Running Head: RECOGNIZING VOICES THROUGH COCHLEAR IMPLANTS

| 1 | [This is a preprint of a manuscript that has been published at JSLHR. |
|----|--|
| 2 | The final version is available at https://doi.org/10.1044/2022_JSLHR-21-00209] |
| 3 | |
| 4 | |
| 5 | |
| 6 | |
| 7 | |
| 8 | Recognizing voices through a cochlear implant: A systematic review of voice perception, talker |
| 9 | discrimination, and talker identification |
| 10 | |
| 11 | |
| 12 | Sarah Colby ¹ , Adriel John Orena ² |
| 13 | |
| 14 | ¹ Department of Psychological and Brain Sciences, University of Iowa, |
| 15 | G60 Psychological and Brain Sciences Building, 340 Iowa Ave., Iowa City, Iowa, 52242, USA |
| 16 | ² Department of Psychology, University of British Columbia, |
| 17 | 2136 West Mall, Vancouver, British Columbia, V6T 1Z4, Canada |
| 18 | sarah-colby@uiowa.edu, aorena@psych.ubc.ca |
| 19 | |

20

ABSTRACT

Objective: Some cochlear implant (CI) users report having difficulty accessing indexical information in the speech signal, presumably due to limitations in the transmission of fine spectrotemporal cues. The purpose of this paper was to systematically review and evaluate the existing research on talker processing in CI users. Specifically, we reviewed the performance of CI users in three types of talker- and voice-related tasks. We also examined the different factors (such as participant, hearing and device characteristics) that might influence performance in these specific tasks.

Design: We completed a systematic search of the literature with select keywords using citation aggregation software to search Google Scholar. We included primary reports that tested: i) talker discrimination; ii) voice perception, and iii) talker identification. Each report must have had at least one group of participants with cochlear implants. Each included study was also evaluated for quality of evidence.

Results: The searches resulted in 1561 references, which were first screened for inclusion and then 33 34 evaluated in full. Forty-three studies examining talker discrimination, voice perception, and talker identification were included in the final review. Most studies were focused on postlingually 35 36 deafened and implanted adult CI users, with fewer studies focused on prelingual implant users. In general, CI users performed above chance in these tasks. When there was a difference between 37 groups, CI users performed less accurately than their normal-hearing (NH) peers. A subset of CI 38 39 users reached the same level of performance as NH participants exposed to noise-vocoded stimuli. Some studies found that CI users and NH participants relied on different cues for talker perception. 40 41 Within groups of CI users, there is moderate evidence for a bimodal benefit for talker processing, 42 and there are mixed findings about the effects of hearing experience.

43 **Conclusion:** The current review highlights the challenges faced by CI users in tracking and 44 recognizing voices and how they adapt to it. While large variability exists, there is evidence that 45 CI users can process indexical information from speech, though with less accuracy than their NH 46 peers. Recent work has described some of the factors that might ease the challenges of talker 47 processing in CI users. We conclude by suggesting some future avenues of research to optimize 48 real-world speech outcomes.

49

- 50 *Keywords:* Cochlear implants, Talker perception, Talker identification, Systematic review
- 51

52

1. Introduction

53 Perceiving the indexical properties of the speech signal is a fundamental communicative ability that is often taken for granted (see Sidtis & Kreiman, 2012 for a review). Indeed, many 54 55 studies have shown that normal-hearing (NH) listeners are adept at perceiving indexical 56 information: they achieve high degrees of accuracy in identifying an unfamiliar talker's age, 57 gender, and their regional background (Bradlow & Bent, 2008; Perry et al., 2001; G. E. Peterson & Barney, 1952). Even infants are able to recognize their mother's voice from birth (DeCasper & 58 Fifer, 1980). Given the intersensory redundancy between faces and voices, it is tempting to dismiss 59 60 voice recognition as a trivial skill. Indeed, in daily conversations, we can often tell who is speaking because we have visual confirmation. However, familiarity with a speaker's voice does not only 61 62 allow us to track who is speaking, it also lends itself to efficient social communication and 63 linguistic processing.

For example, familiarity with a talker's voice allows listeners to deal with talker-specific 64 variability in speech. In quiet speech, the same acoustic sound could be distinguished as one of 65 66 two different phonemes depending on who the listener thinks they are listening to (Johnson et al., 1999). When linguistic and speaker cues conflict, listeners are slower at categorizing speech 67 68 sounds into phonemes (Apfelbaum et al., 2014), indicating an integration of talker and linguistic information in speech processing. These findings have implications on speech perception: 69 70 familiarity with a talker's voice results in improved word and sentence intelligibility, improved 71 recognition memory, and decreased processing time (Clarke & Garrett, 2004; Theodore et al., 2015). These benefits are heightened in noisy situations (Nygaard & Pisoni, 1998), highlighting, 72 73 once again, that in real-life contexts, knowing who is speaking leads to efficient communication.

74 Voice recognition does not come easily to all individuals. One group that may have 75 difficulty with aspects of talker processing is cochlear implant (CI) users. A CI is a sensory aid for 76 individuals with severe-to-profound sensorineural hearing loss. CI devices function by gathering 77 information about the fluctuation of sound energy over time within frequency bandwidths, 78 converting them into electrical impulse patterns, and directing them to specific electrodes located 79 along the cochlea. While the CI does not restore acoustic hearing, it provides recipients with hearing sensitivity within the speech range. Many individuals with CIs receive gains in decoding 80 linguistic information, including in speech perception and word recognition (e.g., Blamey et al., 81 82 2012; N. R. Peterson et al., 2010). With experience, CI users become adept at discriminating voices from environmental sounds (Massida et al., 2011). Some children with CIs even perform 83 84 comparably to their NH peers when categorizing human vocalizations from environmental sounds 85 (Berland et al., 2019). Nonetheless, it has been argued that there are limitations on the use of cochlear implants for talker processing. 86

Why might this be? The electric speech signal transmitted via modern CIs are degraded, 87 88 especially when compared to the acoustic speech signal (see Baskent et al., 2016 for review). Different CI devices employ different processing strategies to transform the speech signal, but a 89 90 common transformation strategy relies on filtering the acoustic signal into bands of frequencies, 91 resulting in blurred frequency variations within bandpass channels. Moreover, there are physiological limitations of electrical stimulation of the auditory nerve. The auditory nerve 92 93 responds differently to electric stimulation than acoustic stimulation. Further, the spatial overlap of the broad stimulation from individual electrodes leads to blurred spatial activation patterns. 94 95 Thus, it is not yet possible to achieve very fine-tuned simulation points with electrodes.

96 This degradation of the speech signal can make a host of speech and sound perception tasks 97 more difficult for CI users (Ciocca et al., 2002; Hopkins & Moore, 2009). Specifically, to 98 distinguish voices, NH listeners rely on a combination of different acoustic cues, but fundamental 99 frequency (F0) and cues to vocal tract length (VTL) appear to be the most helpful for talker 100 perception (Skuk & Schweinberger, 2014). Both of these cues rely on the harmonic and formant 101 structure of speech, but how they are encoded by the CI device and how they are perceived by CI 102 users is quite limited. For example, pitch can be coded through simulation rate, temporal pattern 103 of stimulation, or place of stimulation – but, even among these coding strategies, the percept of 104 pitch is consistently reported as being weak (Moore & Carlyon, 2005). Likewise, the blurred 105 spectral resolution appears to negatively affect CI users' ability to perceive VTL cues (Gaudrain 106 & Baskent, 2015). Taken together, device and physiological constraints appear to not be conducive 107 for talker perception.

