen as 150–400 ms across the NH listeners. The decoders are trained using these chosen parameters to reconstruct the envelope from the EEG in both delta and theta bands and some instances of improved correlation with the inclusion of band-pass filtered versions (delta and theta) of speech envelopes in the same listener shown previously (**Figure 6**) are presented in **Figure 9**.

Figure 10 presents the envelope entrainment in response to the speech passages spoken by two groups of talkers (CI and NH), where the speech envelopes were also band-pass filtered into the delta (top panel) and theta (bottom panel) frequency bands. Overall, the range of correlation values observed in the delta band were broader (**Figures 10A,B**) compared to the range observed in the theta band (**Figures 10C,D**). Comparing the envelope entrainment between the two groups of talkers in the delta band, a higher variation in the median values of the envelope entrainment within the NH talker group (**Figure 10B**) as compared to the CI talker group (**Figure 10A**). In contrary with the results of entrainment to speech envelope ($<$40 Hz) (**Figure 7A**), the variability in the median value of the envelope

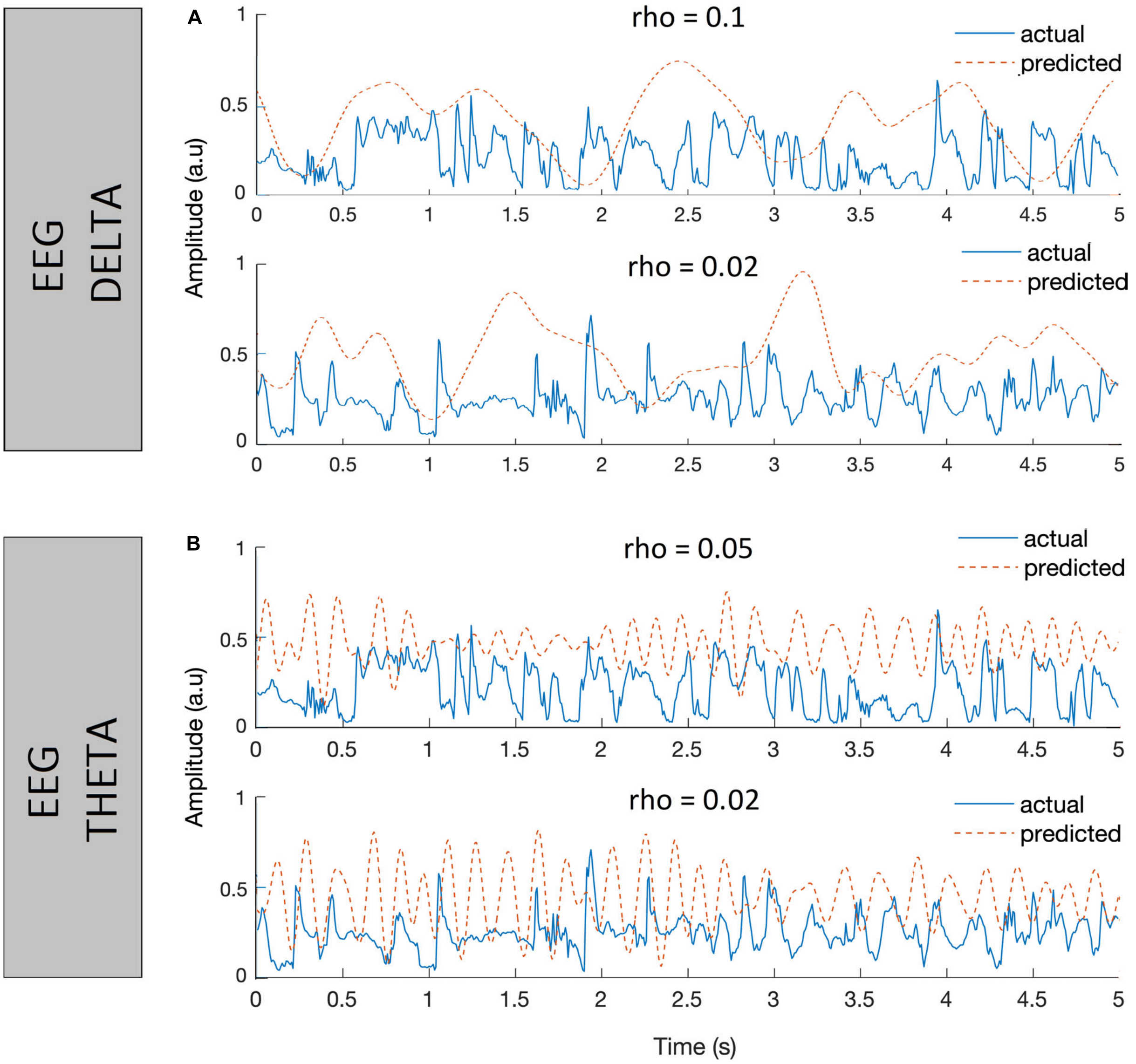

FIGURE 6
Plots showing instances of actual speech envelopes (blue) and the reconstructed envelopes (red dotted) and bootstrapped Spearman correlation (*rho*) between them. The decoders were trained using EEG in delta **(A)** and EEG in theta bands **(B)**, with speech envelope < 40 Hz.

entrainment is less obvious within the CI talker group in the delta band (**Figure 10A**). Looking at the behavioral sound quality, between the talker #1, #2, and #3 in both the talker groups, the speech passage spoken by talker #2 had a higher sound quality compared to that of talker #1 and #3. A similar trend was also observed in the entrainment data, i.e., median value of the envelope entrainment for the speech passage produced by talker #2 is higher compared to that of talker #1 and #3. The above observation was true in both delta and theta bands and in both groups of talkers. Compared to the decoder trained using the actual speech envelope (<40 Hz), the decoders trained using the speech envelopes filtered to match delta and theta EEG bandwidths were able to achieve a higher correlation and better quantify the difference observed from the speech passages spoken by CI talkers and NH talkers.

Statistical analysis

A paired-samples t-test was performed on the Spearman correlation values for the 11 NH listeners between the CI and NH talker groups in the delta and theta bands separately. In the theta band, the t-test did not reveal a significant difference in the envelope entrainment between the two talker groups [$t(10) = 1.47$, $p = 0.17$] that the mean value of envelope entrainment across the NH talker group was 0.09 and that of the CI talker group was 0.08. Whereas in the delta band, the

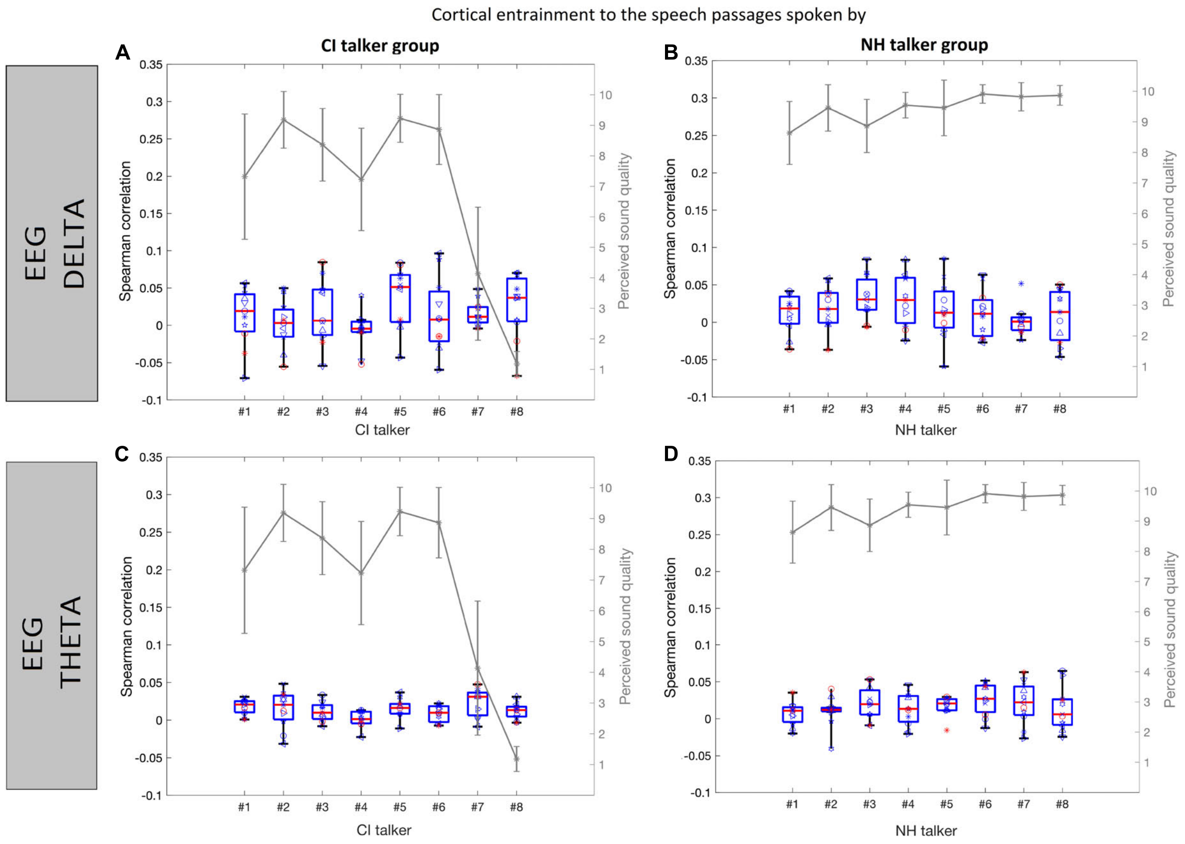

FIGURE 7
Cortical entrainment to speech envelope (<40 Hz) in the normal-hearing listeners is shown for each of the speech passages from the CI talker group **(A,C)** and NH talker group **(B,D)** using EEG in delta and EEG in theta bands, respectively. The data are represented in the form of the boxplots depicting medians (center mark) and interquartile ranges (the bottom and top edges of the box indicate the 25th and 75th percentiles). The boxplot shows the entrainment data variance across listeners, and the individual listener's data points are plotted over it. Each listener is shown with a different symbol. The perceived sound quality curve (gray) is also shown for comparison.

t-test results showed a significant difference in the envelope entrainment between the two talker groups [$t(10) = 3.2$, $p = 0.009$]. The mean value of the envelope entrainment across the NH talker group ($\bar{x}$: 0.12) was significantly higher than the mean value of envelope entrainment across the CI talker group ($\bar{x}$: 0.08). Furthermore, in the delta band, we performed the similar lmer analysis presented before by replacing the outcome variable with Spearman correlation between the actual speech envelope and the reconstructed envelope. We assessed how the fixed factors TALKERGROUP, TALKERGENDER, and DURATION modulated Spearman correlations by fitting two separate hierarchical models: m1 and m2. Model comparison was carried out between the simpler model with the only fixed factor: TALKERGROUP, and the complex model comprising all the fixed factors: TALKERGROUP, TALKERGENDER, and DURATION. The ANOVA results show that there was no significant difference between the simpler model m1's AIC (AIC = −533.29) and the AIC of the complex model m2 (AIC = −529.33) [chi-square(6) = 0.0318, $p < 0.98$]. The addition of the TALKERGENDER and DURATION fixed factors did not affect the goodness of fit of the model. With the simpler model m1, the lmer analysis results showed the TALKERGROUP had a significant effect on envelope entrainment [beta = −0.21, $F(1,162) = 20.84$, $p < 0.001$].

Discussion

Cochlear implant talker group vs. their speech quality

The database (Mendel et al., 2017) provided limited information about the individual CI talkers and this information is summarized in **Table 4**. As can be seen in **Table 4** showing the demographic details of the chosen CI talkers, CI talkers #2, #5, and #6 were aided later than the others (marked in gray), presumably meaning a later onset of hearing loss. Therefore, better perceived speech quality likely reflects their

TABLE 3 Individual best λ for NH listeners.

					DELTA						
Listener #	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	**#9**	**#10**	**#11**
λ value	10^3	10^2	10^3	10^3	10^2	10^2	10^2	10^3	10^2	10^2	10^2
					THETA						
Listener #	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	**#9**	**#10**	**#11**
λ value	10^2	10^2	10^2	10^2	10^2	10^2	10^2	10^2	10^2	10^2	10^2

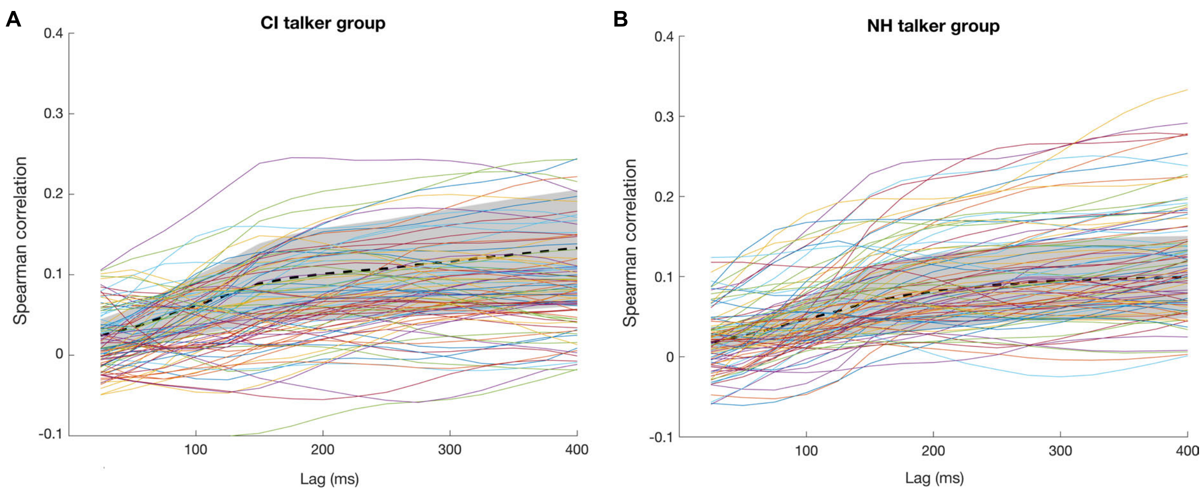

FIGURE 8
Spearman correlations between the actual speech envelope and the reconstructed envelope for **(A)** CI talker group and **(B)** NH talker group across the time lags. The thick dotted lines represent the mean values, and the shaded areas represent the standard deviation across the listeners.

greater degree of time in sound. The opposite occurred for CI talkers #7 and #8. These two talkers were aided earlier (marked in yellow), presumably indicating an earlier onset of hearing loss. Further, CI talker #8 used sign language only for communication whereas the others used oral communication and/or oral communication plus sign language. Their poorer speech quality likely reflects these factors.

Cortical entrainment to speech envelope

Correlation between the actual speech envelope and the reconstructed envelope was higher using the speech envelope filtered to match delta (0.5–4 Hz) and theta (4–8 Hz) bandwidth, when compared to the speech envelope with fluctuations < 40 Hz (Vanthornhout et al., 2018). In this scenario, between the delta and theta bands, in general, a closer cortical tracking to speech envelope was observed in the delta band. The observation was found in both talker groups. Previous study (Vanthornhout et al., 2018) has shown that cortical activity in delta band can serve as an indicator of how well a listener can recognize speech in the presence of noise. It is also known to carry the prosodic information (Goswami and Leong, 2013) to predict speech intelligibility (Ding and Simon, 2013; Vanthornhout et al., 2018). In alignment with the above studies, our results also suggest the cortical entrainment in delta band can serve as an indicator of the perceived sound quality by the listeners.

This preliminary research employed a linear model to predict and reconstruct the speech envelope from the EEG signal. The correlation values between the actual speech envelope and the reconstructed envelope obtained mainly ranged from −0.1 to 0.2. The reason for observing low correlation values between the two envelopes could be the assumption of a linear relationship between the speech envelope and the evoked neural response. A simple linear decoder is probably not a good fit to handle all the complexity of the auditory system and the brain (Vanthornhout et al., 2018).

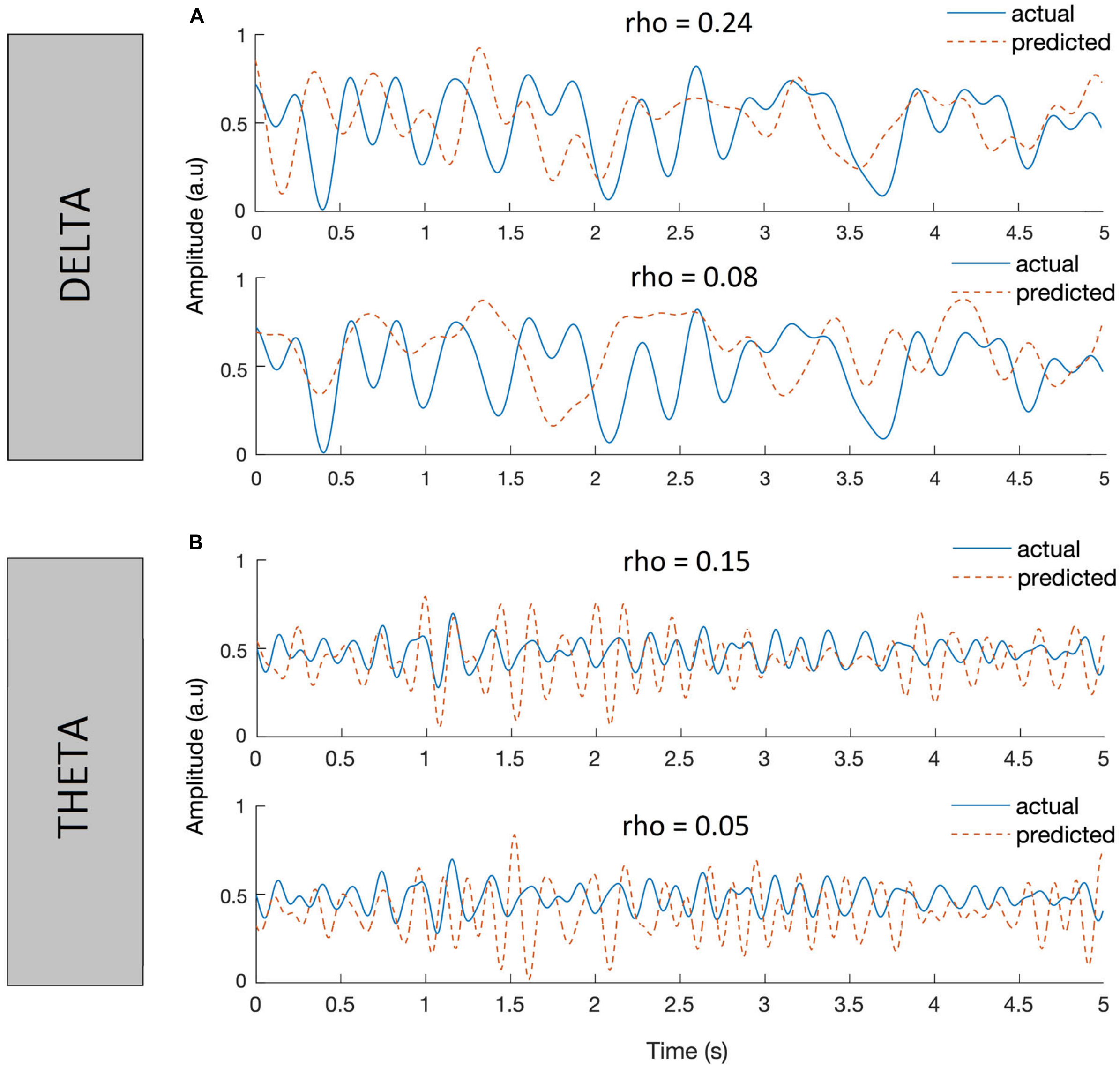

FIGURE 9
Plots showing instances of actual speech envelopes (blue) and the reconstructed envelopes (red dotted) and bootstrapped Spearman correlation (rho) between them in delta **(A)** and theta **(B)** bands.