108 Beyond the ability to encode acoustic cues, access to a language's sound structure assists 109 talker recognition by allowing listeners to distinguish between variability in speech sounds and 110 variability in different talker's voices (see Creel & Bregman, 2011 for review). This in turn helps 111 individuals track and adapt to the idiosyncrasies of a person's voice. Indeed, prior work has shown 112 a gradient influence of phonological processing on talker recognition tasks. Listeners have 113 heightened talker recognition skills for talkers who speak their native language, compared to 114 talkers with a different accent (Vanags et al., 2005) or talkers who speak the listener's second 115 language (Bregman & Creel, 2014). Further, individuals who perform worse on phonological 116 processing tasks, such as individuals with dyslexia, tended to have difficulty with talker 117 recognition (Kadam et al., 2016; Perrachione et al., 2011).

118 Based on these findings, the predictions for the performance of CI users on talker 119 processing tasks are clear. The degradation of the speech signal should make talker processing a 120 challenging task for CI users, especially compared to listeners with typical hearing. Here, we 121 review the literature to systematically compile the evidence. How well do CI users perform on 122 tasks of talker discrimination, voice perception, and talker identification, compared to their 123 normal-hearing peers? What factors (e.g., participant, hearing, device) affect their performance on 124 these tasks? Is their performance on these tasks related to their performance on other linguistic 125 tasks?

126 We chose to focus on studies with participants who are CI users, instead of focusing on 127 experiments that used CI-simulated stimuli to investigate talker processing (i.e., through degrading 128 stimuli with noise vocoding). Because NH participants are easier to recruit, presenting degraded 129 stimuli can serve as a first step in uncovering interesting avenues of investigation for improving 130 CI outcomes (Krull et al., 2012). These studies provide a useful foundation to further examine perceptual skills of CI users. However, these studies have two limitations. First, the vocoded 131 132 manipulations that researchers use in these experiments are based on the manipulations that CI 133 devices do in their processors, but what CI users actually experience were reported to differ than 134 how the simulations sound. Thus, while vocoded stimuli are a good approximate, they are not 135 always analogous. Second, CI users and NH listeners rely on different cues for speech or talker 136 perception since their prior experience with the speech signal differs (e.g., Fuller et al., 2014). On 137 that note, it is important to acknowledge that CI users are a heterogeneous group. For example, CI users can differ by chronological age, device characteristics, onset and duration of deafness, length 138 139 of CI use, and communication mode - each of which have been shown to contribute to a CI user's

level of success in speech perception (Belzner & Seal, 2009; Roberts et al., 2013). It remains a
question of whether the same effects hold for talker processing.

142 Of particular interest, CI users can also have different configurations of devices, which will 143 provide them with different levels of access to sounds. Individuals with unimodal CIs (whether 144 unilaterally, or bilaterally) typically have no residual acoustic hearing in either ear, and thus must 145 rely exclusively on the electric stimulation from their implant. Conversely, individuals with 146 bimodal or hybrid CIs may have residual acoustic hearing. In these contexts, a hearing aid or the 147 hybrid CI could amplify low-frequency sounds. If talker processing relies extensively on the 148 fundamental frequency, then we would expect improved voice recognition for both bimodal and 149 hybrid CI users. But, if talker processing also relies on higher frequency spectral information, then 150 bimodal or hybrid CI configurations might not result in performance levels equal to those of NH 151 participants.

152 In the present review, we will summarize the current knowledge on talker processing in 153 individuals with cochlear implants, with a focus on talker discrimination and identification. 154 Reflecting the heterogeneity of CI users in the general population, these studies vary in the 155 participants that they have recruited (i.e., in age, device configuration, and hearing experience). 156 Further, they vary in the methodologies that they use to assess talker processing. This has produced 157 a wealth of evidence in various ways, which we now aim to analyze to extract overall coherent 158 findings. First, we scanned the literature, and we tracked the following characteristics in order to 159 identify any research gaps in the literature:

161

160

(1) *Participant characteristics:* As previously mentioned, there is large heterogeneity in the hearing experiences of CI users. Here, we summarized the characteristics of CI users who

- have taken part in studies on talker processing in order to identify which groups the current
 results are relevant for, and to identify potential gaps in the literature.
- 164 (2) *Task characteristics:* Talker processing can be assessed in various ways, and prior work
 165 has suggested that different paradigms can lead to different conclusions (see Perrachione,
 166 2017). Here, we reviewed the different tasks that are being used with CI users, with a focus
 167 on the task paradigms and the types of stimuli being used. We use the term talker
 168 processing to refer to the broad topic of this review. When referring to results that apply to
 169 a task, we specify talker discrimination, talker identification or voice perception. The
 170 breakdown of these categories is described in the Results section.

171 Further, we reviewed the literature to examine the following research questions:

- (1) *Comparison with NH listeners:* NH listeners are often used as control groups in
 experiments assessing CI users' performance. Thus, we reviewed the literature to examine
 the performance of CI users, as a group, in talker processing tasks, compared to NH
 listeners.
- 176 (2) Comparison between device configurations: Here, we reviewed the studies that examine 177 whether different hearing configurations impact the manner with which CI users process 178 talker information. If residual acoustic hearing can help with voice perception, then we 179 would expect improved talker processing for bimodal CI users, compared to unimodal 180 users. If access to these low-frequency sounds drive voice perception, we might even 181 expect bimodal CI users to perform as well as NH listeners. In contrast, if talker processing also relies on high-frequency spectral information, then bimodal or hybrid CI 182 183 configurations would not result in performance levels equal to those of NH participants.

(3) *Role of different acoustic cues:* NH listeners take advantage of various cues, such as F0
and acoustic correlates of VTL (e.g., distribution of formant peaks), to encode talker
identity (Smith & Patterson, 2005). Given that CI users do not have the same access to
these cues, an open question is whether CI users make use of different cues for identifying
voices. Here, we report the findings from studies that investigated how various acoustic
cues influence talker information processing in CI users.

190 (4) Relationship to participant characteristics: Tracking the trajectory of talker processing 191 for different CI users is important for developing expectations about speech processing 192 performance. However, this task is complicated by the wide heterogeneity in demographic 193 and hearing experiences of CI users. Indeed, individuals with different hearing experience, 194 such age at implantation, can have different hearing outcomes (e.g., Manrique et al., 2004). 195 For example, in an investigation of melodic contour identification, Tao and colleagues 196 (2015) found that post-lingual CI users outperformed pre-lingual users. Here, we explored how developmental age and hearing experience (including age of onset of deafness, age at 197 198 implantation, and duration of CI use) might map onto talker processing abilities.

(5) Relationship to linguistic tasks: Prior work has shown that linguistic and indexical
 processing is intertwined in speech processing (e.g., Creel & Bregman, 2011). While some
 individuals might broadly perform well across a variety of speech tasks, it is equally likely
 that certain strengths and weaknesses will arise within the broad domain of speech
 processing. Thus, it is of interest to examine how performance in talker processing tasks
 might relate to their performance in other linguistic tasks, such as consonant/vowel
 perception or word/sentence recognition.

206

207 208 *2*.

2. Methods

208 2.1. Search strategy

Figure 1 summarizes the procedural outcomes of this systematic review. The authors ran several searches using the software Publish Or Perish (version 7.15.2643.7260; Harzing, 2007) on Google Scholar. The authors used the search terms "cochlear implant" and every combination of "voice", "talker", and "gender", with either "identification", "generalization", or "discrimination". The searches took place between February 12, 2019 and March 3, 2019. This resulted in 1499 potentially relevant articles. After removing duplicates, 1124 articles remained for further review. The same searches were run again between May 13, 2020 and May 22, 2020 to find any new



Figure 1. PRISMA chart of the study selection process.

216 research published in the last year. These searches returned 157 results, which resulted in an 217 additional 115 articles to review after duplicates were removed. This led to a total of 1239 articles. 218 The following inclusion criteria was used to identify relevant articles. We included (1) only 219 primary reports with an experimental study (excluding reports in specialized format, such as 220 theses, dissertations, secondary reports, and conference abstracts); (2) only reports written in 221 English; (3) studies that have at least one group of participants that use cochlear implants; (4) a 222 measure of talker or gender perception, either via an identification or discrimination task. No 223 restrictions were imposed on patient characteristics.

224 *2.2. Coding the studies*

In the first triage, the authors scanned the titles and abstracts of 1239 records. Each 225 226 author was responsible for reviewing half of the records. Articles were coded as either "Include", 227 "Exclude", or "Maybe" following our criteria laid out in the previous section. Because the initial 228 triage was based on the title and abstract alone, some articles were ambiguous in whether they met our inclusion criteria (i.e., did not explicitly state if their participants were CI users or did 229 230 not clearly fit our task criteria). These were coded as "Maybe" and were reviewed by both 231 authors on the basis of title and abstract, and a mutual decision was made regarding their 232 inclusion status. Following review of the abstracts, 48 articles were included for further review.

Each author was responsible for independently reviewing the full-text of half of the remaining articles. Upon more careful review, nine additional articles were excluded after reading the full text for not meeting the inclusion criteria. This left 39 studies in our full review. Participant characteristics (sample size, age, onset of deafness, duration of deafness, age of implantation, length of use, hearing configuration) and study characteristics (goals, outcome measures, stimuli, experimental task, major findings) were summarized for each study. Note that some of these 239 descriptive data were not directly available from the text of manuscripts. In cases where 240 participant-level data was available in tables, we calculated their averages (see footnote of Tables 241 1, 2 and 3). Two research assistants independently checked the tables for content accuracy. Any 242 clarification or uncertainty was discussed between authors and research assistants until a mutual 243 decision was reached. Studies are reported in three separate tables based on their experimental 244 task: Table 1 summarizes the participant and study characteristics of the studies investigating talker discrimination, Table 2 summarizes the voice perception studies, and Table 3 summarizes the 245 246 talker identification studies.

247 2.3. Additional search

Per the suggestion of a reviewer, we conducted a third search with the additional search terms of "voice perception", "talker perception", and "gender perception", along with "cochlear implant". This search returned 483 articles, and after removing duplicates, resulted in 332 unique articles. Of these, 12 articles were identified as being relevant after reviewing abstracts. After a full-text review, only 4 of these articles were included in the systematic review. We thus include 43 articles in this review.

254 2.3. Quality assessment

To assess the quality level of the selected studies, we used a methodological quality appraisal tool based on the Grades of Recommendation, Assessment, Development, and Evaluation Working Group (GRADE) approach (Higgins & Green, 2006). Specifically, we adapted an appraisal tool, developed by Downs & Black (1998), for this particular systematic review. The following questions were asked: **Q1**) Was the objective of the study clearly defined? **Q2**) Was the participant inclusion criteria clearly described? **Q3**) Are the main study findings, as pertains to talker discrimination or identification, clearly stated? **Q4**) Are the main outcome measures, as pertains to talker discrimination or identification, clearly stated? **Q5**) Were the investigators blinded to the participant characteristics to reduce bias? **Q6**) Is there a clarification for the appropriateness of the sample size studied? Note that this scale is not necessarily assessing the quality of the evidence, rather our ability to interpret findings based on what was presented in the manuscripts.

267 Each author assessed $\sim 75\%$ of the selected studies (30 papers), such that approximately 50% of the papers were assessed by both authors (21 papers). Each study was given one point for 268 each "Yes" to the questions above, for a total of 6 points. Studies that received 0-2 points were 269 270 categorized as "Weak"; studies that received 3-4 points were categorized as "moderate", and 271 studies that received 5-6 points were categorized as "strong". Thus, a "weak" manuscript is one 272 that did not include details that would allow for clear interpretation, while a "strong" study is one 273 that included details that allows for clear interpretation. In total, there were 3 "weak" studies, 24 "moderate" studies, and 16 "strong" studies. Inter-coder ratings were consistent, as none of the 274 275 rating categories among the papers reviewed by both authors differed. The ratings are included in 276 each study's entry in their respective tables (Table 1, 2, or 3).

277

3. Results

278 *3.1. Participant characteristics*

Tables 1, 2, and 3 display the various characteristics of participants in studies covered by this systematic review. Records included a total of 1165 participants with CIs and 381 participants with NH. Certain sets of studies acknowledge being subgroups of one another, resulting in overlapping samples, which means that there were fewer than 1546 unique participants overall. The reviewed studies tended to have small sample sizes, with a median sample size of 15. Group sample sizes for each study are reported in Tables 1, 2, and 3.

285 In general, there were two main groups of CI users being tested: those focused on adults 286 (i.e., mean age over 18 years of age; n = 30), and those focused on children (i.e., mean age under 18 years of age; n = 13). As is typical in the CI literature, studies with adult participants were 287 288 mostly focused on middle to late adulthood with postlingual deafness (see Figure 2). For studies 289 with adults that report these characteristics, the mean onset of deafness was 38.5 years of age (SD = 11.87) and the mean age of implantation was 48.59 years of age (SD = 13.94). There was only 290 291 one study that specifically focused on adult CI users with a prelingual onset of deafness (Zaltz et 292 al., 2018). Participants in this study had an onset of deafness before 1.5 years, but a range of 293 implantation ages from 2.25 - 33.3 years old. Thus, this study is the only study to include early 294 deafened, late implanted individuals. One study had a mix of prelingual and postlingual adult CI 295 users, with one child CI user (Skuk et al., 2020), although these factors were not a focus of 296 investigation. Studies with child participants focused on older children or adolescents with prelingual deafness. The study with the youngest participants was conducted by van Heugten et 297 298 al. (2014), who recruited children ages 4- to 7-years old. For child studies that report these 299 characteristics, the mean onset of deafness was 0.56 years of age (SD = 0.30) and the mean age of 300 implantation was 3.75 years of age (SD = 2.37). As will be discussed later, these descriptive 301 analyses reveal several gaps of research, including a lack of research on several age ranges that is 302 typical in the CI literature (i.e., children, adolescents, and early adulthood).