Also, neurons are known to respond to a complex stimulus like speech in a non-linear manner (Theunissen et al., 2000) and most likely, a non-linear decoder is needed to reconstruct the speech feature more accurately from the cortical responses. This idea is also supported by an EEG study conducted by Yang et al. (2015) which showed that compared with the linear regression model, the reconstructed spectrograms from the deep neural network achieved a higher average correlation with the actual spectrograms. Additionally, it is still not clear whether the envelope was a good representation of speech relevant to the perception of speech quality in normal-hearing listeners. Therefore, in the future, other speech features such as the spectrogram and phoneme-related features can be included as well in an attempt to improving the performance of a decoder (e.g., Di Liberto et al., 2015; Lesenfants et al., 2019a).

Perceived sound quality vs. envelope entrainment to spoken speech passages by cochlear implant talkers and normal-hearing talkers

By looking at the results of behavioral sound quality ratings, we inferred that the higher average correlation

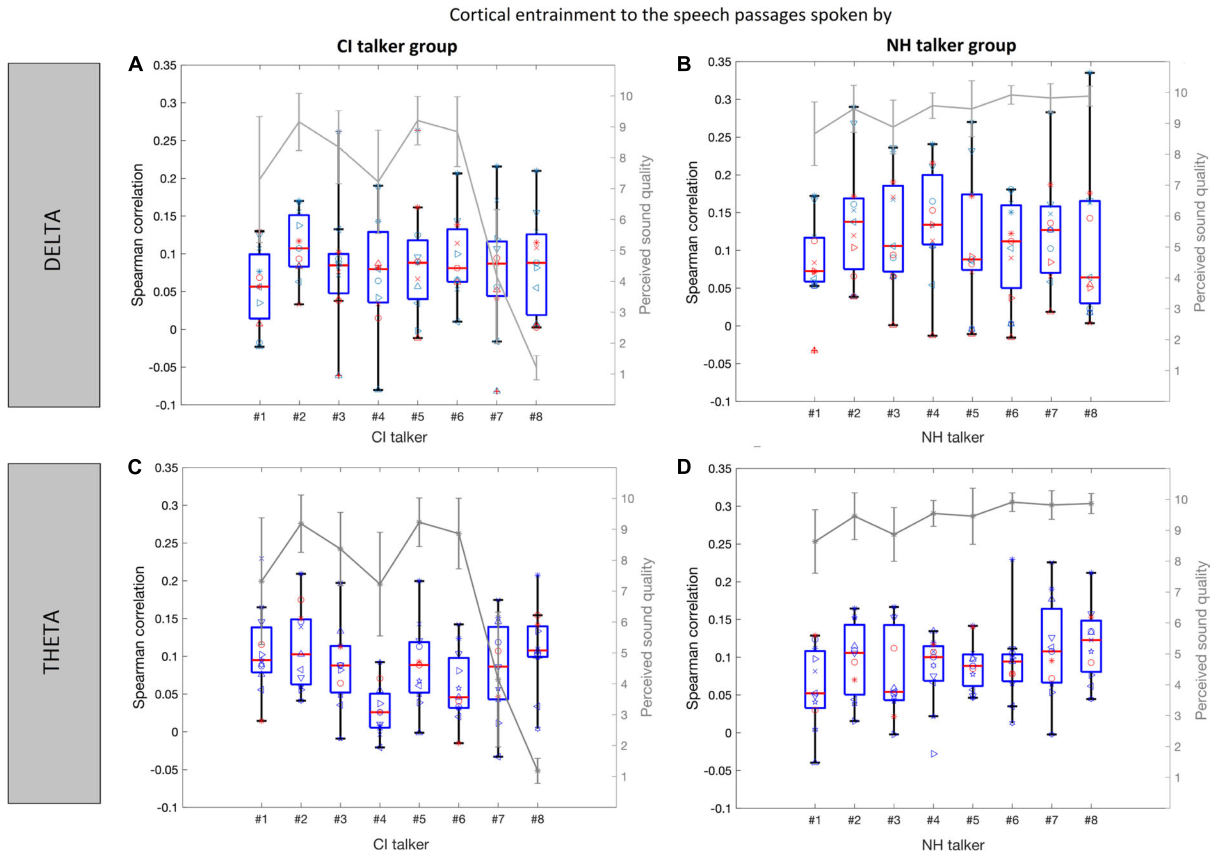

FIGURE 10
Cortical entrainment to speech envelope (filtered to match delta and theta bandwidth) in the normal-hearing listeners is shown for each of the speech passages from the CI talker group **(A,C)** and NH talker group **(B,D)** in delta and theta band, respectively. The data are represented in the form of the boxplots depicting medians (center mark) and interquartile ranges (the bottom and top edges of the box indicate the 25th and 75th percentiles). The boxplot shows the entrainment data variance across listeners, and the individual listenerŠs data points are plotted over it. Each listener is shown with a different symbol. The perceived sound quality curve (gray) is also shown for comparison.

between the actual speech envelope and the reconstructed envelope was found for the NH talker group who produced the speech with higher sound quality. This helped to prove our hypothesis that closer cortical tracking of speech envelope (higher correlation) was observed when speech was of higher perceived sound quality. However, we did not associate the sound quality ratings with the envelope entrainment results to find the relationship between them due to the difference in which the behavioral experiment was conducted compared to the EEG experiment. The behavioral data were collected when listeners were actively listening to speech passages, whereas EEG data were obtained when listeners were listening passively. Cortical entrainment has been seen in both active and passive paradigms of listening (for review, see Ding and Simon, 2014). A number of studies have shown that the cortical entrainment to speech is strongly modulated by attention and it has been shown that the reconstructed envelope depends strongly on the attentional focus of the listener and resembles the envelope of the attended speech (Kerlin et al., 2010; Ding and Simon, 2012). Hence, it is known that passive/active listening (Di Liberto et al., 2015; O'Sullivan et al., 2015) to speech affects the degree of cortical entrainment, and the difference in the two experimental setups would make it more difficult to extract appropriate conclusions about differences or similarities between behavioral and EEG results and this is one of the limitations of this study. In the future, the neural responses of the same normal hearing listeners using EEG while they paid attention to the speech passages can be recorded and how well the perceived sound quality ratings are correlated with the entrainment results can be analyzed. Our previous study (Akbarzadeh et al., 2021) has psychophysically validated the perceived sound quality measure with consistency to speech recognition scores and cortical entrainment outcome with normal hearing and hearing impaired listeners. The perceived sound quality measure

10.3389/fnins.2022.927872

TABLE 4 Demographic details of the selected CI talkers from the deaf speech corpus showing the mean sound quality rating of their spoken speech passages.

CI talker	Age (years)	Gender	Age of first amplification use (years)		Onset of hearing loss	Current type of amplification		Communication mode[4]	Mean perceived sound quality
			Right ear	Left ear		Right ear	Left ear		
#1	37	Male	5	5	Post-lingual	CI[2]	CI	Oral	7.3
#2	38	Female	28	NA[1]	Post-lingual	CI	NA	Oral	9.2
#3	16	Female	0.5	2	Pre-lingual	HA[3]	CI	Oral	8.4
#4	77	Male	18	73	Post-lingual	HA	CI	Oral and sign	7.2
#5	62	Female	38	38	Post-lingual	CI	HA	Oral and sign	9.2
#6	60	Male	52	52	Post-lingual	HA	CI	Oral and sign	8.9
#7	57	Female	3	58	Post-lingual	HA	CI	Oral	4.1
#8	33	Male	NA	3	Post-lingual	NA	CI	Sign only	1.2

[1]NA, not applicable. [2]CI, cochlear implant. [3]HA, hearing aid. [4]Oral indicates that the talker used oral speech and language. Sign indicates that the talker used sign language. Gray highlight: Higher mean perceived sound quality. Yellow highlight: Lower mean perceived sound quality.

adopted in the study is limited to the same perspective as previously studied. More work on the perceived sound quality measure will continue to cover a wider perspective of the measure.

Conclusion

The present study shows that speech envelope is well-represented neurophysiologically. The speech envelope reconstructed from EEG using the regenerative model (Crosse et al., 2016) shows similarities to the actual speech envelope, particularly when the speech envelope is filtered to match the EEG bandwidth (delta and theta) of interest. Perceived sound quality ratings by the 11 normal-hearing listeners were found to be associated with the cortical activity involved in tracking the speech envelope. Closer tracking of the speech envelope, with higher correlations between the actual speech envelope and the envelope reconstructed from the EEG, was obtained in response to speech produced by NH talkers relative to CI talkers. Our results also show that the perceived sound quality differences rated by the normal-hearing listeners between speech passages spoken by CI talkers and NH talkers can be seen in the cortical tracking of the speech envelope in the same listeners.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board at the University of Texas at Dallas. The patients/participants provided their written informed consent to participate in this study.

Author contributions

SR, BM, SL, and C-TT were involved in conceptualization, data collection and analysis, and manuscript preparation. HC

was involved in data collection. All authors contributed to the article and approved the submitted version.

Funding

BM was supported by PSC-CUNY.

Acknowledgments

We would like to thank the authors of Mendel et al. (2017) especially Dr. Lisa Lucks Mendel (School of Communication Sciences and Disorders, University of Memphis) for the data courtesy.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aiken, S. J., and Picton, T. W. (2008). Human cortical responses to the speech envelope. *Ear Hear.* 29, 139–157. doi: 10.1097/AUD.0b013e31816453dc

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Control* 19, 716–723.

Akbarzadeh, S., Lee, S., Chen, F., and Tan, C.-T. (2021). The effect of speech and noise levels on the quality perceived by cochlear implant and normal hearing listeners. *Speech Commun.* 132, 106–113.

Antons, J. N., Porbadnigk, A. K., Schleicher, R., Blankertz, B., Möller, S., and Curio, G. (2010). "Subjective listening tests and neural correlates of speech degradation in case of signal-correlated noise," in *Proceedings of the Audio Engineering Society (AES)*, Montreux.

Antons, J. N., Schleicher, R., Arndt, S., Möller, S., Porbadnigk, A. K., and Curio, G. (2012). Analyzing speech quality perception using electroencephalography. *IEEE J. Sel. Top. Signal Process.* 6, 721–731.

Bates, D. M., Mäechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48.

Bigdely-Shamlo, N., Kreutz-Delgado, K., Kothe, C., and Makeig, S. (2013). "Eyecatch: data-mining over half a million EEG independent components to construct a fully-automated eye-component detector," in *Proceedings 35th Annual International Conference IEEE Engineering in Medicine and Biology Society*, Manhattan, NY. doi: 10.1109/EMBC.2013.6610881

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Hillsdale, NJ: Erlbaum.

Crosse, M. J., Di Liberto, G. M., Bednar, A., and Lalor, E. C. (2016). The Multivariate Temporal Response Function (mTRF) Toolbox: a MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. *Front. Hum. Neurosci.* 10:604. doi: 10.3389/fnhum.2016.00604

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21.

Di Liberto, G. M., O'Sullivan, J. A., and Lalor, E. C. (2015). Low-frequency cortical entrainment to speech reflects phoneme-level processing. *Curr. Biol.* 25, 2457–2465. doi: 10.1016/j.cub.2015.08.030

Ding, N., and Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11854–11859. doi: 10.1073/pnas.1205381109

Ding, N., and Simon, J. Z. (2013). Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. *J. Neurosci.* 33, 5728–5735. doi: 10.1523/JNEUROSCI.5297-12

Ding, N., and Simon, J. Z. (2014). Cortical entrainment to continuous speech: functional roles and interpretations. *Front. Hum. Neurosci.* 8:311. doi: 10.3389/fnhum.2014.00311

Gautam, A., Naples, J. G., and Eliades, S. J. (2019). Control of speech and voice in cochlear implant patients. *Laryngoscope* 129, 2158–2163.

Goswami, U., and Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives? *Lab. Phonol.* 4, 67–92. doi: 10.1515/lp-2013-0004

Kerlin, J. R., Shahin, A. J., and Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a "cocktail party". *J. Neurosci.* 30, 620–628. doi: 10.1523/JNEUROSCI.3631-09.2010

Kim, H., Kang, W. S., Park, H. J., Lee, J. Y., Park, J. W., Kim, Y., et al. (2018). Cochlear implantation in postlingually deaf adults is time-sensitive towards positive outcome: prediction using advanced machine learning techniques. *Sci. Rep.* 8:18004. doi: 10.1038/s41598-018-36404-1

Kong, Y. Y., Somarowthu, A., and Ding, N. (2015). Effects of spectral degradation on attentional modulation of cortical auditory responses to continuous speech. *J. Assoc. Res. Otolaryngol.* 16, 783–796. doi: 10.1007/s10162-015-0540-x

Lalor, E. C., and Foxe, J. J. (2010). Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *Eur. J. Neurosci.* 31, 189–193. doi: 10.1111/j.1460-9568.2009.07055.x

Lesenfants, D., Vanthornhout, J., Verschueren, E., Decruy, L., and Francart, T. (2019a). Predicting individual speech intelligibility from the cortical tracking of acoustic- and phonetic-level speech representations. *Hear. Res.* 380, 1–9. doi: 10.1016/j.heares.2019.05.006

Lesenfants, D., Vanthornhout, J., Verschueren, E., and Francart, T. (2019b). Data-driven spatial filtering for improved measurement of cortical tracking of multiple representations of speech. *J. Neural Eng.* 16:066017.

Loizou, P. (2011). "Speech quality assessment," in *Multimedia analysis, processing and communications*, eds W. Lin, D. Tao, J. Kacprzyk, Z. Li, E. Izquierdo, and H. Wang (Berlin: Springer-Verlag), 623–654.

Martin, B. A., Sigal, A., Kurtzberg, D., and Stapells, D. R. (1997). The effects of decreased audibility produced by high-pass noise masking on cortical event-related potentials to speech sounds/ba/and/da/. *J. Acoust. Soc. Am.* 101, 1585–1599. doi: 10.1121/1.418146

Martin, B. A., and Stapells, D. R. (2005). Effects of low-pass masking on auditory event-related potentials to speech. *Ear Hear.* 26, 195–213.

Mendel, L. L., Lee, S., Pousson, M., Patro, C., McSorley, S., Banerjee, B., et al. (2017). Corpus of deaf speech for acoustic and speech production research. *J. Acoust. Soc. Am.* 142:EL102.

Oostenveld, R., and Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clin. Neurophysiol.* 112, 713–719. doi: 10.1016/s1388-2457(00)00527-7

O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinncunningham, B. G., et al. (2015). Attentional selection in a cocktail party environment can be decoded from single-trial eeg. *Cereb. Cortex* 25, 1697–1706. doi: 10.1093/cercor/bht355

Peelle, J. E., and Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Front. Psychol.* 3:320. doi: 10.3389/fpsyg.2012.00320

Porbadnigk, A. K., Treder, M. S., Blankertz, B., Antons, J. N., Schleicher, R., Möller, S., et al. (2013). Single-trial analysis of the neural correlates of speech quality perception. *J. Neural Eng.* 10:056003. doi: 10.1088/1741-2560/10/5/056003

Ruff, S., Bocklet, T., Nöth, E., Müller, J., Hoster, E., and Schuster, M. (2017). Speech production quality of cochlear implant users with respect to duration and onset of hearing loss. *ORL* 79, 282–294. doi: 10.1159/000479819

Sangthong, M. (2020). The Effect of the Likert Point Scale and Sample Size on the Efficiency of Parametric and Nonparametric Tests. *Thailand Stat.* 18, 55–64.

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303–304.

Theunissen, F. E., Sen, K., and Doupe, A. J. (2000). Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J. Neurosci.* 20, 2315–2331.

Uhrig, S., Mittag, G., Möller, S., and Voigt-Antons, J.-N. (2019a). P300 indicates context-dependent change in speech quality beyond phonological change. *J. Neural Eng.* 16:066008. doi: 10.1088/1741-2552/ab1673

Uhrig, S., Mittag, G., Möller, S., and Voigt-Antons, J.-N. (2019b). Neural correlates of speech quality dimensions analyzed using electroencephalography (EEG). *J. Neural Eng.* 16:036009. doi: 10.1088/1741-2552/aaf122

Vanthornhout, J., Decruy, L., Wouters, J., Simon, J. Z., and Francart, T. (2018). Speech intelligibility predicted from neural entrainment of the speech envelope. *J. Assoc. Res. Otolaryngol.* 19, 181–191.

Yang, M., Sheth, S. A., Schevon, C. A., McKhann, I. I. G. M., and Mesgarani, N. (2015). "Speech reconstruction from human auditory cortex withdeep neural networks," in *Proceedings of the Sixteenth Annual Conference of the International Speech Communication Association*, Dresden.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 02 September 2022
DOI 10.3389/fpsyg.2022.927854

OPEN ACCESS

EDITED BY
Fei Chen,
Southern University of Science and Technology, China

REVIEWED BY
Hilal Dincer D'Alessandro,
Hacettepe University, Turkey
Andreu Paredes-Gallardo,
Technical University of Denmark, Denmark

*CORRESPONDENCE
Yingjiu Nie
nieyx@jmu.edu

SPECIALTY SECTION
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

RECEIVED 25 April 2022
ACCEPTED 11 August 2022
PUBLISHED 02 September 2022

CITATION
Matz AF, Nie Y and Wheeler HJ (2022) Auditory stream segregation of amplitude-modulated narrowband noise in cochlear implant users and individuals with normal hearing.
Front. Psychol. 13:927854.
doi: 10.3389/fpsyg.2022.927854

Auditory stream segregation of amplitude-modulated narrowband noise in cochlear implant users and individuals with normal hearing

Alexandria F. Matz[1], Yingjiu Nie[2]* and Harley J. Wheeler[3]

[1]Department of Otolaryngology, Eastern Virginia Medical School, Norfolk, VA, United States, [2]Department of Communication Sciences and Disorders, James Madison University, Harrisonburg, VA, United States, [3]Department of Speech-Language-Hearing Sciences, University of Minnesota, Twin Cities, Minneapolis, MN, United States

Voluntary stream segregation was investigated in cochlear implant (CI) users and normal-hearing (NH) listeners using a segregation-promoting objective approach which evaluated the role of spectral and amplitude-modulation (AM) rate separations on stream segregation and its build-up. Sequences of 9 or 3 pairs of A and B narrowband noise (NBN) bursts were presented which differed in either center frequency of the noise band, the AM-rate, or both. In some sequences (delayed sequences), the last B burst was delayed by 35ms from their otherwise-steady temporal position. In the other sequences (no-delay sequences), the last B bursts were temporally advanced from 0 to 10 ms. A single interval yes/no procedure was utilized to measure participants' sensitivity (d') in identifying delayed vs. no-delay sequences. A higher d' value showed the higher ability to segregate the A and B subsequences. For NH listeners, performance improved with each spectral separation. However, for CI users, performance was only significantly better for the condition with the largest spectral separation. Additionally, performance was significantly poorer for the largest AM-rate separation than for the condition with no AM-rate separation for both groups. The significant effect of sequence duration in both groups indicated that listeners made more improvement with lengthening the duration of stimulus sequences, supporting the build-up effect. The results of this study suggest that CI users are less able than NH listeners to segregate NBN bursts into different auditory streams when they are moderately separated in the spectral domain. Contrary to our hypothesis, our results indicate that AM-rate separation may interfere with the segregation of streams of NBN. Additionally, our results add evidence to the literature that CI users build up stream segregation at a rate comparable to NH listeners, when the inter-stream spectral separations are adequately large.