Among unimodal CI users, seven studies included only participants with unilateral CIs and one study included only participants with bilateral CIs. Seven studies included a mixed group of unilateral and bilateral CIs. There were three studies that only recruited bimodal CI users, and nine studies recruited a mix of bimodal and unimodal CI users. The remaining twelve studies did not report information about the CI configuration of their participants.



Figure 2. Descriptive statistics of the age of participants in papers with A) talker discrimination tasks, B) voice perception tasks, and C) talker identification tasks. Filled circle dots indicate the mean age, solid horizontal lines indicate the standard deviation of age, and dotted horizontal lines indicate the age range. Orange dots and lines represent data for CI participants, while blue dots and lines represent data for NH participants. Numbers in parentheses following the citation indicate different CI groups within the same study.

309 *3.2. Task characteristics*

310 Researchers have noted that different talker processing tasks vary in their sensitivity of detecting 311 differences between groups (if any exists, e.g., Levi, 2019; Perrachione, 2017). In this review, we 312 separated studies into one of three categories: talker discrimination, voice perception, and talker 313 identification. Note, however, that Carmel et al. (Carmel et al., 2011) included both a talker 314 discrimination and talker identification task, and Abdeltawwab et al. (Abdeltawwab et al., 2016) 315 included both a talker discrimination and voice perception task. These categories were informed 316 by the goals of the relevant experimental task. For example, if a study uses a discrimination task 317 with a continua of stimuli along an acoustic dimension, and the main finding is in regards to 318 sensitivity to acoustic cues, this falls under voice perception. In this section, we briefly introduce 319 the tasks that were used in the reported studies, as well as the types of stimuli used, to provide 320 context when discussing findings in later sections.

In *talker discrimination* tasks (n = 15; see Table 1), participants were assessed on their 321 322 ability to tell different voices apart. This includes both discrimination between and within genders. 323 One study assessed this indirectly by using a parental survey (Carmel et al., 2011), wherein parents 324 were asked if their child had difficulty discriminating between voices. Other laboratory studies 325 generally used a same/different task, in which participants heard two different stimuli and had to 326 respond whether these two tokens were spoken by the same person or by different people. A major 327 feature of this task is that listeners do not need to encode or recall information beyond what is 328 presented in a single trial.

In *voice perception* tasks (n = 21; Table 2), participants were assessed on their ability to perceive differences in vocal cues. This was perhaps the broadest grouping, but these studies all required listeners to perceive differences between talkers or between continua of manipulated

332 talkers. For the most part, these studies are concerned with how CI users perceive acoustic features 333 (F0 and VTL) that cue gender differences, although one study examines cues to vocal age (Skuk 334 et al., 2020). In many cases, these tasks used a two-alternative forced choice paradigm in which 335 participants heard a stimulus and had to respond whether the token was more likely to come from 336 a female or male talker. In one study, participants were asked to judge the "femaleness" or 337 "maleness" of the talker's voice using a rating scale (Meister et al., 2016). Several studies (El 338 Boghdady et al., 2019, 2021; Gaudrain & Baskent, 2018; Nogueira et al., 2021) used an adaptive 339 three-alternative forced choice task to measure the threshold at which participants could perceive 340 the difference between acoustic cues to vocal gender (i.e., just noticeable differences).

341 In *talker identification* tasks (n = 7; see Table 3), participants were assessed on their ability 342 to identify voices based on previous experience. These include both recognition of familiar or 343 trained voices and identification of voices into discrete categories. Two studies used parental 344 surveys to assess participants' talker recognition skills (Carmel et al., 2011; Morris et al., 2015). 345 Other lab experiments used anywhere from a 2- to 10-alternative forced choice task, typically 346 involving a training and a test phase. During training, participants learned different face-voice or 347 name-voice pairings, in which they received feedback regarding their accuracy. During the test 348 phase, participants were asked to identify the person who was speaking. In this cognitively more 349 demanding task, listeners have to encode the relevant features of voices from unfamiliar 350 individuals, hold these in memory, and recall them at a later test phase. Indeed, young children 351 tend to succeed in talker or gender discrimination tasks, but less so in unfamiliar talker identification tasks (Fecher et al., 2019). Given the differences across these tasks, we are careful 352 353 to label the task when discussing findings below.

| 354 | Across the different laboratory tasks, different types of stimuli are also used. The tokens |
|-----|---|
| 355 | themselves were real words (e.g., Consonant-Vowel-Consonant words; $n = 13$), or short sentences |
| 356 | (n = 19). Other studies used isolated vowels $(n = 2)$, non-words $(n = 6)$ or a combination of words |
| 357 | and sentences $(n = 1)$. The vast majority of these tokens were produced naturally and used as |
| 358 | stimuli in their natural form ($n = 29$), while other studies used synthetic or synthesized speech (n |
| 359 | = 12). For example, Barone et al. (2016) recorded two voices (one female and one male) producing |
| 360 | a word, and using a specialized speech manipulation software, they morphed these two voices to |
| 361 | a continuum of 11 voices, ranging from 100% female to 100% male and 9 gender-interpolated |
| 362 | voices in between. |

In the vast majority of cases, CI users' accuracy in talker discrimination, voice perception, and talker identification were above chance levels. There are certainly some studies showing individual differences in performance (to be discussed in more detail later). For example, Kovačić et al. (2010) found that a subset of their CI participants had above-chance performance at identifying the gender of voices, while another subset had below-chance performance. Nonetheless, the fact that most CI users are reaching above-chance levels in these variety of tasks indicates that some talker information is encoded through the CI device.

Some studies set out to examine whether certain forms of speech would affect talker processing. These studies found that CI users were better at picking up on indexical information when presented with neutral speech compared to whispered (Hazrati et al., 2015), speech presented through broadband (compared to presented through the telephone; Horng et al., 2007), and with original stimuli (compared to speech transformed with VTL processing algoritms, Wilkinson et al., 2013). These studies highlight how different real-life situations might impact talker processing for CI users.

377 *3.3. Comparison to individuals with typical hearing*

378 A large proportion of studies (n = 22) included comparisons between CI users and NH 379 listeners. When differences were detected between groups, results showed that CI users had more 380 difficulty in the tasks than NH participants. For example, compared to NH participants, CI users 381 had lower accuracy performance in discriminating between unfamiliar talkers' voices (Carmel et 382 al., 2011; Geers et al., 2013), in identifying the gender of voices (Kovačić & Balaban, 2009; 383 Meister et al., 2009) and in learning to identify voices (van Heugten et al., 2014; Vongpaisal et al., 384 2010). Differences in talker discrimination appear to be more pronounced when tasks used 385 speakers of the same gender (Muhler et al., 2009), or when using varied stimuli within trials 386 (Cleary et al., 2005; Cleary & Pisoni, 2002). CI users also had more difficulty identifying the 387 gender of voices, compared to NH participants (Massida et al., 2013). In some cases, the best 388 performing CI users overlapped in performance with NH listeners (Cleary et al., 2005; Fu et al., 2005). 389

Other studies found significant differences between groups in the manner by which listeners encoded vocal properties (Barone et al., 2016; Kovačić & Balaban, 2009; Meister et al., 2016; Zaltz et al., 2018). Another study examined whether top-down information, such as visual information, would influence listeners' gender rating of voices (Barone et al., 2016). Findings from this study show that, compared to NH participants, CI users were more affected by visual information when distinguishing between male and female voices.