KEYWORDS

auditory stream segregation, cochlear implants, narrowband noise, spectral separation, amplitude modulation, build-up effect

Introduction

Auditory stream segregation (also known as auditory streaming) refers to the process that allows listeners to interpret multiple sounds coming from different sources and assign those sounds to individual sound generators (Moore and Gockel, 2012). For example, normal-hearing (NH) listeners use stream segregation abilities to separate a talker at a noisy party or isolating the violin among the other instruments in an orchestra (Bregman, 1990, Chapter 1). Stream segregation has been shown to be related to the degree of the perceptual differences across sound streams (Moore and Gockel, 2002, 2012) in various domains such as frequency (e.g., Bregman and Campbell, 1971; Warren and Obusek, 1972; Dannenbring and Bregman, 1976), amplitude-modulation rate (AM-rate; e.g., Grimault et al., 2002; Nie and Nelson, 2015), pitch (e.g., Assmann and Summerfield, 1987; de Cheveigné, 1997), etc. When the acoustical inter-stream differences are adequately prominent, NH listeners may perceive separated auditory streams without voluntarily directing their attention to segregating the streams (e.g., van Noorden, 1975). This process is referred to as obligatory segregation and is generally noted to be driven by the stimulus (Bregman, 1990, Chapter 4). Conversely, when these differences are indistinct, listeners will experience obligatory integration, where they perceive only one auditory stream even when attempting to segregate signals into different streams (Bregman, 1990, Chapter 4). When the salience of these differences is ambiguous for the obligatory processing, NH listeners can intentionally direct their attention to perceptually separating or integrating auditory streams (van Noorden, 1975; Bregman, 1990, Chapter 4). These top-down processes are referred to as voluntary segregation and voluntary integration, respectively (Bregman, 1990, Chapter 4).

In everyday life, listeners are in complex auditory scenes where the differences between concurrent signals are often ambiguous, such as conversing in a restaurant or a cocktail party with varying background noises. Therefore, voluntary segregation is frequently employed by listeners to differentiate the interested auditory stream from interferences. Presumably, cochlear implants (CI) users need to engage in voluntary segregation in more listening conditions than do NH listeners, as the degraded auditory cues from cochlear implants may result in increased ambiguity of the differences between auditory streams (e.g., Fu and Nogaki, 2005). Despite such frequent adoption of voluntary stream segregation by CI users, this process remains poorly understood in many aspects. Particularly, the literature is lacking in comparisons between CI users and NH listeners in their ability to voluntarily segregate sound sequences based on the same inter-sequence acoustical differences. Findings of such comparisons may improve the understanding of the cues CI users utilize for sequential segregation with the facilitation of focused attention, which is relevant to speech perception in noise in their daily life (Nogueira and Dolhopiatenko, 2022). The current study was aimed to fill in this gap by comparing the two groups with manipulations of three acoustic attributes of the sound sequences—frequency, amplitude-modulation rate (referred to as AM-rate hereafter), and duration of the sequence.

In laboratory research, auditory stream segregation has been assessed using both subjective and objective paradigms. In subjective paradigms, listeners report their perception of the number of sound streams perceived. In objective paradigms, stream segregation is indexed by behavioral performance in the purportedly-designed listening tasks that can assess either voluntary or obligatory stream segregation. To assess voluntary segregation, the listening tasks may be arranged to be segregation-facilitating, such that they presumably require listeners to make effort to segregate auditory streams to achieve better performance. Thus, better performance in the segregation-facilitating objective paradigms indexes stronger voluntary segregation. To assess obligatory segregation, the listening tasks may be arranged to be integration-facilitating, in other words, segregation-hindering, such that listeners are awarded with better task performance for their mental effort made to integrate the auditory streams. Specifically, better performance in the integration-facilitating objective paradigms indexes stronger obligatory segregation.

One of the drawbacks with using subjective paradigms is the subject bias, such as listeners adopting different perceptual criteria for reporting stream segregation. For CI users, this bias may be partly attributed to the uncertainty of their discernment of auditory streams. CI users are provided signals by way of electrical stimulations with degraded auditory cues, unlike NH listeners. As a result, it is unclear whether CI users comprehend the concept of auditory streams consistently both within the group and when compared with NH listeners. Hence, the subjective reports of stream segregation may not be based on the same perception between CI users and NH listeners as well as among CI users. In contrast, with an objective paradigm, listeners are not required to comprehend the concept of auditory streams. Rather, the listening task typically requires listeners to follow the elements of the stimulus sequences over the course of each presentation and perceptually group relevant elements sequentially and separate the groups into different running auditory streams. As a result, the objective paradigms can reduce the subject bias associated with the subjective paradigms. Additionally, listeners' desire of providing highest possible performance in the objective paradigms tend to motivate them to execute at their highest capacity. When a separation-facilitating task is used as an objective approach to study voluntary stream segregation, this motivational aspect would elicit stream segregation ability to its highest level. Considering the aforementioned advantages of an objective paradigm, this study employed a segregation-facilitating task which was modified from the task reported in Nie and Nelson (2015). The direction of focused attention on segregation for better performance in the task resembled the top-down processing for speech perception in background noise where listeners direct their attention selectively to interested speech instead of background noise.

Auditory stream formation has been shown to be dependent on the amount of time the target sequence is presented (e.g., Anstis and Saida, 1985; Cusack et al., 2004). The tendency for segregation to occur increases with longer exposure time to the sound sequence. In other words, auditory stream segregation builds up over time. Using subjective methods, researchers have estimated that stream segregation builds up rapidly over about 10 s, then builds more slowly up to at least 60 s (Moore and Gockel, 2012). While the time course of build-up segregation has not been assessed in CI users using objective methods, our past study (Nie and Nelson, 2015) has shown that the build-up can be observed in a period of 3.1 s for NH listeners with a segregation-facilitating objective paradigm. Nie and Nelson (2015) also revealed that both spectral and AM-rate separations could be cues for the build-up; that is, larger inter-stream separations in either spectrum or AM-rate are associated with greater increases of stream segregation over the course of 3.1 s.

It has been hypothesized that CI users are inferior to NH listeners in stream segregation abilities (e.g., Hong and Turner, 2006; Böckmann-Barthel et al., 2014). Growing evidence has supported this hypothesis for obligatory segregation. For example, using objective segregation-hindering paradigms, some works (e.g., Tejani et al., 2017) have shown that CI users experience weaker perceptual segregation than NH listeners for the same amount of inter-stream frequency separation. Additionally, other works (Cooper and Roberts, 2007, 2009) have argued the absence of obligatory segregation in CI users. Generally, the aforementioned findings suggest that CI users have a lower capacity of perceiving salience of inter-stream differences than NH listeners, which can be attributed to the degradation of auditory signals through cochlear implants. Degraded auditory signals tend to present ambiguous cues for stream segregation which may be modulated by voluntary attention. A few works (e.g., Böckmann-Barthel et al., 2014) have used subjective paradigms to show that CI users are able to form segregated auditory streams even with these ambiguous cues. To our knowledge, only one research group (Paredes-Gallardo et al., 2018a,b,c) has specifically studied voluntary stream segregation in CI users. Using an objective segregation-facilitating paradigm with direct electrical pulse stimuli, this group concluded that CI users were able to voluntarily segregate auditory streams of pulses when the inter-stream difference is in electrode position (Paredes-Gallardo et al., 2018b) or in pulse rate (Paredes-Gallardo et al., 2018c), even when the differences were ambiguous.

These findings suggest that CI users are able to segregate auditory streams based on spectral differences—which is related to the inter-stream electrode-position differences, and temporal-pitch separations—a cue elicited by the inter-stream pulse rate separation. While the direct electrical stimuli in Paredes-Gallardo et al. studies (Paredes-Gallardo et al., 2018b,c) allowed more robust control of stimuli, the acoustic signals and direct electrical stimuli do not activate the electrodes in the identical manner. Thus, the ability to utilize spectral and temporal-pitch cues remains to be examined with acoustic stimuli in CI users. Segregating auditory streams has been suggested to contribute to speech recognition in noise in CI users (Hong and Turner, 2006). A wealth of research has shown that CI users are more vulnerable than NH listeners to distractions when recognizing speech (e.g., Cullington and Zeng, 2008) and investigated various underlying mechanisms for CI users' higher vulnerability (e.g., Oxenham and Kreft, 2014; Goehring et al., 2021). However, no known research has compared voluntary segregation between CI users and NH listeners to study the mechanism of sequential processing. Examining whether the two groups can attain a comparable level of segregation based on the identical acoustical cues would have implications, from the sequential processing aspect, on understanding CI users' vulnerability when recognizing speech in noise.

The current study was aimed to evaluate the spectral and temporal-pitch cues for voluntary stream segregation and for the build-up of stream segregation in CI users in comparison to NH listeners. We adopted stimulus constructs and procedures similar to those in Nie and Nelson (2015) with modifications. Specifically, the study was conducted using a segregation-facilitating objective paradigm with stimulus sequences of narrowband noise (NBN) that was amplitude modulated. The stimulus sequences differed in either the frequency region of the NBN or AM-rate for the purpose of examining spectral and temporal-pitch cues for stream segregation, respectively. Each noise band was manipulated such that its bandwidth was constrained within the single excited auditory peripheral filter for NH listeners. For CI users, the bandwidth was restricted to be within the frequency passband of the assigned electrode according to their individual clinical MAP. This manipulation allowed the degrees of inter-stream spectral separation between NH and CI users to be similar for the acoustic stimuli, while limiting the cochlear regions stimulated by the electrical output of the NBN for CI users. However, the inter-stream spectral separations in the internal electrical stimulation for CI users are effectively reduced resulting from the nature that certain noise bands activating other electrodes beside the assigned one (for details, see the electrodogram in Materials and methods). Findings of the current study revealed the effects of reduced salience of internal inter-stream spectral separation on stream segregation.

The aim of studying the build-up stream segregation was motivated by poorly understood listening challenges faced by CI users. For example, whether CI users take a longer time to separate running auditory sequences, which has important implications for CI users' listening in real life. Studying the build-up stream segregation in both CI and NH groups will allow us to compare how fast the two groups can separate auditory streams of the same acoustic stimuli. Together, the aims of this study include using a segregation-facilitating objective approach to compare voluntary stream segregation abilities with NBN noise in NH listeners and CI users based on inter-stream spectral separations and AM-rate separations, as well as build-up segregation.

Materials and methods

Participants

Ten adult listeners between 18 and 69 years of age, three female, seven male, participated in the study. They were divided into two groups: post-lingually deafened cochlear implant (CI) users (6 participants) aged 24–69 years with a mean age of 52.5 years, and normal-hearing (NH) listeners (4 participants) aged 22–60 with a mean age of 37.8 years. All NH listeners had symmetric hearing thresholds no greater than 25 dB HL at audiometric frequencies of 250 to 8,000 Hz, and no greater than a 10 dB difference between ears at the same frequency. All CI users wore only one cochlear implant; if they were bilateral users, they wore the CI on the side perceived to be dominant. All CI participants had no residual hearing, expect one who was a bimodal listener. This bimodal listener did not use their hearing aid in the other ear that was blocked with a foam earplug to avoid the effect of residual acoustic hearing. Table 1 illustrates the demographics of each CI user.

Apparatus

The stimuli were generated live using a customized MATLAB (R2013a) script at a sampling rate of 44,100 Hz, then processed through a Lynx 22 soundcard installed in a Dell Optiplex 9010 computer, which ran through a DAC1 device. The analog output of the DAC1 was amplified *via* a Tucker Davis Technologies, TDT RZ6 system and presented through a Klipsch RB-51 bookshelf speaker. Stimulus presentation and response recording was controlled by the MATLAB script in conjunction with PsychToolbox (version 3; Brainard, 1997; Pelli, 1997). To record the participants' responses, an RTbox (Li et al., 2010) was used as the hardware interface. Participants were seated in a sound-attenuated booth at 0° azimuth at a 1-meter distance from the speaker.

Stimulus sequences

The stimulus paradigm consisted of sequences of 9 or 3 pairs of A and B noise bursts, in the pattern of ABABAB…. The paradigm is illustrated in Figure 1. The A and B bursts were NBN generated by passing broadband Gaussian noise through 10th order Butterworth filters with center frequencies and bandwidths described in the following paragraph. In some conditions, amplitude modulation was superimposed on A and B bursts. The A and B bursts differed in either center frequency of the noise band, the AM-rate, or both. The duration of each A or B burst was 80 ms including 8 ms rise/fall ramps. A 50-ms silent gap was included between the offset of a burst to the onset of the next with the A bursts (except the initial one in a stimulus sequence) jittering from their nominal temporal location. The amount of jitter was randomly drawn for each jittered A burst from a rectangular distribution between 0 to 40 ms. In other words, the B bursts were presented steadily with a 180-ms gap between the two consecutive ones, except for the last B burst, in any sequence, while the offset-to-onset gap between an A burst and either adjacent B burst ranged between 10 and 90 ms. In some sequences, namely *delayed sequences*, the last B bursts were delayed from their otherwise-steady temporal position by 35 ms; in the other sequences, namely *no-delay sequences*, the last B bursts were temporally advanced by an amount randomly drawn from the rectangular distribution ranging from 0 to 10 ms. As a result, the *delayed sequences* were 2.325 and 0.665 s in duration when the sequences, respectively, consisted of 9 and 3 pairs of A and B bursts, while the *no-delay sequences* were 2.28–2.29 s and 0.62–0.63 s when consisting of 9 and 3 AB pairs, respectively.

The longer sequences were shortened by approximately 1 s from those in Nie and Nelson (2015) to allow the study of build-up within the time course of 3 s. This was designed to address the question raised in Böckmann-Barthel et al. (2014). Using a subjective approach, Böckmann-Barthel et al. examined the time course for CI users to build up stream segregation of sequences of interleaved harmonic tone complexes differed by varied amounts of f0. They noted that the CI users rarely provided first response

TABLE 1 Participants' demographics and electrodes to which A and B noise bands were mapped in the moderate and large A-B spectral separations. The center frequencies (CF) of the noise band is also shown.

Participant code	Age (Years)	CI brand	Noise band (B)		Noise band (A)			
					Moderate A–B spectral separation		Large A–B spectral separation	
			Electrode #	CF (Hz)	Electrode #	CF (Hz)	Electrode #	CF (Hz)
CI1	53	Cochlear	11	1808	8	2,927	2	6,418
CI2	69	Cochlear	12	1,683	8	2,871	2	6,485
CI3	69	MED-EL	7	1,632	10	3,064	12	7,352
CI4	43	Cochlear	11	1741	7	3,092	2	6,828
CI5	24	Cochlear	12	1,683	8	2,871	2	6,485
CI6	57	Cochlear	12	1,683	8	2,871	2	6,485

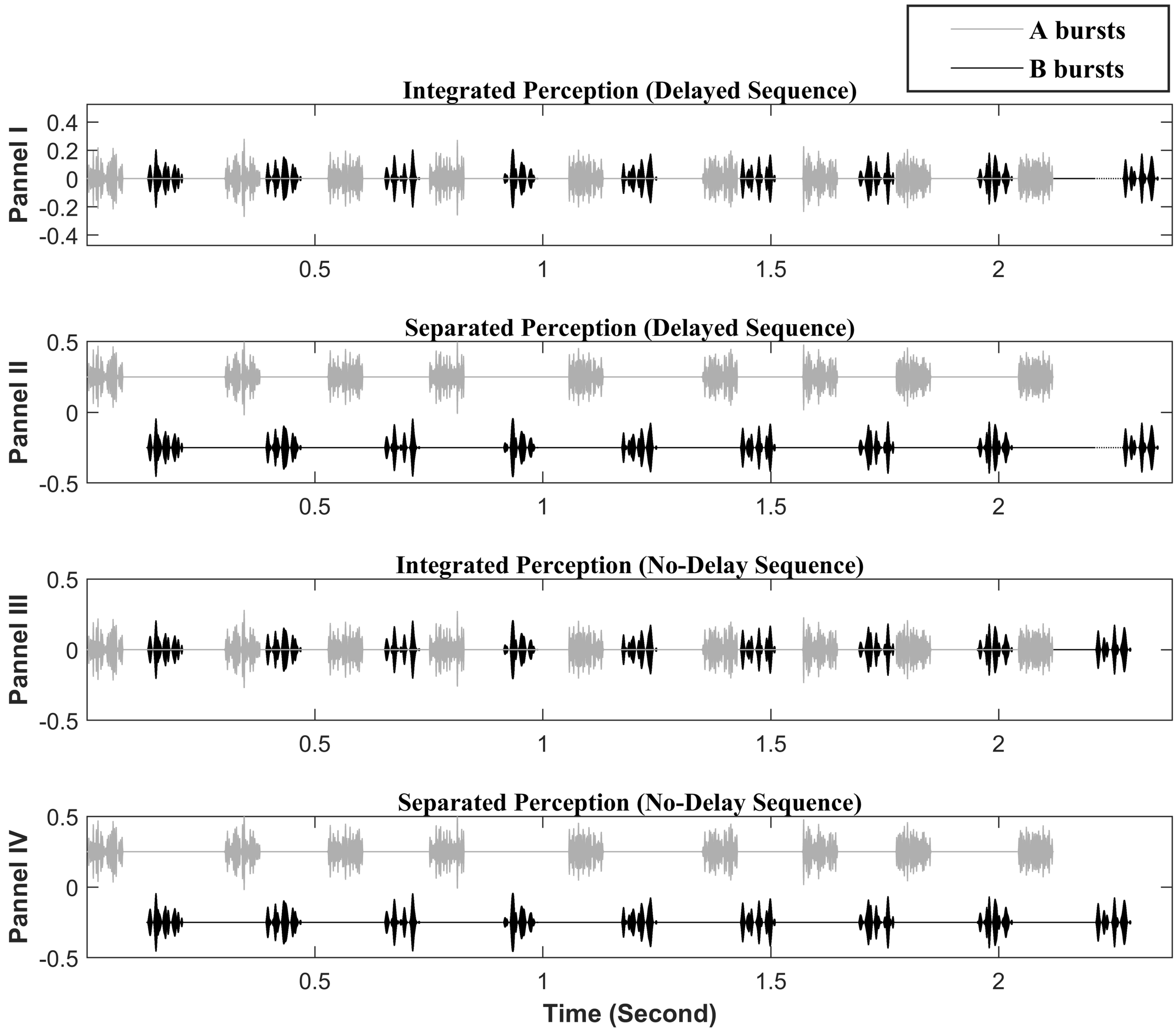

FIGURE 1
Illustration of the stimulus paradigm (adopted from Nie and Nelson, 2015, with permission). The black segments represent the B bursts that form the attended target subsequence. The light gray segments represent the A bursts that form the unattended distracting subsequence. Panels I and II illustrate the delayed sequences: The dark dotted lines to the left side of the last B burst show the delay of 35ms for the last B burst. Panels III and IV illustrate the no-delay sequences. The I and III panels depict a visual representation of the integrated perception, while the II and IV panels depict the segregated perception. The large A–B spectral separation is depicted here. The AM rates shown on the A bursts and B bursts are 300Hz and 50Hz, respectively. The depicted sequences all consist of 9 pairs of A and B bursts with a duration of 80ms for each burst. The onset-to-onset time between he first A and B bursts is 130ms consisting of a silent gap of 50ms between the offset of A and the onset of B. The B–B onset-to-onset time is 260ms; in other word, the B–B offset-to-onset time is 180ms as noted in the text. Note, the last eight A bursts are temporally jittered between the two consecutive B bursts resulting in the offset-to-onset gap between an A burst and an adjacent B burst ranging between 10 and 90ms.

within 3 s of the stimulus onset to indicate the number of streams they perceived. The authors argued that the initial time point for studying the course of build-up, in turn, should be normalized to the timepoint when the first response was provided. With the approach of normalizing the initial timepoint, the study has shown that CI users may not require time to build up segregation when the f0 difference between the two streams is substantially large. The authors noted that this trend was comparable to NH listeners who showed the no build-up required when segregating streams of harmonic tone complexes which had large frequency differences (Deike et al., 2012). However, Deike et al. (2012) did not undertake the normalized initial timepoint approach and showed that NH listeners reported perceiving segregation within 2–3 s of the onset of a stimulus sequence. Thus, it remained unclear how CI users would compare with NH listeners in build-up segregation within 2–3 s of stimulus onset. A sequence shorter than 3 s with an objective paradigm allowed us to examine this process in the current study.