396 Some studies compared CI users' performance to NH listeners' performance with CI 397 simulations (vocoded speech), rather than with unmanipulated speech (Sjoberg et al., 2017; 398 Stickney et al., 2004). In these cases, CI users performed comparably to NH listeners exposed to 399 stimuli that were filtered to a certain amount of channels or frequency bands. For example, Vongphoe et al. (2005) found that CI users performed comparably to NH listeners exposed to
stimuli filtered to 1-band through amplitude modulation, while Fu et al. (2005) presented NH
listeners with 4- or 8-bands, and van Heugten et al. (2014) with 24-bands.

403 *3.4. Comparison between device configurations*

404 An important question is whether the talker processing abilities of CI users are affected by 405 the configuration of their CI devices. For instance, some studies were interested in the potential 406 advantage of acoustic hearing through hearing aids (HA) to provide indexical cues for talker 407 processing. Cullington and Zeng did not find any significant differences between bimodal users 408 and bilateral CI users on either the identification of specific talkers or the identification of talkers 409 into categories (Cullington & Zeng, 2011, 2010). On the other hand, Hay-McCutcheon et al. (Hay-410 McCutcheon et al., 2018) found some benefit of acoustic hearing in a talker discrimination task, 411 as bimodal CI users outperformed unilateral CI users. Davidson et al. (2019) found that bimodal 412 children with longer HA use and better Pure Tone Averages (PTA) had a higher suprasegmental 413 speech perception score (a composite score that included performance on talker discrimination), 414 suggesting that prolonged acoustic experience is beneficial as long as hearing loss is not too severe.

415 Given that between-group comparisons may introduce some confounds, a better test of the 416 hypothesis that residual acoustic hearing may improve voice perception is if bimodal CI users 417 show improved voice perception abilities when listening to combined electric-acoustic stimuli 418 versus just receiving electric or acoustic input. Surprisingly, these studies did not find any significant difference in talker discrimination abilities across the three conditions (Davidson et al., 419 420 2015; Dorman et al., 2008). Zhang et al. (2012) administered auditory training to bimodal CI 421 participants, and also found no significant difference in magnitude of improvement in their gender 422 identification performance between electric-alone trials and electric-acoustic stimulation trials. In a talker discrimination task, Abdeltawwab et al. (Abdeltawwab et al., 2016) found an advantage
for bimodal stimulation compared to acoustic alone (i.e., HA only), but not compared to electrical
alone (i.e., CI only). Taken together, there appears to be minimal evidence that residual acoustic
hearing aids in voice perception.

427 Several studies compared the effects of different CI devices or processing strategies on 428 voice perception. Some studies found that talker processing abilities were not affected by electrode 429 configuration or speech coding strategies (Landwehr et al., 2014). Nevertheless, others have found 430 some key differences in performance based on which processing strategy CI participants were 431 using (i.e., processors from Cochlear Limited, Advanced Bionics, and Med El; Spahr et al., 2007; Spahr & Dorman, 2004, 2003). For instance, Fuller et al. (Fuller et al., 2014) found that users with 432 433 devices that had a higher stimulation rate were more likely to categorize stimuli as female across 434 a continuum of voices. Geers et al. (2013) found better performance for children who used the most recent CI processor in their study, and Dillier et al. (Dillier et al., 1994) found that the strategy 435 436 that preserved speech quality features and had a high continuous stimulation rate (High Spectral 437 Transmission) resulted in the best performance in a voice perception task.

438 *3.5. Role of different acoustic cues*

MH listeners take advantage of various cues, such as fundamental frequency (F0; related to the pitch of the voice) and the distribution of formant peaks (related to VTL and thus to the height of the talker) to encode talker identity (e.g., Smith & Patterson, 2005). Several studies have investigated how various acoustic cues might influence talker information processing in CI users, including contrasting F0 and VTL, and the spectral and temporal cues that comprise F0. In general, CI users have poorer access to fine acoustic detail, as evidenced by CI users having larger *Just*

Noticeable Differences for both F0 and VTL cues compared to NH listeners (Gaudrain & Başkent,
2018).

447 Given the nature of CI processing, it follows that CI users would adapt to rely more on F0 448 than on the complex VTL cues. Indeed, findings show that CI users rely more heavily on F0 than 449 VTL for voice gender perception compared to NH listeners, who strongly weigh both cues (Fuller 450 et al., 2014; Meister et al., 2009; Skuk et al., 2020). For instance, CI users are capable of 451 categorizing stimuli from an artificial F0 continuum similarly to NH listeners (Meister et al., 2009) and the discriminability of voices by CI users is correlated to the difference in F0 of the voices, 452 453 with larger differences in F0 being easier for CI users to discriminate (Muhler et al., 2009). Fu et al. (2005) found that CI users can take advantage of temporal periodicity cues to discriminate 454 455 gender when voices have distinct F0. However, if the F0 of the voices overlap, CI users were not 456 able to effectively discriminate between genders. It has been suggested that voices that differ by more than an octave, like would be found between a typical male and typical female voice, should 457 458 be discriminable by CI users, but differences of less than an octave results in ambiguity (see Moore 459 & Carlyon, 2005 for a review of pitch processing in CI users).

Accurate VTL perception requires perception of formant peaks which are more obscured by the limited spectral resolution of a CI. There is some evidence that CI users can use combined F0-VTL cues in sentence contexts (Meister et al., 2016), but it is clear from the limitations of CI processing that it is easier for CI users to access F0 information. For prelingually-deafened adult CI users, age at implantation was related to use of VTL cues, but not F0 (Zaltz et al., 2018). Those who were implanted earlier than age 4 had improved VTL discrimination, suggesting early exposure is important for taking advantage of VTL cues. A recent study (Skuk et al., 2020) found evidence that CI users do use timbre cues to perceive age differences in voices. This study used a combination of aperiodicity, spectrum level, and formant frequencies to represent timbre and found that high-performing CI users were able to use timbre and F0 to judge age in a similar patter to NH controls (although the controls still outperformed the CI users). The poorer performing CI users, however, showed no evidence of being able to use timbre and consistently judged gender differences as age differences (female voices were categorized younger, while male voices were categorized older).

In summary, CI users adapt their use of acoustic cues from the use of spectral (VTL) information to the greater reliance on F0 to compensate for the degradation of their input compared to NH listeners. Some CI users still take advantage of VTL cues, but this is predicted by early auditory experience. For the most part, CI users seem to rely on F0 as a more reliable cue to distinguishing voices.

479 *3.6. Relationship to hearing experience*

480 Individual differences in hearing history might also affect task performance. These 481 variables include age at onset of deafness, age of implantation, and duration of CI use - all of 482 which have been found to affect anatomical and neurophysiological properties of the auditory 483 system. As noted above, the majority of studies carried out were with either postlingually-deafened 484 adults or prelingually-deafened children, so disentangling the effects of onset of deafness from the 485 age of participants is difficult. Some studies used a mixed group with participants whose age at 486 onset of deafness ranged from early childhood through adulthood (Fu et al., 2004, 2005; Vongphoe & Zeng, 2005), but they did not investigate the role of age at onset of deafness on the performance 487 488 of participants.