Three levels of spectral separation between A and B bursts—no-separation, moderate separation, and large separation—were examined. For the NH listeners, by holding the center frequency of B bursts constant at 1,803 Hz, these A–B spectral separations were achieved by setting the center frequency of the A bursts, respectively, at 1,803, 3,022, or 6,665 Hz. As a result, the center frequencies of A and B bands differed by 0.75 and 1.89 octaves for the moderate and large A–B spectral separations, respectively. The three center frequencies coincided with the center frequencies mapped to the 10, 13, and 16th electrodes through typical 16-channel signal processing strategies of the Advanced Bionics technology, thus were selected for NH listeners to simulate separations of assigned channels between A and B bursts around 3 or 6 electrodes for CI users. Noise bursts were set to the narrowest bandwidths allowing a steady presentation level in the sound field in the participant's location (Walker et al., 1984). Subsequently, bandwidths of 162 Hz were applied for the noise bands centered at 1,803 and 3,022 Hz, and 216 Hz for noise band centered at 6,665 Hz.

For the CI users, the center frequency of B bursts was customized for each listener so that it coincided with the center frequency of the signal processing channel in which the 1,803 Hz was allocated in the listener's clinical MAP. Likewise, the center frequency of the A bursts was customized for each CI users to achieve the three spectral separations. As a result, at the moderate and large A–B spectral separations, the center frequencies of the noise bands separated by 0.70–0.91 octaves and 1.83–2.17 octaves, respectively. Table 1 shows the center frequencies and electrode numbers for the A and B bursts at three levels of A–B spectral separation for each CI user. Through the CI processing, in addition to the assigned electrode, a given noise band effectively activated a number of other electrodes. An electrodogram was computed for the B burst and the two frequency regions of A bursts based on the most common frequency allocation among all CI users. Figure 2 shows the electrodogram demonstrating that a given NBN stimulus activated a total of four or five neighboring electrodes adjacent to the assigned electrode. Specifically, the B burst activated electrodes #9, 10, 11, 12, and 13, while the A bursts activated electrodes #6, 7, 8, 9, and 10 in the moderate spectral separation and electrodes #1, 2, 3, and 4 in the large spectral separation.

The AM-rate alternatives for A bursts were 0 (i.e., no amplitude modulation), 200, or 300 Hz. B bursts were presented either at an AM-rate of 0 or 50 Hz. Three possible AM-rate separations between A and B bursts included no separation wherein both bursts were not modulated (AM0-0), 2-octave separation with A bursts modulated at 200 Hz and B bursts at 50 Hz (AM200-50), and 2.59-octave separation with A bursts modulated at 300 Hz and B bursts at 50 Hz. The depth of AM was 100%.

Procedure

To account for perceived loudness differences in presentation of frequency-varied stimuli, each participant performed loudness balancing through an adaptive procedure (Jesteadt, 1980) to begin testing. Through the procedure, the levels of an A burst perceived to be equally loud as 60 dB SPL for a B burst were derived separately for the moderate and large spectral separations.

To measure stream segregation abilities based on listeners' behavioral responses, a single interval yes/no procedure was adopted. On each trial, either a delayed sequence or a no-delay sequence was presented. In a sequence, each B burst was presented at 60 dB SPL and each A burst at the level derived in the loudness balancing procedure. The task was to determine whether the sequence was delayed (i.e., the signal sequence) or no-delay (i.e., the reference sequence). Two graphic boxes on a computer screen, one showing "1 Longer" and one showing "2 Shorter," respectively, for the delayed and no-delay sequences. The participants responded by pressing number 1 or 2 on the RTbox (Li et al., 2010) to indicate their identification of delayed or no-delay sequence. Feedback was provided following each response by illuminating the box corresponding to the correct answer on the screen. Participants were allowed to take as much time as they needed to make the selection for each trial.

Two blocks of 65 trials were run for each condition with a 50% chance of signal sequences. The first 5 trials served to familiarize participants with the task. The last 60 trials were used to compute the hit rate and false alarm rate from both of which participants' sensitivity d' to the signal sequence was derived from Equation 1, yielding two d' scores per experimental condition.

$$d' = Z(h) - Z(f) \tag{1}$$

where $Z(h)$ and $Z(f)$, respectively, represent the Z transforms of hit rate and false alarm rate. Our previous studies (Nie et al., 2014; Nie and Nelson, 2015) have shown that this stimulus paradigm encouraged participants to segregate the A and B subsequences to achieve higher d' values. In those studies, the baseline performance for the rhythm-based stream segregation, as described at the end of the Procedure, was estimated with a d' value of 1.5 on average with stimulus sequences of 12 pairs of broadband noise bursts.[1] Before the experimental sessions, each participant completed a number of 40-trial training blocks that reflected the task of all the experimental conditions. To proceed to the experimental sessions, all participants were required to achieve a d' score of 1.5 or higher in at least one training block for the condition of large spectral separation without AM-rate separation, which suggested their ability to perform the stream segregation task.

A total of 18 experimental conditions were examined, with two sequence durations (9-pair and 3-pair), three levels of A–B spectral separation (no-separation, moderate separation, and large separation), and three AM-rate separations (AM0-0, AM200-50, and AM300-50). The participants undertook these conditions in

1 The rhythm-based baseline was anticipated to be lower than 1.5 due to the shorter sequence of 9 pairs AB bursts versus the 12 pairs in our previous reports.

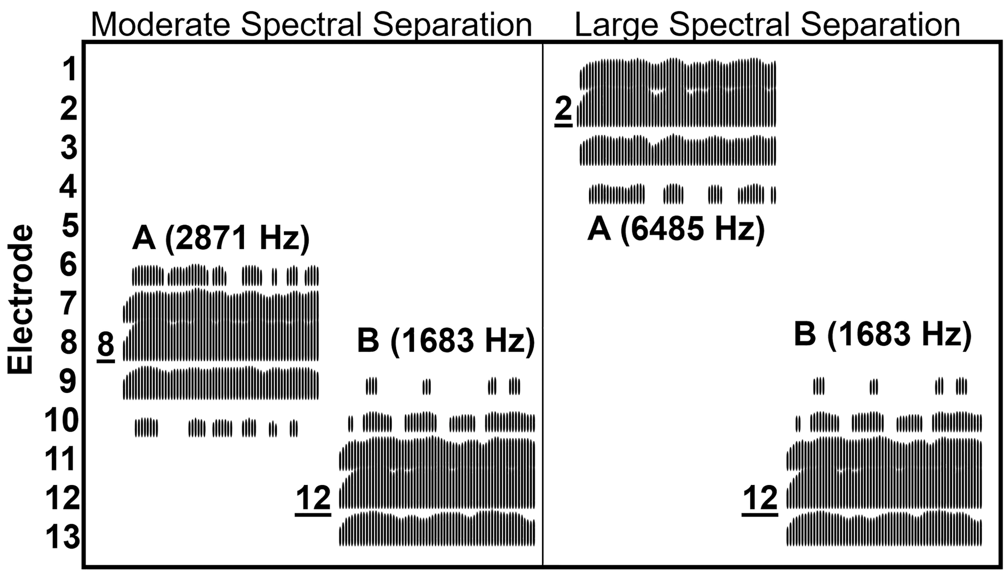

FIGURE 2
Electrodogram illustrating the electrodes activated by A and B narrowband bursts when they were at the moderate and large spectral separations for the three CI users, for whom the frequencies of stimulus bursts were identical. The electrodogram was generated based on the ACE processing strategy all three CI users used with a default frequency allocation (188–7,938Hz), stimulation rate (900Hz), and eight maxima. The A and B bursts are labeled with their corresponding center frequencies in the parentheses. The electrodes are ordered vertically starting from the most basal electrode #1 at the top through electrode #13 at the bottom. Direct labeling with underlined numbers indicate those electrodes to which the A and B bursts were assigned. Note that an A or B burst activated four or five electrodes and did not produce any activation on electrodes #5 and #14–20. The electrodes #14–20 are not shown in the figure.

a pseudorandomized order, such that duration/spectral separation conditions were randomized first, followed by the random order of the AM-rate separations nested under the duration/frequency separation conditions. The two repetitions of the same condition were conducted in two consecutively blocks.

It is noteworthy that the rhythm embedded in the stimulus sequences has been shown to enable voluntary stream segregation (Devergie et al., 2010; Nie et al., 2014). Thus, the stimulus sequences with no inter-stream spectral separation (i.e., the no-separation condition) and no AM-rate separation (i.e., the AM0-0 condition) effectively served as the control condition that provided the baseline performance in the absence of both spectral and AM-rate separations between the A and B streams.

Data analysis

The statistical package of IBM SPSS for Windows (Version 27.0) was used for data analysis. With either the Shapiro–Wilk test or graphical inspection of histograms, the d' scores were found to have violated the assumption of normal distribution overall or for most of the groupings. Thus, a complex Linear Mixed Effects (LME) model was fitted to the d' scores of both listener groups. The d' data of each listener group were also fitted with an LME model separately to examine the effects of spectral separation separately for each group. The residuals of these LME models were found normally distributed for most of the groupings. For readability, variables fitted in the LME models are specified in the Results section. When pairwise comparisons were performed, the reported p values have been corrected with the Bonferroni approach to control for the familywise error rate.

Results

Figure 3 depicts the d' values for both CI users and NH listeners in different experimental conditions. The d' values were higher for the 9-pair than for the 3-pair sequences and highest for the large A–B spectral separation condition out of the three separations. The effectiveness of the listeners training was investigated by comparing the full model with a simpler one. This simpler model excluded the *Repetition* effect but was otherwise identical to the full model with random effects of individual participants and their intercepts.[2] The simpler model is referred to as the complex model and assessed the fixed-effects factors of *Listener Group (or Group), Sequence Duration (or Duration), Spectral Separation, AM-rate Separation (or AM Separation)*, and all their two-way interactions, two three-way interactions (i.e., *Group* X *Spectral Separation* X *Duration* and *Group* X *AM Separation* X *Duration*).

With respect to the main fixed effects, three out of four were found to be significant, *Spectral Separation* [*F*(2,

2 The effect of *Repetition* was not significant in the full model, indicating limited learning in the experimental session.

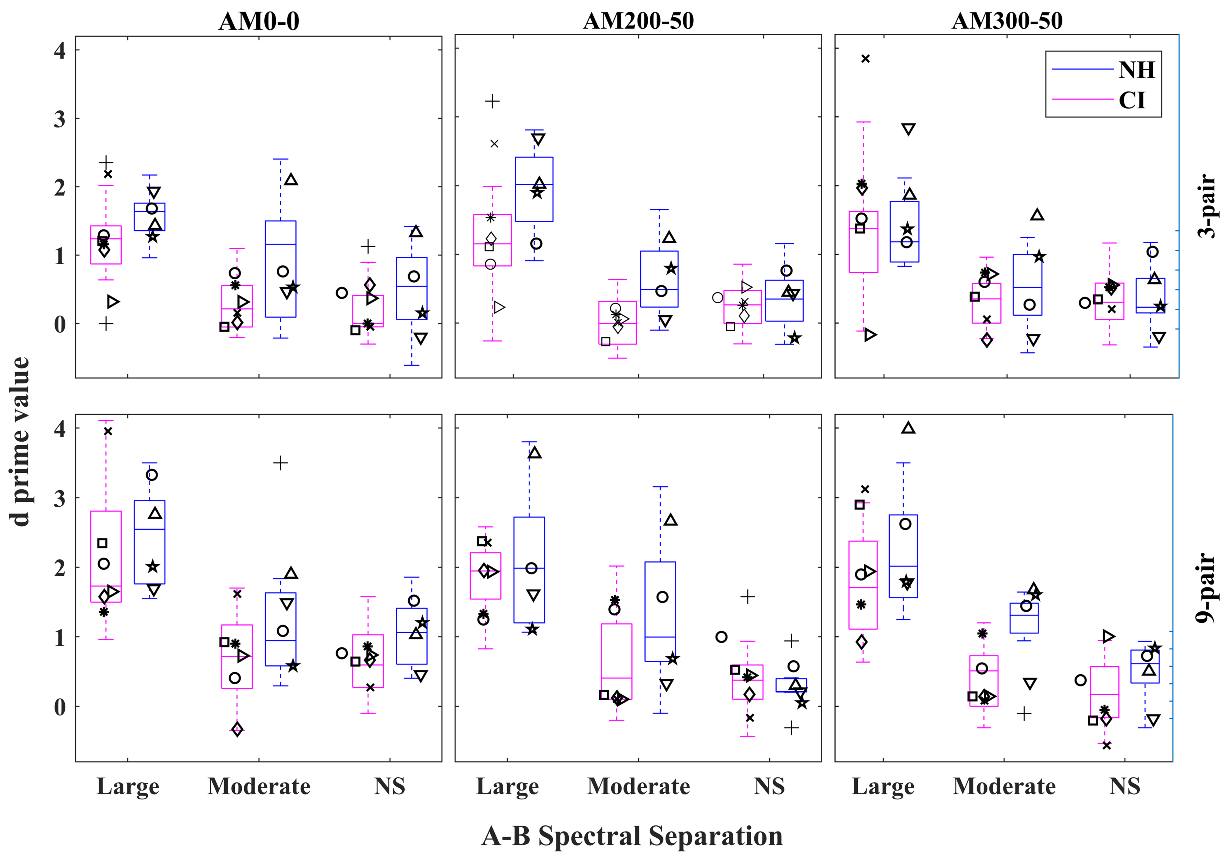

FIGURE 3
Boxplots illustrating d' values for each listener group in different experimental conditions. The boxes show the 25–75th percentile, the error bars show the 5th and 95th percentile, the plus signs show outliers, the solid horizontal line shows median performance. The rows show different sequence durations while the columns show the condition of AM-rate separation. NS=No-Separation. Each participant's d' mean of the two repetitions in each condition is also depicted with a unique symbol for the same listener across all conditions.

328) = 165.508, $p < 0.001$], *AM-rate Separation* [F(2, 328) = 4.197, $p = 0.016$], and *Duration* [F(1, 328) = 45.167, $p < 0.001$]. *Post-hoc* pairwise comparisons revealed that d' values progressively increased with the spectral separation, increasing by 0.254 ($p = 0.004$) from no-separation to the moderate separation and 1.079 ($p < 0.001$) from the moderate separation to large separation, as illustrated in Figure 4. The main effects of *AM Separation* and *Duration* are shown in Figure 5: The d' score was significantly poorer ($p = 0.018$) for the largest AM separation (AM300-50) than for the condition with no AM separation (AM0-0) by 0.216; no other pairwise comparisons among the AM separations were found to be significant. In addition, the d' mean difference of 0.427 was consistent with higher stream segregation ability with the 9-pair sequences than the 3-pair sequences. The *Group* effect was not significant [F(8, 328) = 4.554, $p = 0.065$].

Results revealed that the interaction of *Group* and *Spectral Separation* was significant, F(2, 328) = 3.611, $p = 0.028$. The data of each listener group were fitted with a separate LME model with the fixed-effects and random-effects terms same as those in the full model excluding any term associated with *Group*. As depicted in Figure 4, the d' score was found to progressively increase with the spectral separation for NH listeners ($p < 0.001$ for any d' increase). While the CI users showed the highest d' for the large spectral separation ($p < 0.001$), their d' score did not differ between the moderate- and no-separations ($p > 0.999$).

Results also revealed a significant interaction in the complex model between *Spectral Separation* and *Duration* [F(2, 328) = 4.552, $p = 0.011$] for both groups. The data of each duration were fitted with a separate LME model with the fixed-effects and random-effects terms same as those in the most complex model excluding any term associated with *Duration*. As depicted in Figure 6, the d' value was found to progressively increase with the spectral separation for the 9-pair sequences ($p < 0.003$ for any d' increase). With the 3-pair sequences, the d' score was highest for the large spectral separation ($p < 0.001$), but did not differ between the moderate- and no- separations ($p = 0.712$).

No significance was found for the interactions of *Duration* X *Group, Duration* X *AM Separation, Group* X *AM Separation, Group* X *Spectral Separation* X *Duration,* and *Group* X *AM Separation* X *Duration.*

Discussion

With a segregation-facilitating objective paradigm, the current study compared CI users and NH listeners in their ability to voluntarily segregate streams of NBN bursts based on the inter-stream spectral separation or AM-rate separation. The build-up of stream segregation was also investigated and compared between the two groups. The results suggest that CI users are less able than NH listeners to segregate NBN bursts into different auditory streams when they are moderately separated in the spectral domain. Contrary to our hypothesis, results indicated that the AM-rate separation interfered with the ability to segregate NBN sequences for both listener groups. Additionally, our results add evidence to the literature that CI users build up stream segregation at a rate comparable to NH listeners within around 2–3 s of stimulus onset.

Stream segregation based on spectral separation and AM-rate separation between groups

Consistent with the literature (e.g., Nie and Nelson, 2015; Tejani et al., 2017), the spectral separation of noise bands was shown to be a cue for stream segregation in both NH and CI groups as indicated by the increased d' values with the increase of A–B spectral separation. The progressively improved d' scores in NH listeners indicate their ability to segregate the A and B subsequences into different streams within 2.3 s when the two streams were at least 0.75 octaves apart. Recall that, the amount of moderate A–B spectral separation was 0.75 octaves for the NH listeners. In contrast, the CI users were unable to segregate A and B subsequences that were moderately separated in spectrum (with the separation of 0.70–0.91 octaves) in 2.3 s, even with the facilitation of focused voluntary attentional effort. Although, when the noise bands are largely separated in spectrum by 1.83–2.17 octaves, the CI participants are clearly able to voluntarily segregate noise streams.

Böckmann-Barthel et al. (2014) used acoustic sequences of harmonic tone complexes to investigate the effect of f0 differences across tone-complex sequences on stream segregation. With a subjective paradigm where CI listeners reported the number of streams perceived, the authors found prevalent stream segregation at the 10-semitone (i.e., 0.83 octaves) f0 difference. This differs from our result, which indicates that CI listeners were unable to segregate streams of NBN separated by 0.70–0.91 octaves. The discrepancy may in part be due to the slower stimulus rate in the current study than in Böckmann-Barthel et al.: In the current study, the stimulus rate was 3.8 AB pairs per second, whereas Böckmann-Barthel

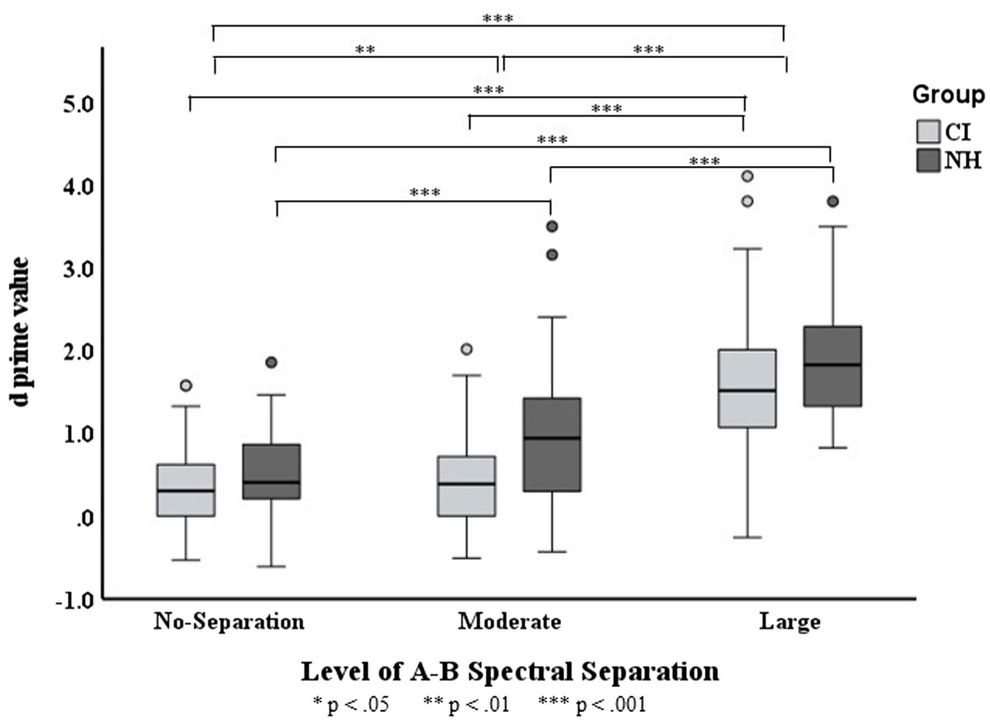

FIGURE 4
Boxplots for d' values across three A–B spectral separations in the CI and NH groups. The boxes show the 25–75th percentile, the error bars show the 5 and 95th percentile, the circle symbols show outliers, the solid horizontal line shows median performance. The statistically significant differences reported in the Results is indicated by asterisks.

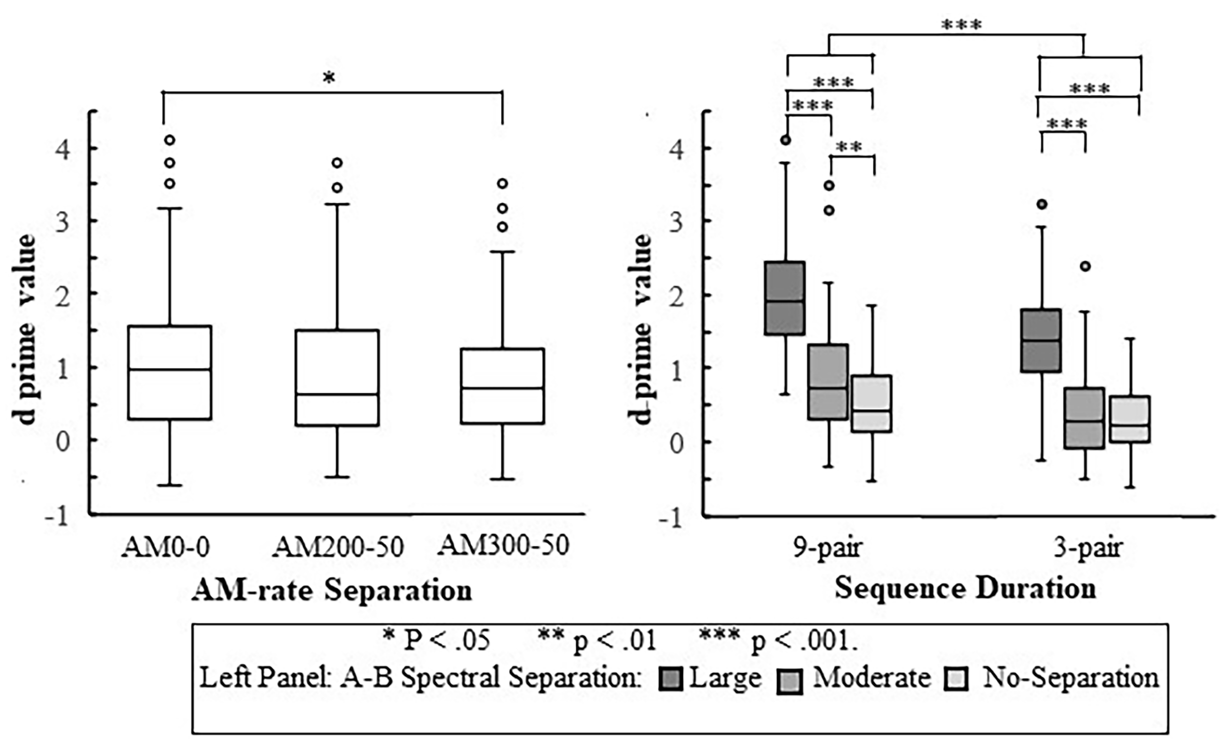

FIGURE 5
Boxplots for d' values in the three AM-rate separations (left panel) and in the two sequence durations at different A–B spectral separations (right panel). The boxes show the 25–75th percentile, the error bars show the 5 and 95th percentile, the circle symbols show outliers, the solid horizontal line shows median performance. The statistically significant differences reported in the Results are indicated by asterisks.

et al. repeatedly presented harmonic tone complexes in the ABAB group at a rate of 6 Hz, resulting a stimulus rate of 12 AB pairs per second. The slower rate in the current study may have led to weaker stream segregation (van Noorden, 1975). The discrepancy may also be attributed to the longer observational time for participants in Böckmann-Barthel et al. than in the current study. That is, almost all CI users provided their first responses 3 s after the stimulus onset, whereas the observational sequence duration was 2.325 s or shorter for the listeners in the current study. This potential effect of observational time is relevant to the build-up of stream segregation. Future studies using a larger range of varying sequence duration will allow studying the build-up stream segregation based on the cue of moderate spectral separation in CI users.

Note that, while the A and B bursts each were manipulated to be assigned to a single electrode according to a CI user's frequency allocation, the electrodogram shows that effectively, each acoustic burst activated a group of 4–5 electrodes. This internal spread of activation substantially reduced the A–B spatial separation in the cochlea. As a result, for the moderate spectral separation, contrasting the 4- or 5-electrode separation between A and B bursts according to the frequency allocation, the electrode groups activated by the two bursts are overlapping; for the large spectral separation, the two bursts activated two electrode groups that are four electrodes apart, instead of two single electrodes separated by 9–10 electrodes (for the MED-EL device, it was 5 electrodes apart).

As alluded in Introduction, using acoustic stimuli does not allow precisely controlling the electrode separation of the streams. In contrast, the use of direct electrical stimulation can constrain each stream to discretely activate a single electrode, in turn providing more precise control on the electrode separation. With direct electrical stimulation, (Paredes-Gallardo et al., 2018b) revealed that, on average, a minimum of 2.8-electrode separation is required for CI users to segregate two auditory streams. The patterns of electrode activation by the acoustic stimuli in our study show a similar trend in that CI users have limited ability to form perceptual streams when the internal activations by acoustic streams are not distinctly separated in the cochlea. It is clear that a distinct four-electrode separation is adequate to allow voluntary segregation. Additionally, using acoustic stimuli allows us to relate our finding to the real-world, which highlights that CI listeners experience sequential interference even when acoustic targets and distractors are apart by more than half octaves.

As the significant interaction of *Group* with *Spectral Separation* reveals, NH listeners are better able to use spectral separation as a cue for the formation of auditory streams. This interaction also indicates that the performance differences between NH listeners and CI users are not comparable across the three spectral separations. There were smaller group differences at the two extreme A–B spectral separations—no-separation and large separation than at the moderate segregation. This suggests that the focused attention may help CI users segregate the auditory streams

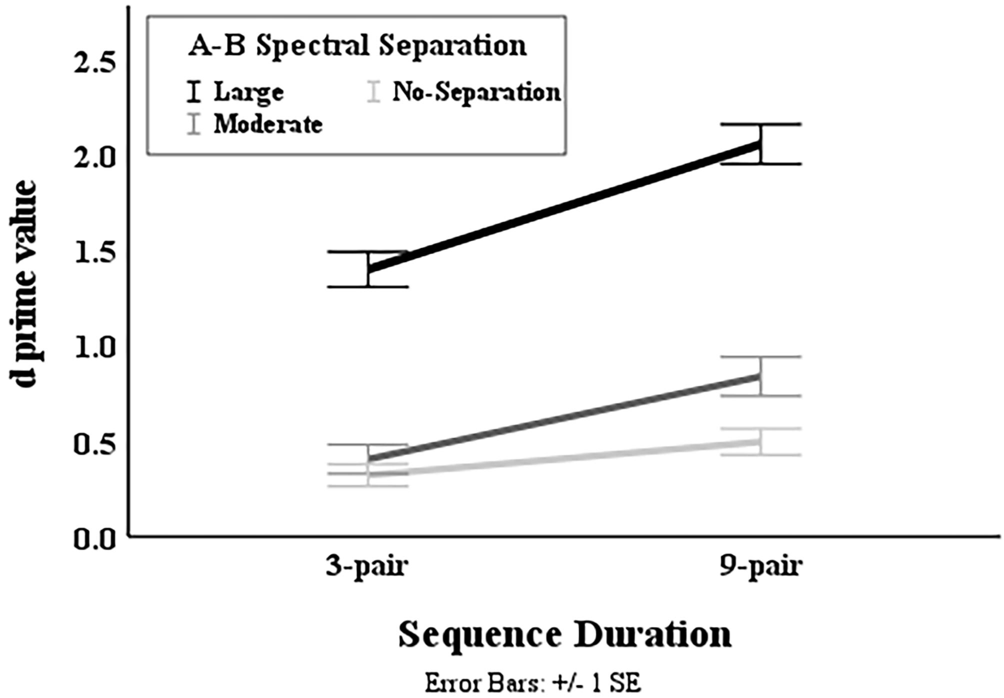

FIGURE 6
Illustration of mean d' value in the two sequence durations at different A–B spectral separations.

when the physical cues, such as spectral difference in this study, become more salient. Recent electrooculography (EEG) studies have reported the facilitation of focused attention for the processing of both non-speech and speech signals in competing maskers for CI users. For example, presenting CI users with melody-like interleaved target and distractor streams through direct stimulations, Paredes-Gallardo et al. (2018a) studied the N1 wave—the neurophysiological marker for the initial sensory registration of auditory stimuli—separately for the target and distractor streams in two attentional conditions—the attentive condition, wherein the CI users selectively focused attention on the target stream to perform a listening task, and the ignore condition, wherein the CI users' attention was focused on a silent video instead of the auditory stimulation. The study revealed that N1 was enhanced in the attentive condition compared to the ignore condition for the target (i.e., attended) stream but remained comparable between the two attentional conditions for the distractor (unattended) stream, demonstrating the facilitative effect of focused attention on the response to non-speech streams. Nogueira and Dolhopiatenko (2022) used EEG to decode attention to two concurrent speech streams—target (i.e., attended) stream and competing (i.e., unattended) stream and found that the EEG-indexed attentional difference between the attended and unattended streams positively correlated with the CI users' perception of the target speech stream, demonstrating the facilitative effect of focused attention on speech perception. Our results suggest that NH listeners may require less focused attention in separating auditory streams, even when spectral separations are small as compared to CI users, who may require both larger spectral separations and focused attention.

When comparing the CI listeners to the NH listeners, they showed overall comparable d' scores. This is likely due to the small sample size which has lowered the statistical power. To examine this proposition, we performed a power analysis which, at the α level of 0.05 and the power of 80% (see the Limitations section for details), estimated the sample size to double in both groups to reveal a significantly higher d' for the NH group than for the CI group. Thus, additional participants would be required to further examine the hypothesized lower overall ability to segregate NBN streams for CI users' than NH listeners.

In contrast to previous results (Nie and Nelson, 2015; Paredes-Gallardo et al., 2018c) that temporal-pitch separations (generated by differences in AM-rate or pulse rate) aid stream segregation, the current study showed that the large AM separation (i.e., AM300-50) significantly interfered with performance with no AM separation (AM0-0). That is, as AM-rate differences increased, performance decreased. In other words, increased AM-rate differences caused stream segregation abilities to decrease. Considering the process of a listener performing the task, this result suggests that the A stream interferes more strongly with participants' ability to follow the steady B stream when the A and B bursts were amplitude modulated at 300 and 50 Hz, respectively, than when the two bursts were both unmodulated. In Nie and Nelson (2015), wideband noise carriers were used, whereas the current study used NBN carriers. In line with the lower sensitivity in detecting AM with noise carriers of narrower bandwidth than those with wider bandwidth (e.g., Viemeister, 1979; Eddins, 1993), Lemańska et al. (2002) reported a significantly worse AM-rate discrimination with NBN carriers than with wideband noise carriers, likely due to the interference of larger intrinsic amplitude fluctuations of the NBN on the amplitude changes resulting from the amplitude modulation. As a result, the A-to-B perceptual difference elicited by the reduced sensitivity to AM may be insufficient for stream segregation. On the contrary, the AM-elicited A-to-B perceptual difference increased the interfering effect of A stream on the B stream. This may be attributed to the other stimulus difference between the current study and previous works (Nie and Nelson, 2015; Paredes-Gallardo et al., 2018c) in that the slower AM-rate (i.e., 50 Hz) was presented in the attended stream (i.e., B stream) in the current study. In contrast, the faster rate (e.g., 300 Hz for AM-rate or 300 pulses per second) was in the attended stream in the other two works. The sensitivity to AM has been evidenced to depend on the observational intervals (e.g., Viemeister, 1979; Lee and Bacon, 1997) such that, based on the multiple-looks theory (Viemeister and Wakefield, 1991), as the number of "looks" (commonly regarded as equivalent to the cycles) of AM increase, the sensitivity increases. With the 80-ms duration for each stimulus burst in the current study, the number of "looks" for 50-Hz AM was four which was substantially less than the 24 "looks" for the 300-Hz AM, which may affect the equivalence of perceptual salience between the amplitude modulated A and B streams. Future studies should confirm or equate the perceptual salience elicited by different AM rates or pulse rates to examine the effect of temporal pitch on stream segregation.

Build-up stream segregation in cochlear implant users compared to normal-hearing listeners

Both CI users and NH listeners showed evidence of build-up stream segregation as performance improved with the 9-pair condition relative to the 3-pair condition. Nie and Nelson (2015) used 12-pair sequences to elicit build-up, whereas the current study used 9-pair sequences. This study has shown that even 9-pair sequences are adequately long to elicit build-up stream segregation when compared to 3-pair sequences. That is, over a course of approximately 2.3 s, both NH listeners and CI users increased their performance in segregating the A and B streams with the facilitation of voluntary attention. These results also address the question raised in Böckmann-Barthel et al. (2014) on the build-up stream segregation of CI users within approximately 3 s of the stimulus onset. Recall, in that study where listeners were asked to report the number of streams perceived throughout the course an acoustic sequence of 30 s, the majority of the CI users did not provide their first response until 3 s post the onset of a sequence, leaving the build-up effect uncertain in the short post-onset period. Here, results show that the CI users in our study made use of the short duration of 2.3 s to build up stronger stream segregation for the levels of inter-stream differences specific to this study. These results also suggest that the current stimulus paradigm may be utilized to objectively study CI users' perception in the earlier period of a listening course in which subjective responses are not readily provided by the listeners.

Our results did not show an interaction between listener group and the duration of the sequence. While this indicates that both participant groups build up stream segregation at comparable rates, the smaller sample size in each group could also be a potential factor. The statistical power of 8.50% estimated in the Limitations section is markedly small, suggesting the likelihood of the non-significant interaction arising from the small sample size is low.

It should note that, the build-up effect revealed in the NH group, at the inter-stream spectral separation of 0.75 octaves or greater in the current study, is not consistent with the Deike et al. (2012) study in which listeners were asked to report the number of streams perceived throughout a course when listening to a sequence composed of two alternating harmonic tone complexes. At the inter-stream f0 separations of 8 semitones (i.e., 0.67 octave) or greater, the likelihood for NH listeners in the Deike et al. study to perceive two separate auditory streams started at the highest level for the first responses (within 1 s of the sequence onset) and remained constant over the entire course of the sequence. In other words, NH listeners did not require time to build up stream segregation in these f0 separations in Deike et al. The faster stimulus rate in the Deike et al., which was 24 AB pairs--the same as in Böckmann-Barthel et al. (2014), may have resulted in stronger segregation which require limited build-up.

Limitations

The main limitation of this study is the small sample size, which raised a question about the adequacy of power for the statistical analyses. To assess this limitation, using an R package SIMR (Green et al., 2016; Green and MacLeod, 2016), we conducted post-hoc power analyses.[3] The power analyses started with fitting the d' values in the R package LME4 (Bates

3 *A priori* sample size had been estimated using G*Power (Faul et al., 2009) based on the repeated measures analysis of variance instead of the LME model that was used in the current study. For a medium effect size in Cohen's f of 0.25, the sample size was 8 in total for the within factors, 2 groups (CI and NH), 36 measures (18 experimental conditions×2 repetitions), alpha of 0.05, and power of 80%. With respect to the between factors, 28 total participants were predicted based on the same aforementioned statistical parameters, except a large effect size in Cohen's f of 0.4 was adopted here. The use of a large effect size for the between factors was assumed based on the findings of substantially worse discrimination of acoustic contours based on frequency differences, such as melodic contour discrimination or prosodic discrimination.

et al., 2015) through the comprehensive LME model with the random and fixed effects as reported in Results. Note that the two analysis software programs—R and SPSS revealed comparable statistics. The power of a given main fixed effect was estimated in SIMR through the function "powerSim" based on 1,000 simulations applying the "anova" test method. At the α level of 0.05, the estimated power values were greater than 80% for the effects of *Duration, Spectral Separation,* and their interaction (*Spectral Separation X Duration*), 75.8% for *AM-rate Separation*, and 67.30% for the interaction of *Group* X *Spectral Separation*. The non-significant difference between the NH and CI groups was against our hypothesis. The estimated power was 46.90%, suggesting inadequate sample size. Thus, the sample size required for both the CI and NH groups was estimated based on a minimum of the conventionally desired power of 80% (Cohen, 1977, Chapter 2). Following the procedure described by Green and MacLeod, this sample size was estimated to be 12 for each group with a total of 24 participants. To assess the potential contribution of the small sample size to the non-significant *Group X Duration* interaction, the statistical power was also estimated to be 8.50%, suggesting the likelihood of the non-significant interaction arising from the small sample size is low. The above power analysis outcomes suggest that most of the significant effects were fair, but non-significant effect of listener group is likely due to the small sample size.

Additionally, it should note that a participant may perform the task based on an alternative mechanism not involving stream segregation; that is, based solely on detecting the gap between the last A and B bursts, instead of following throughout the entire course of a sequence. The d' value in the 3-pair condition with AM0-0 and no A–B spectral separation may approximate the sensitivity to the signal sequences based on this mechanism of gap discrimination. In this condition, the listeners' rhythm-based stream segregation was limited, if not none, as a result of markedly low observational intervals with three pairs of bursts. With the identical A and B bursts, stream segregation based on the dissimilarity between A and B bursts was not possible. Thus, the participants performed the task primarily by discriminating the A–B gaps and the d' (mean = 0.32, SD = 0.54) in this condition can be considered the baseline sensitivity through this mechanism. Participants achieved d' values substantially higher than this baseline, in conditions with robust cues for stream segregation, such as the 9-pair conditions (mean = 1.73, SD = 0.83) and the large spectral separations (mean = 1.13, SD = 0.98), which suggests that the mechanism of gap discrimination contributed to the task performance to a modest extent.

Summary

In summary, NH listeners were able to separate two NBN streams when their spectral separation was moderate or large within the given conditions. In contrast, CI users appeared only to be able to segregate these streams when their spectral separation was large. Additionally, the significant effect of sequence duration in both groups indicates listeners made more improvement with lengthening the duration of stimulus sequences, supporting the build-up effect within the course of approximately 2.3 s. The results of this study suggest that CI users are less able than NH listeners to segregate NBN bursts into different auditory streams when they are moderately separated in the spectral domain. Contrary to our hypothesis, our results indicate that AM-rate separation somewhat may interfere segregation of streams of NBN. Additionally, our results extend previous findings that cochlear implant users do show evidence for the build-up of stream segregation, which appeared to be comparable to NH listeners.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors upon reasonable request, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board at James Madison University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YN conceptualized the study. YN, AM, and HW designed the study. AM and HW collected the data. AM and YN conducted the data analysis. AM, YN, and HW contributed to the writing of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported, in part, by the CHBS Research Grant, CSD Faculty Fund and Ruth Memorial Student Research Grant from James Madison University.

Acknowledgments

We thank all the participants for their participation in this study. We would also like to thank Chris Clinard and Lincoln Gray for their helpful comments and Qian-Jie Fu for generating the electrogram.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Anstis, S., and Saida, S. (1985). Adaptation to auditory streaming of frequency-modulated tones. *J. Exp. Psychol. Hum. Percept. Perform.* 11, 257–271. doi: 10.1037/0096-1523.11.3.257

Assmann, P. F., and Summerfield, Q. (1987). Perceptual segregation of concurrent vowels. *J. Acoust. Soc. Am.* 82:S120. doi: 10.1121/1.2024632

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Böckmann-Barthel, M., Deike, S., Brechmann, A., Ziese, M., and Verhey, J. L. (2014). Time course of auditory streaming: do CI users differ from normal-hearing listeners? *Front. Psychol.* 5:775. doi: 10.3389/fpsyg.2014.00775

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.