25

For individuals who were deaf at a young age, much of the discussion on the effects of hearing experience on voice perception surrounds critical periods. There are two possible predictions. First, if early access to acoustic speech allows for the development of the auditory cortex, then we might expect that children with later onset of deafness might have improved voice perception. However, Kovačić and Balaban (2010) found no substantive effect of age at onset of deafness in the gender identification performance of children ages 5 to 18, with the caveat that cochlear implantation happens early enough (i.e., hearing loss is not left untreated).

496 Another hypothesis is that children who receive their CIs earlier may have improved voice 497 perception, especially if there is a sensitive period for voice perception. Indeed, children who were 498 implanted at a younger age have better talker discrimination (Geers et al., 2013) and children with 499 a shorter duration of deafness before implantation showed better gender identification (Kovačić & 500 Balaban, 2010). These studies have vastly different ranges of implantation age (1 - 3.2 years versus 501 2.1 - 15.3 years, respectively) which could explain the different findings with respect to age of 502 implantation and duration of deafness. A benefit of early implantation (before age 4) was also seen 503 in adult prelingually-deafened CI users' discrimination of voice cues (Zaltz et al., 2018). One study 504 only found weak effects, albeit in the expected direction (Cleary et al., 2005). Others did not find 505 any effect of age of implantation in their respective tasks (Cleary & Pisoni, 2002; Morris et al., 506 2015).

507 Beyond critical periods, some studies suggest that hearing experience in general 508 contributes to talker processing abilities of CI users. As mentioned previously, Davidson et al. 509 (2019) found that longer hearing aid use benefits suprasegmental perception (by their definition, a 510 composite score that includes performance on talker discrimination) in children with CIs as long 511 as their hearing loss is not too profound, providing evidence for the important role of early auditory 512 input for talker discrimination abilities. Massida et al. (2013) also found that the voice perception 513 skills of CI users can improve over time. Specifically, CI users had poor performance when they 514 were tested on the first month of CI use, but improved over time. Only one study specifically 515 looked at the effects of hearing experience on voice perception in older adult CI users, and they 516 found no relation between factors such as age, duration of deafness or duration of CI experience 517 and participants' performance on a gender identification task (Barone et al., 2016).

518 *3.7. Relationship to linguistic tasks*

519 Here, we investigate how performance in talker processing tasks might relate to their 520 performance in other linguistic and cognitive tasks, such as word recognition, vowel/consonant 521 perception, pitch/prosody perception, speech-on-speech masking, and other cognitive tasks.

522

3.7.1 Word recognition

523 An important question is whether talker processing abilities are related to overall word 524 recognition abilities in CI users. Some studies do not find a statistical relationship between 525 performance in a speech recognition task and a talker processing task. For example, Cullington 526 and Zeng (2011) found no relationship between talker identification performance and scores on a 527 standardized word recognition in noise task (Hearing-in-Noise Test). Similarly, Massida et al. 528 (2013) found no relationship between gender categorization and word recognition performance. These studies suggest that improvements in gender identification or talker identification are 529 530 independent of improvements in word recognition.

That said, there is some evidence that word recognition abilities are related to talker and voice cue discrimination. Sjoberg et al. (2017) found that talker discrimination was related to speech recognition abilities in adult CI users, and Cleary and Pisoni (2002) found the same relationship in children. Zaltz et al. (2018) found a moderate relationship between the ability to discriminate VTL cues and speech recognition in adult prelingually-deaf CI users. Davidson et al.
(2019) found that a composite suprasegmental perception score (that incorporated talker
discrimination ability) accounted for variation in both receptive vocabulary and language scores
in children. Specifically, the better a child could perceive suprasegmental information, the stronger
their receptive language and vocabulary skills would be (measured by the CELF-4 and PPVT-4
respectively). More work is required to investigate the cause of this relationship, but these studies
suggest that talker discrimination and word recognition skills are intertwined in speech processing.

542

3.7.2 Vowel and consonant perception

543 Vongphoe and Zeng (2005) did not find evidence for a relationship between talker and 544 vowel perception. VTL is closely tied to formant peaks, as are vowels, so it is perhaps surprising 545 that no relationship was found. However, VTL seems to be particularly difficult for CI users (at 546 least compared to F0) so perhaps that CI users seem to move away from relying on VTL is the 547 cause of the dissociation between these skills.

However, Li and Fu (2011) finds that, under certain conditions, vowel and consonant perception is correlated to voice gender discrimination. Particularly, recognizing speech sounds in noise was positively correlated to gender discrimination when talkers only differed by a small average F0 (10 Hz). This finding suggests that some CI users can pick up on subtle acoustic differences in the signal, and take advantage of this in different ways (i.e., recognizing a voice or recognizing phonemes).

554

3.7.3 Speech masking

555 Several studies investigated the relationship between speech-on-speech understanding and 556 sensitivity to vocal cues or talker identification (Cullington & Zeng, 2011; El Boghdady et al., 557 2019, 2021; Nogueira et al., 2021). In a speech-on-speech task, listeners must track a target voice

558 that is masked by a competing voice (or voices). If the target and masker voices are similar, the 559 task becomes more difficult. El Boghdady et al. (2019) found that, on average, CI users with better 560 access to vocal cues (i.e., smaller JNDs to both F0 and VTL) have better performance recognizing 561 single-talker masked speech. Cullington & Zeng (2011), on the other hand, found no correlation 562 between comprehension of single-talker masked sentences and talker identification ability. These 563 conflicting findings suggest that better perception of acoustic cues to voice can benefit speech comprehension in challenging conditions, but that more difficult talker-based tasks do not 564 565 necessarily relate to one another.

566 *3.7.4 Other abilities*

Several stand-alone studies have investigated various other abilities that might relate to 567 568 talker processing, including musical experience, cognitive abilities, and affective prosody 569 discrimination. Zaltz et al. (2018) found no relationship between scores on a variety of cognitive tasks (auditory working memory, visual attention, task switching, and nonverbal intelligence) and 570 571 ability to perceive vocal cues, suggesting that these cognitive functions are not related to 572 discriminability of acoustic features in experienced adult CI users. Cullington & Zeng (2011) 573 found no correlation between talker identification and a standardized assessment of music ability 574 (the Montreal Battery of Evaluation for Amusia). However, prior music experience and musical 575 pitch perception was positively correlated to talker discrimination in children with CIs (Sjoberg et al., 2017). With increasing music experience, CI users performed better at talker discrimination 576 577 and pitch perception tasks, suggesting that music training can improve pitch perception in CI users 578 which may then enhance processing of talker information (Sjoberg et al., 2017).

579 Cullington and Zeng (2011) also investigated the relationship between affective prosody 580 discrimination (that is, discrimination of prosody for emotions) and talker identification. There was a strong positive correlation between these two abilities, suggesting that there is at least some relationship between talker identification abilities and learned prosodic processing abilities, like emotion. More work is required to pin down the nature of these relationships, both to further investigate the role of broader cognitive abilities on talker processing and to confirm whether musical training can improve talker perception.

586

4. Discussion

As a group, CI users performed above chance in talker discrimination and identification tasks, suggesting that they are fairly capable processors of talker information. However, there are clear differences in performance when compared to individuals with normal hearing. Compared to NH participants, CI users are less accurate in these talker processing tasks, are less sensitive to the acoustic cues to voice (F0 and VTL), and they use different cues to encode and retrieve indexical properties of voices. These differences highlight the challenges and adaptability of CI users in processing others' voices via electric stimulation.

594 One clear finding from this review is that there are wide individual differences in 595 performance among CI users. A few studies report that a subset of CI users reach the same level 596 of performance in talker processing tasks as NH participants. Some factors appear to correlate with 597 heightened performance in talker processing tasks. For example, there appears to be some benefit 598 of maintaining residual acoustic hearing for processing talker information. Particularly, bimodal 599 CI users (who have a contralateral HA) tended to perform better at talker discrimination tasks. 600 Additionally, several hearing factors - such as age of onset of deafness, age of implantation, and 601 duration of hearing aid use - is related to talker processing abilities. These findings suggest that 602 early intervention and maintaining acoustic input may be beneficial for talker processing.

603 Across the studies evaluated here, discrimination abilities relate more clearly to other 604 factors. Particularly, talker discrimination abilities were tied to pitch perception, music experience, 605 and speech recognition abilities in both adults and children (Cleary & Pisoni, 2002; Sjoberg et al., 606 2017), but the same relationships were not apparent for talker identification abilities. The disparity 607 in findings suggest that identification and discrimination might draw on different perceptual skills. 608 Indeed, it has been suggested that performance in discrimination paradigms rely more on low-level processing of acoustics (see Perrachione, 2017). If such is the case, it may be that this underlying 609 610 low-level mechanism ties CI users' performance in these different tasks. On the other hand, 611 performance in talker identification tasks has been described as being a better representation of the 612 psychological processes that contribute to voice recognition. It is therefore not surprising that the 613 one significant relationship related to performance in talker identification was with the ability to 614 discriminate affective prosody (Cullington & Zeng, 2011). Further work on this issue in both CI 615 and NH populations would help us better understand the mechanisms that contribute to ecological 616 voice recognition behaviours.

617 *4.1. Limitations*

618 There are several limitations in summarizing this body of literature that makes it difficult 619 to come to firm conclusions. First, many studies reported here fell prey to small samples sizes, 620 which make generalizing their findings difficult. Indeed, the median number of participants for 621 both CI and NH participants in this review was 15. This is an issue that plagues much of clinical 622 language research, as recruiting patients to participate in research can be challenging. In some cases, a small sample size is not a concern if the effect size is large enough for the study to remain 623 624 well-powered. With a population as variable in outcomes as CI users, larger samples will become 625 necessary when attempting to understand this variability, especially when it comes to investigating

interactions between talker processing and broader language or cognitive functioning. In order to
predict outcomes for CI users, individual differences approaches require larger samples. The goal
of this review was to amalgamate evidence across studies in an attempt to paint a coherent picture
of talker processing in CI users.

630 Second, comparisons between CI and NH participants are often confounded with age. As 631 is typical within the CI literature, the CI participants recruited in the studies listed here tended to 632 be older and included wider age ranges than groups of NH participants (see Figure 2). When direct 633 comparisons are involved, the ages of CI adult participants (M = 56.1 years) were often older than 634 the NH adult participants (M = 33.1 years). Direct comparisons for child participants were more equivalent, the mean ages of child CI participants was 9.5 years compared to 7.4 years for NH 635 636 child participants. The imbalance in adult comparisons is, in large part, due to the difficulties of 637 recruiting CI participants, as well as in recruiting age-matched participants with normal hearing. 638 Findings from studies with imbalanced ages may be difficult to interpret as there is some evidence 639 that NH older adults perform poorer in talker identification tasks compared to NH younger adults 640 (Best et al., 2018). In other words, it is unclear whether differences in performance between groups 641 are due to age or hearing configuration. When possible, future studies should be more careful about 642 selecting the age ranges of participants.

A related recommendation for researchers is to provide more detailed age and other demographic characteristics of their participants. Four of the studies in this review did not explicitly indicate any age information. Some studies indicate wide age ranges without indicating the mean age, while others indicate the mean age without providing the standard deviation or age range. In addition, several studies provide detailed age information about their participants with 650 Lastly, as is apparent from the summary tables in this paper, there is a wide variety of tasks 651 used to assess talker processing. The variability in tasks and outcome measures prohibits the use 652 of a meta-analysis to summarize this work. Nevertheless, as indicated above, these task differences 653 also provide answers to slightly different questions. Discrimination tasks address the differences 654 in talkers that CI users are able to perceive, while identification tasks assess CI users' ability to 655 explicitly label talkers. The difference in perspective between these tasks provides complementary 656 information about talker processing. On one hand, discrimination tasks shed light on the acoustic information that CI users can perceive, while identification tasks address the information that users 657 658 actually take advantage of to successfully label talkers.

659 Similarly, this review is limited in the scope of included tasks. We did not include studies 660 whose primary goal was investigating other aspects of talker processing, including speech masking 661 and talker familiarity (although some talker identification studies with training phases may touch 662 on familiarity). Speech masking represents an increased challenge to talker processing which 663 incorporates issues from speech in noise processing, speech streaming, and discriminating 664 competing auditory sources. An extended review of CI performance on these tasks would provide 665 additional framing for how CI users process talker information. Indeed, to successfully recognize 666 a masked talker, one must be able to attend to the target speech stream.

667 *4.2. Gaps in research*

668 This systematic review highlights several gaps in research on indexical processing in CI 669 users. First, while most age ranges are well represented in the literature, there remain gaps of 670 knowledge through the age ranges. There are very few studies examining the talker learning skills 671 of young children, and there is a lack of longitudinal studies within the literature. The studies 672 described in this review provide a solid foundation for understanding the talker learning skills of 673 children with CIs, but they also motivate examining the developmental trajectory of talker learning 674 skills at younger ages. Indeed, prior research has shown that NH infants can recognize highly 675 familiar voices (DeCasper & Fifer, 1980), and that infants can encode and retain unfamiliar voices 676 by 8-months of age (Orena & Werker, 2020). Nonetheless, as with speech perception, the 677 development of voice perception is a gradual process that continues to mature through childhood 678 and adolescence (Creel & Jimenez, 2012; Nagels et al., 2020, 2021; Rigler et al., 2015). Identifying 679 the time points in which young children with CIs become sensitive to indexical information could 680 highlight the plasticity available to child CI users and would be beneficial for clinicians and 681 caregivers of young CI patients.

682 In a similar vein, there is also a lack of research directly comparing prelingually- and 683 postlingually-implanted CI users. This comparison could provide insight into the different 684 strategies that CI users employ based on their language and acoustic hearing experience. For 685 instance, postlingually-deafened CI users have mental representations of pitch from when they 686 were able to hear acoustic sounds; thus, they may be able to use their memory of acoustic sounds 687 to process pitch through an implant. Prelingual users, on the other hand, have little to no acoustic 688 language experience to draw from, and learn strategies that are informed entirely by input from 689 their implant. Nonetheless, there are a host of confounding variables that could arise when 690 comparing pre- and postlingual CI users. Age, for instance, is likely to be unbalanced across the 691 two groups, as prelingual users will likely be younger than postlingual users.