Bregman, A. S., and Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *J. Exp. Psychol.* 89, 244–249. doi: 10.1037/h0031163

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. ProQuest Ebook Central. Available at: https://ebookcentral.proquest.com

Cooper, H. R., and Roberts, B. (2007). Auditory stream segregation of tone sequences in cochlear implant listeners. *Hear. Res.* 225, 11–24. doi: 10.1016/j.heares.2006.11.010

Cooper, H. R., and Roberts, B. (2009). Auditory stream segregation in cochlear implant listeners: measures based on temporal discrimination and interleaved melody recognition. *J. Acoust. Soc. Am.* 126, 1975–1987. doi: 10.1121/1.3203210

Cullington, H. E., and Zeng, F. G. (2008). Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects. *J. Acoust. Soc. Am.* 123, 450–461. doi: 10.1121/1.2805617

Cusack, R., Decks, J., Aikman, G., and Carlyon, R. P. (2004). Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 643–656. doi: 10.1037/0096-1523.30.4.643

Dannenbring, G. L., and Bregman, A. S. (1976). Stream segregation and the illusion of overlap. *J. Exp. Psychol. Hum. Percept. Perform.* 2, 544–555. doi: 10.1037/0096-1523.2.4.544

de Cheveigné, A. (1997). Concurrent vowel identification. III. A neural model of harmonic interference cancellation. *J. Acoust. Soc. Am.* 101, 2857–2865. doi: 10.1121/1.419480

Deike, S., Heil, P., Böckmann-Barthel, M., and Brechmann, A. (2012). The build-up of auditory stream segregation: a different perspective. *Front. Psychol.* 3:461. doi: 10.3389/fpsyg.2012.00461

Devergie, A., Grimault, N., Tillmann, B., and Berthommier, F. (2010). Effect of rhythmic attention on the segregation of interleaved melodies. *J. Acoust. Soc. Am.* 128, EL1–EL7. doi: 10.1121/1.3436498

Eddins, D. A. (1993). Amplitude modulation detection of narrow-band noise: effects of absolute bandwidth and frequency region. *J. Acoust. Soc. Am.* 93, 470–479. doi: 10.1121/1.405627

Faul, F., Erdfelder, E., Buchner, A., and Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149

Fu, Q. J., and Nogaki, G. (2005). Noise susceptibility of cochlear implant users: the role of spectral resolution and smearing. *J. Assoc. Res. Otolaryngol.* 6, 19–27. doi: 10.1007/s10162-004-5024-3

Goehring, T., Archer-Boyd, A. W., Arenberg, J. G., and Carlyon, R. P. (2021). The effect of increased channel interaction on speech perception with cochlear implants. *Sci. Rep.* 11:10383. doi: 10.1038/s41598-021-89932-8

Green, P., Mac Leod, C. J., and Alday, P. (2016). *Package 'simr'*. Available at: https://cran.r-project.org/web/packages/simr/simr.pdf

Green, P., and MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods Ecol. Evol.* 7, 493–498. doi: 10.1111/2041-210X.12504

Grimault, N., Bacon, S. P., and Micheyl, C. (2002). Auditory stream segregation on the basis of amplitude-modulation rate. *J. Acoust. Soc. Am.* 111:1340. doi: 10.1121/1.1452740

Hong, R. S., and Turner, C. W. (2006). Pure-tone auditory stream segregation and speech perception in noise in cochlear implant recipients. *J. Acoust. Soc. Am.* 120, 360–374. doi: 10.1121/1.2204450

Jesteadt, W. (1980). An adaptive procedure for subjective judgments. *Percept. Psychophys.* 28, 85–88. doi: 10.3758/BF03204321

Lee, J., and Bacon, S. P. (1997). Amplitude modulation depth discrimination of a sinusoidal carrier: effect of stimulus duration. *J. Acoust. Soc. Am.* 101, 3688–3693. doi: 10.1121/1.418329

Lemańska, J., Sęk, A., and Skrodzka, E. (2002). Discrimination of the amplitude modulation rate. *Arch. Acoust.* 27, 3–22.

Li, X., Liang, Z., Kleiner, M., and Lu, Z.-L. (2010). RTbox: a device for highly accurate response time measurements. *Behav. Res. Methods* 42, 212–225. doi: 10.3758/BRM.42.1.212

Moore, B., and Gockel, H. (2002). Factors influencing sequential stream segregation. *Acta Acust. Acust.* 88, 320–333.

Moore, B. C. J., and Gockel, H. E. (2012). Properties of auditory stream formation. *Philos. Trans. R. Soc. B* 367, 919–931. doi: 10.1098/rstb.2011.0355

Nie, Y., and Nelson, P. B. (2015). Auditory stream segregation using amplitude modulated bandpass noise. *Front. Psychol.* 6:1151. doi: 10.3389/fpsyg.2015.01151

Nie, Y., Zhang, Y., and Nelson, P. B. (2014). Auditory stream segregation using bandpass noises: evidence from event-related potentials. *Front. Neurosci.* 8:277. doi: 10.3389/fnins.2014.00277

Nogueira, W., and Dolhopiatenko, H. (2022). Predicting speech intelligibility from a selective attention decoding paradigm in cochlear implant users. *J. Neural Eng.* 19:026037. doi: 10.1088/1741-2552/ac599f

Oxenham, A. J., and Kreft, H. A. (2014). Speech perception in tones and noise via cochlear implants reveals influence of spectral resolution on temporal processing. *Trends Hear.* 18:2331216514553783. doi: 10.1177/2331216514553783

Paredes-Gallardo, A., Innes-Brown, H., Madsen, S., Dau, T., and Marozeau, J. (2018a). Auditory stream segregation and selective attention for cochlear implant listeners: evidence from behavioral measures and event-related potentials. *Front. Neurosci.* 12:581. doi: 10.3389/fnins.2018.00581

Paredes-Gallardo, A., Madsen, S. M. K., Dau, T., and Marozeau, J. (2018b). The role of place cues in voluntary stream segregation for cochlear implant users. *Trends Hear.* 22:233121651775026. doi: 10.1177/2331216517750262

Paredes-Gallardo, A., Madsen, S. M. K., Dau, T., and Marozeau, J. (2018c). The role of temporal cues in voluntary stream segregation for Cochlear implant users. *Trends Hear.* 22:233121651877322. doi: 10.1177/2331216518773226

Pelli, D. G. (1997). The video toolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366

Tejani, V. D., Schvartz-Leyzac, K. C., and Chatterjee, M. (2017). Sequential stream segregation in normally-hearing and cochlear-implant listeners. *J. Acoust. Soc. Am.* 141, 50–64. doi: 10.1121/1.4973516 van Noorden, L. (1975). *Temporal Coherence in the Perception of Tone Sequences*. Unpublished Ph. D. dissertation, Eindhoven University of Technology, Eindhoven.

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *J. Acoust. Soc. Am.* 66, 1364–1380. doi: 10.1121/1.383531

Viemeister, N. F., and Wakefield, G. H. (1991). Temporal integration and multiple looks. *J. Acoust. Soc. Am.* 90, 858–865. doi: 10.1121/1.401953

Walker, G., Dillon, H., and Byme, D. (1984). Sound field audiometry: recommended stimuli and procedures. *Ear Hear.* 5, 13–21. doi: 10.1097/00003446-198401000-00005

Warren, R. M., and Obusek, C. J. (1972). Identification of temporal order within auditory sequences. *Percept. Psychophys.* 12, 86–90. doi: 10.3758/BF03212848

frontiers | Frontiers in Neuroscience

TYPE Original Research
PUBLISHED 27 September 2022
DOI 10.3389/fnins.2022.900879

Check for updates

OPEN ACCESS

EDITED BY
Jing Chen,
Peking University, China

REVIEWED BY
Tatiana Rocha Silva,
Federal University of Minas Gerais, Brazil
Magdalena Sosna-Duranowska,
Institute of Physiology and Pathology of Hearing (IFPS), Poland
Leo Ling,
University of Washington, United States

*CORRESPONDENCE
Lei Xu
sdphxl@126.com

SPECIALTY SECTION
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

RECEIVED 21 March 2022
ACCEPTED 29 August 2022
PUBLISHED 27 September 2022

CITATION
Wang R, Luo J, Chao X, Wang H, Fan Z and Xu L (2022) Minimally invasive surgical techniques in vestibular function preservation in patients with cochlear implants.
Front. Neurosci. 16:900879.
doi: 10.3389/fnins.2022.900879

Minimally invasive surgical techniques in vestibular function preservation in patients with cochlear implants

Ruijie Wang[1,2], Jianfen Luo[1,2], Xiuhua Chao[1,2], Haibo Wang[1], Zhaomin Fan[1] and Lei Xu[1,2]*

[1]Department of Otolaryngology-Head and Neck Surgery, Shandong Provincial ENT Hospital, Shandong University, Jinan, China, [2]Department of Auditory Implantation, Shandong Provincial ENT Hospital, Jinan, China

Background: Cochlear implantation (CI) is an effective and successful method of treating individuals with severe-to-profound sensorineural hearing loss (SNHL). Coupled with it's great clinical effectiveness, there is a risk of vestibular damage. With recent advances in surgical approach, modified electrode arrays and other surgical techniques, the potential of hearing preservation (HP) has emerged, in order to preserve the inner ear function. These techniques may also lead to less vestibular damage. However, a systematic study on this at different follow-ups after CI surgery has not been documented before.

Aims: To investigate changes of vestibular function systematically in recipients at short and long follow-ups after a minimally invasive CI surgery.

Methods: In this retrospective study, 72 patients (72 ears) with minimally invasive CI were recruited. All participants selected had bilateral SNHL and pre-operative residual hearing (RH) and underwent unilateral CI. They were treated to comprehensive care. All patients underwent vestibular function tests 5 days prior to CI. During the post-operative period, follow-up tests were performed at 1, 3, 6, 9, and 12 months. The contemporaneous results of caloric, cervical vestibular-evoked myogenic potential (cVEMP), ocular vestibular-evoked myogenic potential (oVEMP), and video head impulse (vHIT) tests were followed together longitudinally.

Results: On the implanted side, the percent fail rate of caloric test was significantly higher than that of vHIT at 1, 3, and 9 months post-operatively ($p < 0.05$); the percent fail rate of oVEMP was higher than vHIT of superior semicircular canal (SSC), posterior semicircular canal (PSC), or horizontal semicircular canal (HSC) at 1, 3, and 9 months ($p < 0.05$); at 3 and 9 months, the percent fail rate of cVEMP was higher than that of SSC and PSC ($p < 0.05$). There were no significant differences in the percent fail rates among all tests at 6 and 12 months post-CI ($p > 0.05$). The percent fail rates showed decreased trends in caloric ($p = 0.319$) and HSC tested by vHIT ($p = 0.328$) from 1–3 to 6–12 months post-operatively. There was no significant difference in cVEMP between 1–3 and 6–12 months ($p = 0.597$). No significant differences on percent fail rates of cVEMP and oVEMP between short- and long-terms post-CI were found in the same subjects ($p > 0.05$). Before surgery, the abnormal cVEMP and oVEMP response rates were both lower in patients with enlarged vestibular aqueduct (EVA) than patients with a normal cochlea ($p = 0.001$, 0.018, respectively).

Conclusion: The short- and long-term impacts on the vestibular function from minimally invasive CI surgery was explored. Most of the vestibular functions can be preserved with no damage discrepancy among the otolith and three semicircular canal functions at 12 months post-CI. Interestingly, a similar pattern of changes in vestibular function was found during the early and the later stages of recovery after surgery.

KEYWORDS

cochlear implantation, minimally invasive surgery, vestibular function, otolith, canal

Introduction

Cochlear implantation (CI) has been widely applied in individuals with severe-to-profound sensorineural hearing loss (SNHL). Although CI is an effective and safe procedure, there is risk of trauma to the vestibular sensor, causing vertigo, balance disorder, or complete deterioration (Ibrahim et al., 2017; Yong et al., 2019; Li and Gong, 2020; Wang et al., 2021; West et al., 2021). Possible reasons include injury during electrode insertion, loss of perilymph, labyrinthitis, endolymphatic hydrops, and electrical stimulation (Fina et al., 2003). However, little is known about the main factors influencing status of vestibular function after CI.

Nowadays, individuals with residual hearing (RH) are also candidates for CI. Optimization of the electrode array and surgical techniques has resulted in a more effective approach to cochlear function preservation. A meta-analysis based on hearing preservation (HP) in cochlear implant surgery showed that a combination of round window (RW) approach with implanting a straight electrode might result in HP (Snels et al., 2019). Glucocorticoids have also been used for their anti-inflammatory and anti-apoptotic properties (Douchement et al., 2015). A recent study on minimally invasive surgery suggested that minimizing intracochlear pressure (ICP) during electrode insertion was effective for HP (Ordonez et al., 2019). Patients with pre-operative RH implanted through these techniques can extensively preserve their cochlear function in the long term (Skarzynski et al., 2019; Sprinzl et al., 2020).

Recently, vestibular function preservation using minimally invasive CI surgery has been addressed. We hypothesized that these techniques could similarly preserve vestibular function because of the proximity of the cochlea and vestibule. The notion rests on the assumption that the primary and secondary effect of insertion trauma might influence peripheral vestibular receptors and cochlear function alike (Stuermer et al., 2019). Tsukada and Usami (2021) found that the risk of vestibular damage could be reduced through less traumatic surgical techniques, such as a RW approach and flexible electrodes. Sosna-Duranowska found that the RW approach in HP techniques was associated with vestibular function protection (Sosna-Duranowska et al., 2021). Other studies have demonstrated that vestibular function can be seriously damaged, even with a RW approach (Li and Gong, 2020; Boje Rasmussen et al., 2021). Nevertheless, systematic studies of the influence of minimally invasive surgery on vestibular function protection are rare.

In previous evaluation of vestibular function in patients with CI, there was an exhaustive analysis of the horizontal semicircular canal (HSC) by caloric stimulation evaluating the low frequency response, otolith function evaluated by cervical vestibular-evoked myogenic potential (cVEMP) and ocular vestibular-evoked myogenic potential (oVEMP), and three semicircular canals evaluated using the video head impulse test (vHIT) with a high frequency stimulation.

We sought to assess the changes in both canal and otolith functions in patients undergoing minimally invasive CI techniques at different follow-up times.

Materials and methods

Participants

Seventy-two patients (72 ears) with pre-operative low-frequency residual hearing (LFRH) and severe-to-profound SNHL who underwent minimally invasive CI surgery at our auditory implant department between June 2017 and November 2020 were included in this retrospective study. The inclusion criteria was at least one low-frequency pure tone threshold (125, 250, or 500 Hz) ≤85 dB HL before surgery. Patients with severe cochlear malformation, peripheral vestibular disease, auditory synaptopathy/neuropathy, cochlear fibrosis, previous otologic surgeries, and those at risk to show poor participation were excluded, except for those with enlarged vestibular aqueduct (EVA). Computed tomography (CT) of the temporal bone and magnetic resonance imaging (MRI) were performed pre-operatively. An EVA was defined as a vestibular aqueduct diameter of >1.5 mm at the midpoint between the posterior cranial fossa and inner ear vestibule (Valvassori and Clemis, 1978).

Of these patients, 23 were female and 49 were male, and the mean age at implantation was 20.35 ± 19.05 years (range, 3–67 years). Young patients comprised 46 participants <18 years (mean age at implantation: 8.17 ± 3.49 years, 3–17 years), and adults comprised 26 participants ≥18 years (mean age at implantation: 41.88 ± 15.91 years, 19–67 years). Pre-operative CT and MRI revealed bilateral EVA in 33 (45.83%) participants. A total of 33 and 39 recipients underwent implantation in the left and right ears, respectively. The Nucleus CI422, CI522, and CI24RECA electrodes were implanted in 25, 8, and 17 patients, respectively. A MedEL Flex 28 electrode was implanted in 18 patients. Four recipients underwent implantation with a Nurotron CS-10A electrode. The RW surgical procedure was applied to the Nucleus CI422/522, Med-EL FLEX 28, and Nurotron CS-10A electrodes in 55 (76.39%) patients. The extended RW approach was applied to the Nucleus CI24RECA electrode in 17 patients.

Patients underwent vestibular assessments through caloric, cVEMP, oVEMP, and vHIT tests 5 days prior to CI and again at 1, 3, 6, 9, and 12 months post-CI. However, some patients were lost to follow-ups because of the limitations encountered in clinical settings. The processors were all switched off during tests after implantation. Detailed demographic information on the study participants is presented in **Tables 1**, **2**.

Minimally invasive surgical techniques

All participants underwent surgery performed by a single surgeon. Full insertion of the electrode was achieved in all patients. In addition to the approach toward the insertion point during surgery and choice of electrode array, other protective measures included the following: (1) the ossicle chain was kept intact during surgery; (2) rotational speed was reduced to a minimum to avoid sound damage when grinding the bone of the RW niche or extending RW; (3) care was taken to avoid aspiration of perilymphatic fluid; (4) sodium hyaluronate was used before opening the RW membrane; (5) the electrode was inserted steadily, gently, and slowly with an insertion time >1 min; (6) after electrode insertion, a small piece of muscle was gently packed around the RW; and (7) systematic glucocorticoids were administered to all patients 1 day before surgery until 1 week after surgery.

Caloric test

The bithermal caloric test was performed. A video-based system was used (Ulmer VNG, v. 1.4; Synapsys, Marseille, France) to acquire and analyze the eye response. Each ear was irrigated alternatively with a constant flow of air at 24 and 49°C for 40 s. The response was recorded over 3 min. A 7-min interval between each stimulus was observed to avoid cumulative effects. We calculated the maximum slow-phase velocity (SPV) of nystagmus after each irrigation to determine unilateral weakness (UW) according to Jongkee's formula. In our laboratory, a value of UW less than 20% was judged normal.

Cervical vestibular-evoked myogenic potential

Cervical vestibular-evoked myogenic potential was recorded using the Neuro-Audio auditory evoked potential equipment (Neurosoft Ltd., Ivanovo, Russia). The test was performed with the patients in seated position. Tone burst stimuli (93–100 dB nHL, 500 Hz) were delivered via a standard headphone. Active recording electrodes with respect to the examination were placed on the region of the upper third of the sternocleidomastoid muscle (SCM) on both sides. The reference electrodes were placed on the upper sternum. The ground electrode was on the nasion. The head was rotated toward the contralateral side of the stimulated ear to achieve tonic contraction of the SCM during recording. The stimulation rate was 5.1 Hz. Bandpass filtering was 30–2000 Hz. An amplitude ratio over 30% was considered abnormal if the weaker response was from the implanted ear. In the event of bilaterally absent responses, the absent response was considered abnormal.

Ocular vestibular-evoked myogenic potential

Ocular vestibular-evoked myogenic potential was recorded using the Neuro-Audio auditory evoked potential equipment (Neurosoft Ltd., Ivanovo, Russia). The electromyographic activity of the extraocular muscle was recorded with the patients in the seated position. Tone burst stimuli (93–100 dB nHL, 500 Hz) were delivered via a standard headphone. The active recording electrodes were placed on the infra-orbital ridge 1 cm below the center of each lower eyelid. The reference electrodes were positioned approximately 1 cm below them. The ground electrode was on the nasion. The results were recorded with eyes open and maximal gaze upward. The stimulation rate

TABLE 1 Demographic information of all subjects who participated in this study.

S	Sex	Side	AAT (year)	Electrode	Imaging	Post-CI (month)	Pre-CI VF	LFRH pre-CI (dB HL)
S1	M	R	6	CI422	M, E	1, 6, 9, 12	−, +, +, +, +, +	N, 70, 95
S2	M	L	11	CS-10A	M, E	12	+, +, +, +, +, +	N, 75, 80
S3	F	R	5	CI422	M, E	1, 3, 6, 9	+, +, +, +, +, +	N, 75, 95
S4	M	R	13	CI422	Normal	1, 9	/, +, +, +, +, +	65, 75, 85
S5	M	L	13	CI422	Normal	1, 3, 6, 9, 12	+, +, +, +, +, +	10, 20, 75
S6	M	R	7	F28	Normal	9	+, +, +, +, +, +	N, 75, 95
S7	F	L	6	CI522	Normal	3, 6, 9, 12	−, +, +, +, +, +	N, 65, 80
S8	M	L	5	CI422	M, E	1, 6, 12	/, +, +, +, +, +	N, 80, 80
S9	F	R	17	CI422	M, E	1, 3, 9, 12	+, +, +, +, +, +	65, 75, 90
S10	M	R	7	CI422	M, E	1, 3, 6, 12	/, +, +, +, +, +	N, 55, 60
S11	M	R	14	CI422	M, E	1, 3, 9, 12	+, +, +, +, +, +	75, 85, 95
S12	M	R	8	CI422	M, E	6, 9, 12	−, +, +, +, +, +	70, 65, 75
S13	F	R	6	CI422	M, E	6	+, +, +, +, +, +	N, 75, 95
S14	F	R	10	F28	Normal	1, 3, 6, 9, 12	+, +, +, +, +, +	50, 65, 85
S15	M	L	19	F28	M, E	3, 9	+, +, +, +, +, +	65, 80, 85
S16	M	R	12	F28	Normal	1, 6, 9, 12	+, +, −, +, +, +	30, 45, 100
S17	F	L	6	F28	M, E	1, 6, 12	/, +, +, +, +, +	N, 65, 55
S18	M	R	7	CI422	Normal	3, 6, 9, 12	/, +, +, +, +, +	55, 75, 100
S19	M	R	3	CI422	Normal	1, 6, 12	/, −, +, +, +, +	N, 60, 70
S20	F	R	11	F28	M, E	3, 6, 9	+, +, +, +, +, +	60, 70, 80
S21	F	L	35	CI422	Normal	1, 3, 6, 9, 12	+, −, +, +, +, +	55, 50, 50
S22	M	L	62	F28	Normal	1, 3, 9, 12	−, −, −, +, +, +	40, 45, 60
S23	F	R	34	CI522	Normal	1, 3, 6, 9	+, +, +, +, +, +	20, 25, 45
S24	M	R	10	CI422	Normal	3, 6, 12	+, +, +, +, +, +	45, 65, 95
S25	M	L	6	F28	Normal	12	−, +, −, +, +, +	N, 80, 85
S26	M	R	7	CI422	M, E	1, 3, 9	+, +, +, +, +, +	N, 70, 90
S27	F	L	67	CI422	Normal	6, 12	−, −, −, +, +, +	55, 65, 70
S28	F	L	15	F28	M, E	9	−, +, −, +, +, +	60, 70, 75
S29	M	R	6	CI422	M, E	1, 12	/, +, +, +, +, +	55, 45, 70
S30	M	L	6	CI522	M, E	1, 12	/, +, +, +, +, +	60, 45, 65
S31	M	L	48	CS-10A	Normal	1, 3, 6	/, −, −, +, +, +	65, 70, 95
S32	M	L	5	CI422	M, E	3, 6	/, +, +, /, /, /	70, 70, 70
S33	M	L	6	F28	M, E	3	/, +, +, +, +, +	60, 65, 75
S34	M	L	7	CI522	M, E	1, 3	+, +, +, +, +, +	70, 70, 70
S35	M	L	41	CS-10A	Normal	1	+, +, +, +, +, +	50, 70, 100
S36	M	R	52	F28	Normal	1, 3	−, +, −, +, −, −	50, 70, 80
S37	F	L	67	CI422	Normal	1, 6, 12	−, −, −, +, +, +	55, 65, 70
S38	F	R	54	CI522	Normal	1, 3, 12	−, +, +, +, +, +	45, 45, 55
S39	M	L	5	CI422	M, E	1, 3, 9	+, +, +, +, +, +	65, 55, 55
S40	M	R	5	CI522	M, E	1, 3, 9	+, −, −, +, +, +	50, 50, 60
S41	F	R	53	CS-10A	Normal	1, 3, 6	−, −, −, +, −, +	40, 55, 75
S42	M	R	6	F28	M, E	1, 3, 6	+, +, +, +, +, +	80, 75, 90
S43	F	L	11	CI422	Normal	1, 3	/, +, +, +, +, +	60, 75, 90
S44	M	R	5	CI422	M, E	1, 6	−, +, +, +, +, +	N, 60, 85
S45	M	R	34	CI422	Normal	6, 12	+, +, −, +, +, +	N, 85, 90
S46	F	R	7	F28	Normal	1, 3	−, +, +, +, +, +	N, 85, 85
S47	F	L	9	F28	M, E	1, 6	+, +, +, +, +, +	55, 65, 95
S48	F	R	9	F28	M, E	1, 6	+, +, −, +, +, +	60, 70, 90
S49	M	L	11	F28	Normal	3, 9, 12	−, +, +, +, +, +	70, 80, 90
S50	M	L	29	CI422	Normal	1	−, −, +, +, +, +	60, 70, 85
S51	M	L	8	F28	E	1, 3, 9	−, −, −, +, +, −	85, 90, 105
S52	F	R	7	CI422	M, E	1, 3, 6	−, +, +, +, +, −	55, 60, 65
S53	M	L	20	CI522	Normal	1, 9	+, +, +, +, +, +	30, 40, 55
S54	M	R	20	CI522	Normal	1, 9	+, +, +, +, +, +	30, 45, 90
S55	M	R	15	F28	Normal	1, 3	+, +, +, +, +, +	75, 85, 95

(Continued)

TABLE 1 (Continued)

S	Sex	Side	AAT (year)	Electrode	Imaging	Post-CI (month)	Pre-CI VF	LFRH pre-CI (dB HL)
S56	M	R	6	CA	M, E	1, 3, 6, 9, 12	+, +, +, +, −, +	N, 65, 75
S57	M	L	8	CA	M, E	1, 3, 9, 12	−, +, +, +, +, +	85, 85, 90
S58	M	R	53	CA	Normal	12	−, −, −, −, −, −	N, 75, 80
S59	F	L	28	CA	Normal	1, 3, 6, 9	−, +, +, +, +, +	15, 45, 80
S60	M	L	19	CA	Normal	1, 3, 6, 9	+, −, /, +, +, +	85, 85, 90
S61	M	R	48	CA	Normal	1, 3	+, −, −, +, −, −	80, 80, 75
S62	F	R	50	CA	Normal	6, 9, 12	−, −, −, +, +, +	65, 75, 85
S63	F	R	6	CA	E	1, 9	+, +, +, +, +, +	N, 80, 90
S64	M	L	30	CA	Normal	1, 3, 9	+, +, +, +, +, +	N, 80, 100
S65	M	R	65	CA	Normal	3	−, −, −, +, +, +	75, 80, 80
S66	M	L	55	CA	Normal	1, 9	+, −, −, +, +, +	N, 80, 65
S67	M	R	6	CA	M, E	3, 9	/, +, +, /, /, /	N, 85, 105
S68	F	L	36	CA	Normal	1, 3	+, +, −, +, +, +	65, 70, 100
S69	M	L	4	CA	E	1, 9	+, +, +, +, +, +	N, 50, 80
S70	M	R	5	CA	M, E	3	−, +, +, +, +, +	N, 80, 95
S71	M	L	51	CA	Normal	1, 3	−, +, +, −, +, +	N, 70, 70
S72	M	R	19	CA	Normal	1, 6	+, −, −, +, +, +	65, 65, 80

AAI, age at implantation; CA, CI24RECA; F, female; M, male; L, left; R, right; M, Mondini; E, enlarged vestibular aqueduct; LFRH, low frequency residual hearing (125, 250, 500 Hz); N, not tested; pre-CI VF, vestibular function pre-operatively in the following order (caloric, cVEMP, oVEMP, SSC, HSC, PSC); /, no tested; +, normal response; −, absent or decreased response; HSC, horizontal semicircular canal; SSC, superior semicircular canal; PSC, posterior semicircular canal.

was 5.1 Hz. Bandpass filtering was 1–1000 Hz. An amplitude ratio over 30% was considered abnormal if the weaker response was from the implanted ear. In the event of bilaterally absent responses, the absent response was considered abnormal (Zhang et al., 2019).

Video head impulse test

The vHIT device (Ulmer II Evolution, France) was used. The patient was instructed to maintain eye fixation on a stationary object on a screen at about 1 m distance while examiner manipulated the patient's head with quick and precise head movements. The vestibulo-ocular reflex (VOR) gain was calculated by vHIT software based on head velocity and eye velocity curves. In a full test, 5–10 head thrusts were completed per canal for the recording. When the head was turned in the plane of the semicircular canal to be tested, the VOR maintained visual fixation. The breaking of visual fixation, revealed by a corrective saccade, indicated a respective canal disorder. This test was possible as soon as the child could hold his head steady. A VOR gain of the HSC less than 0.8 was considered to be abnormal. For both the superior semicircular canal (SSC) and posterior semicircular canal (PSC), a VOR gain less than 0.7 was considered to be abnormal (Sichnarek et al., 2019).

Statistical analyses

Data were analyzed using the Statistical Package for the Social Sciences (SPSS), version 23.0 (SPSS, Inc., Chicago, IL, USA). The Chi-square test was used to compare the percent fail rates. Statistical significance was set at $p < 0.05$.

Results

Comparison of percent fail rates among each vestibular end-organ senor on the implanted side at different time points

Of the 72 patients, 25 were evaluated with all four assessments before CI and 1 month after CI. Caloric responses in CI ears were normal in 18 cases pre-operatively and abnormal in 44.44% (8/18) of cases 1 month post-operatively. Similarly, the percent fail rates were 27.27% (6/22) for cVEMP, 47.62% (10/21) for oVEMP, 8.33% (2/24) for SSC, 16.00% (4/25) for vHIT of HSC, and 8.33% (2/24) for PSC at 1 month post-operatively. The

TABLE 2 Summary the number of patients tested at all the time points.

Tests	Number of patients tested among all 72 patients					
	Pre	1 Month	3 Months	6 Months	9 Months	12 Months
Caloric	59	28	27	20	21	20
cVEMP	72	43	36	28	30	24
oVEMP	71	43	36	27	29	24
SSC	70	45	40	32	32	28
HSC	70	45	40	32	32	28
PSC	70	45	40	32	32	28

10.3389/fnins.2022.900879

chi-square test showed that the percent fail rate of caloric was significantly higher than that of SSC, HSC, and PSC tested under vHIT (p = 0.019, 0.040, and 0.019, respectively). The rate of oVEMP was higher than that of SSC, HSC, and PSC (p = 0.003, 0.020, and 0.003, respectively).

At 3 months post-operatively, the percent fail rate was significantly higher in caloric than in SSC, HSC, and PSC (p = 0.002, 0.036, and 0.009, respectively); the rate of oVEMP was higher than that of SSC and PSC (p = 0.011, 0.039); and the rate of cVEMP was higher than that of SSC and PSC (p = 0.015, 0.049).

At 9 months post-operatively, the percent fail rate was significantly higher in caloric than in SSC, HSC, and PSC (p = 0.002, 0.016, and 0.004, respectively); the rate of oVEMP was higher than that of SSC, HSC, and PSC (p = 0.004, 0.029, and 0.007, respectively); and the rate of cVEMP was higher than that of SSC and PSC (p = 0.020, 0.032).

No statistically significant differences were observed in the percent fail rates among all vestibular function tests at 6 and 12 months after implantation ($p > 0.05$).

In this part, a different set of patients contributed at each time point for each test. All the post-operative vestibular damages were new damages. cVEMP and oVEMP results are shown in **Figure 1**. The percent fail rates of all the four tests on the implanted side at each time point are listed in **Table 3** and **Figure 2**.

Changes in percent fail rate in each vestibular end-organ function on the implanted side at short (1–3 months) and long (6–12 months) follow-up times after surgery

To study the same set of patients longitudinally we divided the follow-up time points into two groups: early (1–3 months) and late (6–12 months). Nineteen patients in caloric, 35 in cVEMP, 35 in oVEMP, and 38 in vHIT were evaluated before CI, 1–3 months and 6–12 months post-CI. All of these patients had pre-operative normal vestibular functions. In this part, for each test the same group of patients participated at all three time points.

For a patient who had multiple assessments at different evaluation time points during his follow-up, the first assessment from 1 to 3 months was chosen as the short-term result and the latest assessment from 6 to 12 months was chosen as the long-term result.

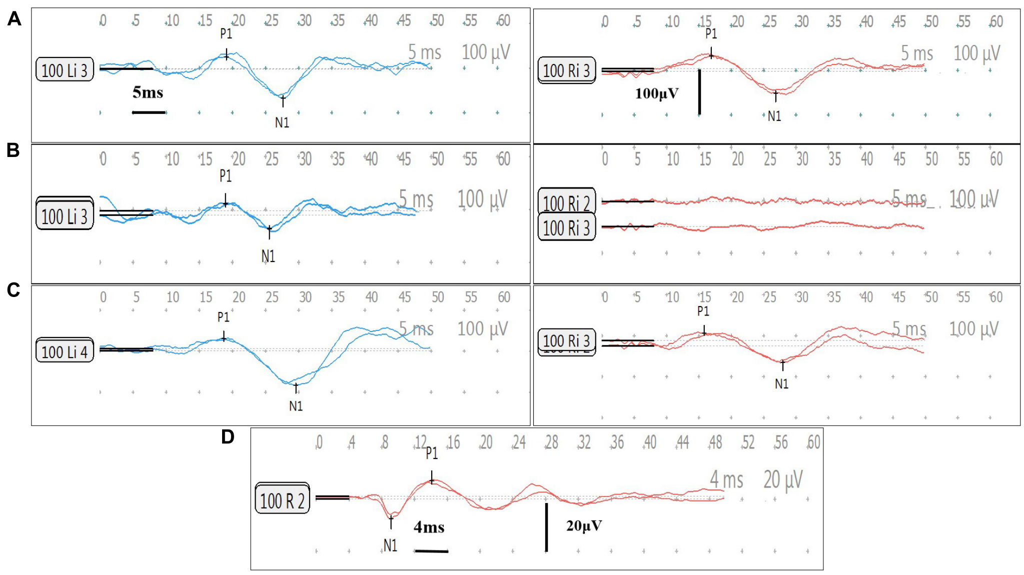

FIGURE 1
Cervical vestibular-evoked myogenic potential (cVEMP) and oVEMP responses of Subject 9. All results are obtained in Subject 9. Her right side is implanted. The left side is blue and the right side is red. The cVEMP response (positive P1 first) and oVEMP response (negative N1 first) are recorded at 100 dB HL (500 Hz tone burst). In cVEMP, the horizontal and vertical calibrations are 5 ms and 100 µV between two adjacent row of points, respectively. In oVEMP, the horizontal and vertical calibrations are 4 ms and 20 µV, respectively. The two traces show the responses of the repeat stimulus. The amplitude ratio (AR) is defined as the difference between the amplitudes of two sides divided by the sum of the amplitudes of two sides. **(A)** Normal bilateral cVEMP responses before surgery. The amplitude is 92.9 µV on the left side and is 82.8 µV on the right, with a AR of 5.7%. **(B)** Normal left cVEMP response and absent right cVEMP response at 1 month after surgery. The amplitude is 60.5 µV on the left and the AR is 100%. **(C)** Normal bilateral cVEMP responses at 12 months post-surgery. The amplitude is 113.2 µV on the left and is 71.9 µV on the right, with a AR of 22.3%. **(D)** Normal right oVEMP response at 12 months post-surgery. The amplitude is 14.6 µV.

TABLE 3 The simultaneous comparison of percent fail rates of vestibular function on the implanted side at all time points.

Tests	Number of patients tested, Percent fail rate (N,%)				
	1 Month $N = 25$	3 Months $N = 25$	6 Months $N = 15$	9 Months $N = 20$	12 Months $N = 16$
Caloric	8/18, 44.44*	9/19, 47.37*	2/11, 18.18	7/15, 46.67*	2/12, 16.67
cVEMP	6/22, 27.27	8/22, 36.36*	1/11, 9.00	5/16, 31.25*	2/11, 18.18
oVEMP	10/21, 47.62*	8/21, 38.10*	4/13, 30.77	7/17, 41.18*	2/11, 18.18
SSC	2/24, 8.33	1/25, 4.00	4/14, 28.57	0/22, 0.00	1/16, 6.25
HSC	4/25, 16.00	4/23, 17.39	3/14, 21.43	1/19, 5.26	0/16, 0.00
PSC	2/24, 8.33	3/25, 12.00	2/14, 14.29	0/19, 0.00	0/16, 0.00

*$p < 0.05$.

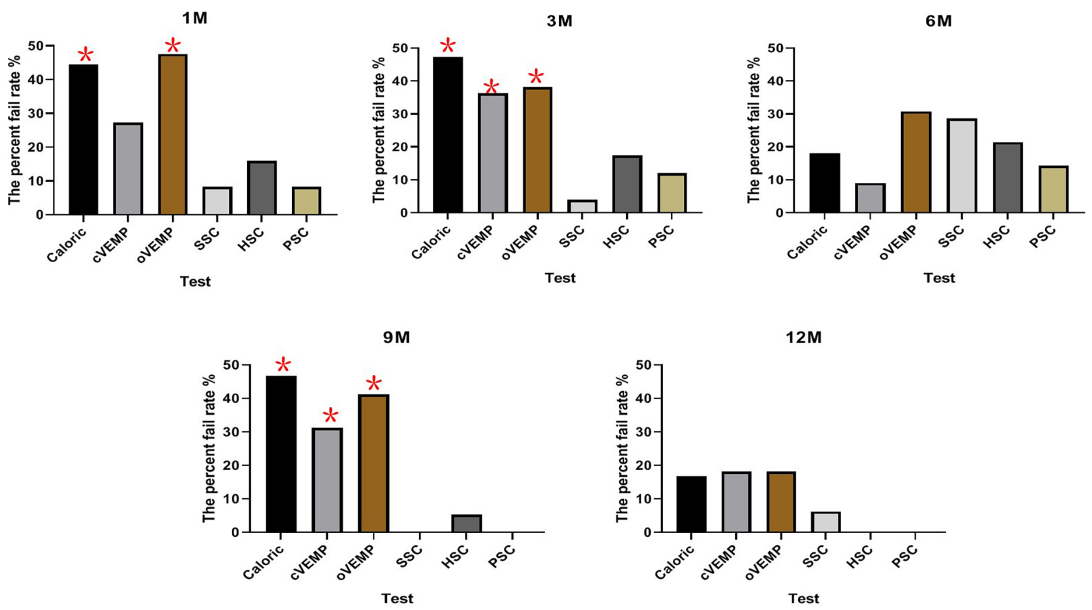

FIGURE 2
The percent fail rates of all the four tests on the implanted side at each time point. The percent fail rate of caloric was significantly higher than vHIT at 1, 3, and 9 months post-operatively ($p < 0.05$). The percent fail rate of oVEMP was higher than that of SSC, HSC, and PSC at 1 and 9 months; the rate was higher than that of SSC and PSC at 3 months post-operatively ($p < 0.05$). At 3 and 9 months post-operatively, the percent fail rate of cVEMP was higher than that of SSC and PSC (*$P < 0.05$).

In the caloric test, the percent fail rates were 47.37% (9/19) and 31.58% (6/19) at 1–3 (1.57 ± 0.90) months and 6–12 (9.63 ± 2.36) months after surgery, respectively. In cVEMP, the rates were 31.43% (11/35) at 1–3 (1.46 ± 0.85) months and were 25.71% (9/35) at 6–12 (9.69 ± 2.31) months. In vHIT of HSC, the rates were 18.42% (7/38) at 1–3 (1.47 ± 0.86) months and were 10.53% (4/38) at 6–12 (9.79 ± 2.28) months.

The percent fail rates showed decreased trends from 1–3 to 6–12 months in caloric ($p = 0.319$) and HSC tested by vHIT ($p = 0.328$), but the trend did not reach statistical significance. There was no significant difference in cVEMP between 1–3 and 6–12 months after surgery ($p = 0.597$). The percent fail rates of oVEMP were the same at 1–3 (1.51 ± 0.89) months and at 6–12 (9.69 ± 2.31) months post-operatively ($p = 1.000$). The percent fail rates of SSC and PSC were the same at 1–3 (1.47 ± 0.86) months and at 6–12 (9.79 ± 2.28) months ($p = 1.000$, respectively). The percent fail rates of each end-organ function on the implanted side at short and long follow-up times after surgery are shown in **Table 4** and **Figure 3**.

Comparison of otolith function variations in the same subjects at short and long follow-up times after surgery

Among these 72 participants, 31 with pre-operative normal otolith functions underwent cVEMP and oVEMP before CI, 1–3 and 6–12 months post-CI simultaneously.

At 1–3 (1.58 ± 0.92) months after surgery, the percent fail rate of cVEMP was 32.26% (10/31), with five children

TABLE 4 The percent fail rate of each end-organ function on the implanted side at all five time points after surgery.

Tests (total patient number)	Number of patients with percent fail rates (*N*, %)	
	1–3 Months	6–12 Months
Caloric (19)	9, 47.37	6, 31.58
cVEMP (35)	11, 31.43	9, 25.71
oVEMP (35)	12, 34.29	12, 34.29
SSC (38)	2, 5.26	2, 5.26
HSC (38)	7, 18.42	4, 10.53
PSC (38)	2, 5.26	2, 5.26

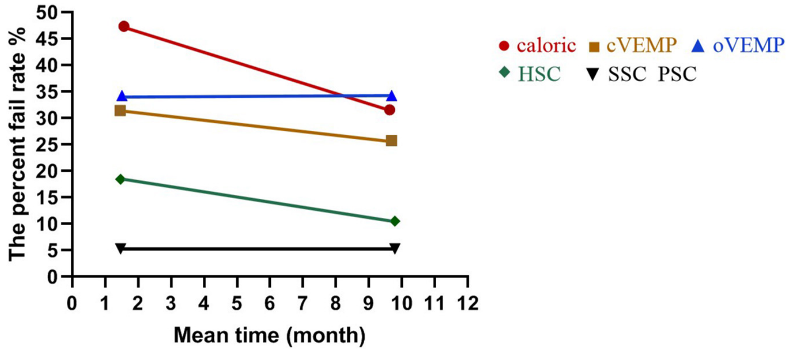

FIGURE 3
The percent fail rate of each end-organ function on the implanted side at short and long follow-up times after surgery. The percent fail rates showed no significant differences from 1–3 (1.57 ± 0.90) to 6–12 (9.63 ± 2.36) months in caloric (p = 0.319) and from 1–3 (1.47 ± 0.86) to 6–12 (9.79 ± 2.28) months in HSC tested by vHIT (p = 0.328) post-operatively. There was no significant difference in cVEMP between 1–3 (1.46 ± 0.85) and 6–12 (9.69 ± 2.31) months (p = 0.597). The percent fail rates of oVEMP were the same at 1–3 (1.51 ± 0.89) and 6–12 (9.69 ± 2.31) months (p = 1.000). The percent fail rates of SSC and PSC were the same at 1–3 (1.47 ± 0.86) and 6–12 (9.79 ± 2.28) months (p = 1.000, respectively).

showing decreased responses and five showing absent responses; the percent fail rate of oVEMP was 35.48% (11/31), with three children having decreased responses and eight having absent responses.

At 6–12 (9.87 ± 2.22) months after surgery, the percent fail rate of cVEMP was 25.81% (8/31), with two children showing decreased responses and six children showing absent responses; the percent fail rate of oVEMP was 35.48% (11/31), with two children having decreased responses and nine having absent responses.

There were no significant differences on percent fail rates of cVEMP and oVEMP between short- and long-terms post-CI ($p > 0.05$).

Comparison of vestibular function between patients with enlarged vestibular aqueduct and patients with a normal cochlea before surgery

Twenty-five patients with EVA and 33 with a normal cochlea underwent all the four tests before surgery. Comparing patients with EVA to normal patients the caloric test showed abnormal responses 32% (8/25) vs. 48.48% (16/33) of the time, respectively; cVEMP 8% (2/25) vs. 36.36% (12/33) of the time, oVEMP 16.00% (4/25) vs. 45.45% (15/33) of the time, vHIT of HSC 4.00% (1/25) vs. 12.12% (4/33) of the time, vHIT of SSC 0.00% (0/25) vs. 3.03% (1/33) of the time, vHIT of PSC 8.00% (2/25) vs. 9.09% (3/33) of the time.

Both the abnormal cVEMP and oVEMP response rates were lower in patients with EVA than patients with a normal cochlea (p = 0.001, 0.018, respectively). There were no significant differences in the abnormal response rates between the two patient groups for caloric test, vHIT of SSC, HSC, and PSC (p = 0.207, 1.000, 0.536, 1.000, respectively).

Discussion

In this report, all 72 patients underwent minimally invasive surgical techniques combining the surgical approach, electrode array, slow insertion of electrode, and systematic glucocorticoids. It is known that this surgical procedure plays an important role in protecting the delicate intracochlear structures (Fina et al., 2003; Eshraghi and Van De Water, 2006; Causon et al., 2015; Kuang et al., 2015; Bruce and Todt, 2018). The concepts of atraumatic electrode insertion include implantation through the RW or extension of the RW (Skarzynski et al., 2002). The RW approach with a straight electrode yielded HP result (Snels et al., 2019). Fifty-one (70.83%) of our patients used flexible electrodes. Although 23.61% of our patients used counter electrodes and four recipients implanted with the Nurotron CS-10A electrode, other protective techniques were used. Only one or two evaluation time points were analyzed in a few previous studies on vestibular function protection with soft surgery (Guan et al., 2021; Sosna-Duranowska et al., 2021; Tsukada and Usami, 2021). The variation in vestibular function at different follow-ups during the first year was analyzed in this study for the first time. Therefore, the trajectory of function variation can be followed dynamically and continuously.

The functional discrepancies in this present study were disparate when all five vestibular end sensor functions were compared simultaneously. Although the subjects' biases were inevitable among different intervals, the results had definite meanings. In this report, at 1 and 3 months post-CI, the percent fail rates of cVEMP and oVEMP were higher than those of a recent report that showed that 19.2% of patients in cVEMP and 17.4% in oVEMP had post-operative function loss at 1–3 months after HP surgery (Sosna-Duranowska et al., 2021). Regarding the vHIT results 4–6 months post-operatively, a similar difference was observed. The main reason for these differences could be our stricter criteria used to judge abnormal responses. In this study, at 6 months after surgery, the percent fail rates of cVEMP and oVEMP were 9.00 and 30.77%, respectively, and were both 18.18% at 12 months. These results were consistent with previous results at 6–12 months after less traumatic CI surgery (Tsukada and Usami, 2021).

Our results revealed that otolith and low-frequency HSC functions were damaged more seriously than high-frequency canal functions in the short-term post-operatively (1 and 3 months). These short-term outcomes agree with the changes affected by conventional surgery, which showed that the otolith sensors were more damaged than three semicircular canal functions and the canal functions were seldom impaired (Ibrahim et al., 2017; Dagkiran et al., 2019; Yong et al., 2019). No selective impairments were found at 6 and 12 months although they existed in the 9th month, indicating that the impairment discrepancies among all sensors began to decrease at nearly 6 months when the distant or secondary effects come into play. Fluctuating hearing changes have been reported in patients after HP surgery. The foreign body response and intracochlear fibro-osseous reaction may contribute to this fluctuation (Foggia et al., 2019; Snels et al., 2019). The reason for our variation in the 9th month may be the individual disparity or other reasons such as foreign body response or fibro-osseous. However, the exact mechanism is unknown. This study discovered no damage discrepancy at 12 months, in contrast with previous reports with the same duration after conventional implantation (Ibrahim et al., 2017; Yong et al., 2019). Our study demonstrated that the atraumatic techniques could diminish the functional impairment at least 1 year post-surgery, being more obvious in the long-term period. It was meaningful and comprehensive to assess function status at multiple time points to display the continuously functional variations.

To deeply explore the status of each function, we evaluated it within the same patient group for each vestibular sensor. Our otolith function damage was less than most of the results through conventional surgery (Verbecque et al., 2017; Wang et al., 2022) and our canal functions were seldom damaged. These results verified the validity of protective surgical techniques. In addition, our results showed a similar status of all the functional variations from 1–3 to 6–12 months although there were decreased tendencies of damage in HSC and saccular functions. Finally, we analyzed the otolith function variation on the same cohort subjects and found the same results. A recent study revealed that conventional surgery could injure all five vestibular end-organ functions and the damage of some functions increased with time (Kwok et al., 2022). Conversely, a different tendency was observed in this report.

With regard to our injury trends, we speculated that the instant mechanical injury produced by electrode insertion might not be the primary damage because the electrode does not come in direct contact with the vestibular organs, although the instant damage could be diminished through our protective methods. Besides this, the secondary or distant effects of surgery may threaten vestibular function, such as inflammation, fibrous tissue formation, or ossification (Fayad et al., 2009). Cochlear fibrosis and new bones can be induced by a traumatic electrode insertion (Fayad et al., 2009; Jia et al., 2016). The acute inflammatory response is due to electrode insertion. Then a chronic phase replaces it because of the foreign body reaction involving macrophages, their derivatives, and lymphocytes (Seyyedi and Nadol, 2014). These influencing factors may mainly participate in vestibular damage (Stuermer et al., 2019). Our surgical methods are believed to diminish the damage from inflammation, fibrous tissue information, or ossification to avoid producing an increased damage with time.

Thirty-three (45.83%) participants revealed an EVA and their abnormal otolith function was less common than patients with a normal cochlea before surgery in this study, consistent with our previous results (Wang et al., 2021). It is hypothesized that the presence of a third window might allow for the activation of VEMP, making the otolith organs more excitable and sensitive to sound stimulation (Govender et al., 2016; Wang et al., 2021). However, the influence of the cochlear anatomical malformation on our results were not analyzed because of the inconsistencies in patients pools at different follow-ups. We mainly focused on the overall effect of protective surgical techniques on vestibular function variations and will explore the influence factors in the future.

Limitations

The main limitation was that some patients were lost to follow-ups after implantation in the clinic. The number of patients evaluated were disparate in the first part of our results. In the next step, we will evaluate vestibular function variations at five continuous follow-up times in the same cohort of subjects for each vestibular sensor.

Conclusion

In this study, the variation of vestibular function in the short and long terms after a minimally invasive CI surgery during a 12 months period was explored. Most of the vestibular functions can be preserved with no damage discrepancy among the otolith and three semicircular canal functions at 12 months post-CI. Interestingly, a similar pattern of changes in vestibular function was found during the early and the later stages of recovery after surgery. Both the instant influence of electrode insertion and the indirect or secondary factors show a similar trend on the functional variation of each vestibular end sensor after a minimally invasive CI surgery.

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Medical Ethics Committee of Shandong Provincial ENT Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

LX, HW, RW, and XC contributed to the conception of the work. JL and ZF contributed to the experimental design. RW selected data and performed the analysis. All authors contributed to the interpretation of the data and were involved in writing the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China under grants (numbers: 81670932, 81600803, and 82071053).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Boje Rasmussen, K. M., West, N., Tian, L., and Thomasen, P. C. (2021). Long-term vestibular outcomes in cochlear implant recipients. *Front. Neurol.* 12:686681. doi: 10.3389/fneur.2021.686681

Bruce, I. A., and Todt, I. (2018). Hearing preservation cochlear implant surgery. *Adv. Otorhinolaryngol.* 81, 66–73. doi: 10.1159/000485544

Causon, A., Verschuur, C., and Newman, T. A. (2015). A retrospective analysis of the contribution of reported factors in cochlear implantation on hearing preservation outcomes. *Otol. Neurotol.* 36, 1137–1145. doi: 10.1097/MAO.0000000000000753

Dagkiran, M., Tuncer, U., Surmelioglu, O., Tarkan, O., Ozdemir, S., Cetik, F., et al. (2019). How does cochlear implantation affect five vestibular end-organ functions and dizziness?. *Auris Nasus Larynx* 46, 178–185. doi: 10.1016/j.anl.2018.07.004

Douchement, D., Terranti, A., Lamblin, J., Salleron, J., Siepmann, F., Siepmann, J., et al. (2015). Dexamethasone eluting electrodes for cochlear implantation: Effect on residual hearing. *Cochlear Implants Int.* 16, 195–200. doi: 10.1179/1754762813Y.0000000053

Eshraghi, A. A., and Van De Water, T. R. (2006). Cochlear implantation trauma and noise-induced hearing loss: Apoptosis and therapeutic strategies. *Anat. Record A Discov. Mol. Cell. Evol. Biol.* 288, 473–481. doi: 10.1002/ar.a.20305

Fayad, J. N., Makarem, A. O., and Linthicum, F. H. (2009). Histopathologic assessment of fibrosis and new bone formation in implanted human temporal bones using 3D reconstruction. *Otolaryngol. Head Neck Surg.* 141, 247–252. doi: 10.1016/j.otohns.2009.03.031

Fina, M., Skinner, M., Goebel, J. A., Piccirillo, J. F., Neely, J. G., and Black, O. (2003). Vestibular dysfunction after cochlear implantation. *Otol. Neurotol.* 24, 234–242. doi: 10.1097/00129492-200303000-00018

Foggia, M. J., Quevedo, R. V., and Hansen, M. R. (2019). Intracochlear fibrosis and the foreign body response to cochlear implant biomaterials. *Laryngoscope Invest. Otolaryngol.* 4, 678–683. doi: 10.1002/lio2.329

Govender, S., Fernando, T., Dennis, D. L., Welgampola, M. S., and Colebatch, J. G. (2016). Properties of 500 Hz air- and bone-conducted vestibular evoked myogenic potentials (VEMPs) in superior canal dehiscence. *Clin. Neurophysiol.* 127, 2522–2531. doi: 10.1016/j.clinph.2016.02.019

Guan, R. R., Wang, Y. Q., Wu, S. S., Zhang, B., Sun, J. W., Guo, X. T., et al. (2021). Vestibular function in children and adults before and after unilateral or sequential bilateral cochlear implantation. *Front. Neurol.* 12:675502. doi: 10.3389/fneur.2021.675502

Ibrahim, I., da Silva, S. D., Segal, B., and Zeitouni, A. (2017). Effect of cochlear implant surgery on vestibular function: Meta-analysis study. *J. Otolaryngol. Head Neck Surg.* 46:44. doi: 10.1186/s40463-017-0224-0

Jia, H., François, F., Bourien, J., Eybalin, M., Lloyd, R. V., Van De Water, T. R., et al. (2016). Prevention of trauma-induced cochlear fibrosis using intracochlear application of anti-inflammatory and antiproliferative drugs. *Neuroscience* 316, 261–278. doi: 10.1016/j.neuroscience.2015.12.031

Kuang, H., Haversat, H. H., and Michaelides, E. M. (2015). Impairment of caloric function after cochlear implantation. *J. Speech Lang. Hear. Res.* 58, 1387–1395. doi: 10.1044/2015_JSLHR-H-15-0010

Kwok, B. Y. C., Rosengren, S. M., Kong, J. H. K., Birman, C. S., Hoskisson, E., Young, A. S., et al. (2022). Impact of Cochlear Implantation on Canal and Otolith Function. *Otol. Neurotol.* 43, 304–312. doi: 10.1097/MAO.0000000000003459

Li, X., and Gong, S. S. (2020). The effect of cochlear implantation on vestibular evoked myogenic potential in children. *Laryngoscope* 130, e918–e925. doi: 10.1002/lary.28520

Ordonez, F., Riemann, C., Mueller, S., Sudhoff, H., and Todt, I. (2019). Dynamic intracochlear pressure measurement during cochlear implant electrode insertion. *Acta Otolaryngol.* 139, 860–865. doi: 10.1080/00016489.2019.1640391

Seyyedi, M., and Nadol, J. B. (2014). Intracochlear inflammatory response to cochlear implant electrodes in humans. *Otol. Neurotol.* 35, 1545–1551. doi: 10.1097/MAO.0000000000000540

Sichnarek, J., Mrazkova, E., Zathurecky, E., and Tomaskova, H. (2019). Comparing results from vestibular caloric stimulation and vHIT from a specialised outpatient clinic. *Int. Tinnitus J.* 23, 1–5. doi: 10.5935/0946-5448.20190001

Skarzynski, H., Lorens, A., D'Haese, P., Walkowiak, A., Piotrowska, A., Sliwa, L., et al. (2002). Preservation of residual hearing in children and post-lingually deafened adults after cochlear implantation: An initial study. *ORL J. Otorhinolaryngol. Relat.* 64, 247–253. doi: 10.1159/000064134

Skarzynski, H., Lorens, A., Dziendziel, B., Rajchel, J. J., Matusiak, M., and Skarzynski, P. H. (2019). Electro-natural stimulation in partial deafness treatment of adult cochlear implant users: Long-term hearing preservation results. *ORL* 81, 63–72. doi: 10.1159/000497060

Snels, C., IntHout, J., Mylanus, E., Huinck, W., and Dhooge, I. (2019). Hearing preservation in cochlear implant surgery: A meta-analysis. *Otol. Neurotol.* 40, 145–153. doi: 10.1097/MAO.0000000000002083

Sosna-Duranowska, M., Tacikowska, G., Gos, E., Krupa, A., Skarzynski, P. H., and Skarzynski, H. (2021). Vestibular Function After Cochlear Implantation in Partial Deafness Treatment. *Front. Neurol.* 12:667055. doi: 10.3389/fneur.2021.667055

Sprinzl, G. M., Schoerg, P., Edlinger, S. H., and Magele, A. (2020). Long-term hearing preservation in electric acoustic cochlear implant candidates. *Otol. Neurotol.* 41, 750–757. doi: 10.1097/MAO.0000000000002627

Stuermer, K. J., Kluenter, H. D., Lang-Roth, R., Schwarz, D., Hüttenbrink, K. B., and Anagiotos, A. (2019). Preservation of vestibular function and residual hearing after round window cochlear implantation. *Otol. Neurotol.* 40, 878–882. doi: 10.1097/MAO.0000000000002257

Tsukada, K., and Usami, S. I. (2021). Vestibular preservation after cochlear implantation using the round window approach. *Front. Neurol.* 12:656592. doi: 10.3389/fneur.2021.656592

Valvassori, G. E., and Clemis, J. D. (1978). The large vestibular aqueduct syndrome. *Laryngoscope* 88, 723–728. doi: 10.1002/lary.1978.88.5.723

Verbecque, E., Marijnissen, T., De Belder, N., Van Rampaey, V., Boudewyns, A., van de Heyning, P., et al. (2017). Vestibular (dys) function in children with sensorineural hearing loss: A systematic review. *Int. J. Audiol.* 56, 361–368. doi: 10.1080/14992027.2017.1281444

Wang, R. J., Chao, X. H., Luo, J. F., Zhang, D. G., Xu, J. L., Liu, X. F., et al. (2022). Objective vestibular function changes in children following cochlear implantation. *JVR* 32, 29–37. doi: 10.3233/VES-190763

Wang, R. J., Zhang, D. G., Luo, J. F., Chao, X. H., Xu, J. L., Liu, X. F., et al. (2021). Influence of cochlear implantation on vestibular function in children with an enlarged vestibular aqueduct. *Front. Neurol.* 12:663123. doi: 10.3389/fneur.2021.663123

West, N., Tian, L., Vang Petersen, L. K., Bille, M., Klokker, M., and Thomasen, P. C. (2021). Objective vestibular test battery and patient reported outcomes in cochlear implant recipients. *Otol. Neurotol.* 42, e416–e424. doi: 10.1097/MAO.0000000000002959

Yong, M., Young, E., Lea, J., Foggin, H., Zaia, E., Kozak, F. K., et al. (2019). Subjective and objective vestibular changes that occur following paediatric cochlear implantation: Systematic review and meta-analysis. *J. Otolaryngol. Head Neck Surg.* 48:22. doi: 10.1186/s40463-019-0341-z

Zhang, D. G., Lv, Y. F., Han, Y. C., Li, Y. W., Li, X. F., Wang, J., et al. (2019). Long-term outcomes of triple semicircular canal plugging for the treatment if intractable Meniere's disease: A single center experience of 361 cases. *J. Vestibul. Res.* 29, 315–322. doi: 10.3233/VES-190682