692Towards the same goal of improving outcomes for users, further investigation into the693individual differences that predict processing of indexical information could shed light on factors

694 that could be targeted for improvement or enhanced through training. For example, there are mixed 695 results as to the impact of improving talker processing for speech outcomes for CI users, so an 696 experimental intervention study examining this specifically could clarify the issue. There is a 697 wealth of evidence that NH listeners use linguistic information for talker processing (for review, 698 Creel & Bregman, 2011), but that same work is lacking with CI users. The work that does 699 investigate the relationship between linguistic and indexical processing in CI users is correlational 700 in nature and offers inconclusive findings. An experimental study manipulating talker 701 identification in CI users' native compared to non-native language could begin to fill this gap.

702 We have also highlighted the importance of residual acoustic hearing, but additional 703 demographic and cognitive factors could prove important. Only one study has investigated the role 704 of executive function on voice perception abilities (Zaltz et al., 2018), leaving room for further 705 investigation. Additionally, an examination of the psychophysical skills that relate to talker 706 perception could address whether improvements need to be made to the quality input to improve 707 talker processing. For child CI users, one might also look at how the variety of input received (e.g., 708 how many speakers children interact with) impacts individuals' ability to process indexical 709 information.

710 4.3. Recommendations

Several studies offer suggestions for improvements to CI processing that could directly benefit talker processing. Broadly, these recommendations include hopes for better algorithms and coding schemes or better optimization strategies for coding algorithms for both children and adults (Geers et al., 2013; Kovačić & Balaban, 2009). For instance, the inclusion of temporal fine structure in CI processors has the potential to improve performance of CI users not just in talker processing, but across a variety of auditory tasks, including speech recognition in noise and

| 717 | perception of music and lexical tone (Stickney et al., 2004). Additionally, VTL perception has |
|-----|---|
| 718 | been singled out as a limitation imposed by CI technology (Zaltz et al., 2018). Better coding |
| 719 | strategies or fitting algorithms that address the shortcomings of VTL discrimination (be it the |
| 720 | inability to discriminate VTL or to use it as a cue for talker-size discrimination) could improve |
| 721 | talker gender identification for CI users (Fuller et al., 2014). Gaudrain and Başkent (2018) suggests |
| 722 | using VTL discrimination as a potential clinical tool for improving CI fit. It could be used as a |
| 723 | measure of spectral resolution that is more related to speech perception than spectral ripple tasks. |
| 724 | Developing a clinical tool with stronger ties to speech perception could have large implications for |
| 725 | CI outcomes related to indexical processing. |

Aside from improvements to the CI processor itself, there is also room for improvement in 726 727 rehabilitation strategies. There is an accumulation of evidence, also supported in this review, that 728 electric-acoustic stimulation (EAS) improves performance for CI users. However, not all CI users 729 have access to residual acoustic hearing. There is some work suggesting that altering the pitch to 730 an audible range for CI users might be beneficial for individuals with limited residual acoustic 731 hearing (Brown et al., 2016), although this work has so far been done with simulations and has yet 732 to be tested with actual CI users. Even further, Huang and colleagues (2017) provide evidence that 733 converting the fundamental frequency of a voice into tactile vibrations (electro-tactile stimulation) 734 improves speech reception thresholds in CI users at a similar magnitude to the benefit seen with 735 EAS. If encoding of talker information cannot be improved at the level of the CI device, then 736 perhaps turning to a different modality (i.e., tactile) to improve new CI users' adaptation to electric input is a worthwhile avenue of investigation. 737

738 *4.5 Conclusion*

| 739 | The goal of this review was to provide a comprehensive overview of talker processing in |
|-----|---|
| 740 | CI users. We found that CI users are able to perceive indexical information, but generally have |
| 741 | more challenges than their NH peers. Several factors were found to relate to perception of talker |
| 742 | information, including residual acoustic hearing and task demands. Future studies could build on |
| 743 | this existing work by investigating training paradigms that might improve long-term talker |
| 744 | processing abilities. Improved talker processing abilities may then impact other speech outcomes. |
| 745 | In the long term, further improvement to devices themselves is crucial to improving talker |
| 746 | processing outcomes for CI users. |

747

748 Acknowledgements

S. Colby is supported by NIH grant P50 000242 awarded to B. Gantz and B. McMurray, and A.
Orena is supported by a Fonds de Recherche Québec Nature et Technologies (FRQNT)
Postdoctoral Fellowship. We thank K. Chen, C. Shum, M. Huffman, and C. Morales for their
assistance in coding the studies. We thank B. McMurray for his comments on an earlier draft of
this paper.

| 754 | References |
|-----|--|
| 755 | Abdeltawwab, M. M., Khater, A., & El-Anwar, M. W. (2016). Contralateral bimodal |
| 756 | stimulation: A way to enhance speech performance in Arabic-speaking cochlear implant |
| 757 | patients. ORL, 78(3), 126-135. https://doi.org/10.1159/000381024 |
| 758 | Apfelbaum, K. S., Bullock-Rest, N., Rhone, A. E., Jongman, A., & McMurray, B. (2014). |
| 759 | Contingent categorisation in speech perception. Language, Cognition and Neuroscience, |
| 760 | 29(9), 1070–1082. https://doi.org/10.1080/01690965.2013.824995 |
| 761 | Barone, P., Chambaudie, L., Strelnikov, K., Fraysse, B., Marx, M., Belin, P., & Deguine, O. |
| 762 | (2016). Crossmodal interactions during non-linguistic auditory processing in cochlear- |
| 763 | implanted deaf patients. Cortex, 83, 259-270. https://doi.org/10.1016/j.cortex.2016.08.005 |
| 764 | Baskent, D., Gaudrain, E., Tamati, T. N., & Wagner, A. (2016). Perception and Psychoacoustics |
| 765 | of Speech in Cochlear Implant Users. In A. T. Cacace, E. de Kleine, A. G. Holt, & P. van |
| 766 | Dijk (Eds.), Scientific Foundations of Audiology. Perspectives from Physics, Biology, |
| 767 | Modelling, and Medicine (pp. 285–319). Plural Publishing Inc. |
| 768 | https://books.google.de/books?hl=de&lr=&id=EtAyDAAAQBAJ&oi=fnd&pg=PA285&dq |
| 769 | $= Scientific + Foundations + of + Audiology \&ots = cfEfTicv7h \&sig = 1 cTQmXsc_FR7oNQiwWith the second state of the second s$ |
| 770 | YgpklOkN0 |
| 771 | Belzner, K. A., & Seal, B. C. (2009). Children with cochlear implants: A review of |
| 772 | demographics and communication outcomes. American Annals of the Deaf, 154(3), 311- |
| 773 | 334. |
| 774 | Berland, A., Collett, E., Gaillard, P., Guidetti, M., Strelnikov, K., Cochard, N., Barone, P., & |
| 775 | Deguine, O. (2019). Categorization of everyday sounds by cochlear implanted children. |
| 776 | Scientific Reports, 9(1), 1-16. https://doi.org/10.1038/s41598-019-39991-9 |

| 777 | Best, V., Ahlstrom, J. B., Mason, C. R., Roverud, E., Perrachione, T. K., Kidd, G., & Dubno, J. |
|-----|--|
| 778 | R. (2018). Talker identification: Effects of masking, hearing loss, and age. The Journal of |
| 779 | the Acoustical Society of America, 143(2), 1085–1092. https://doi.org/10.1121/1.5024333 |
| 780 | Blamey, P., Artieres, F., Başkent, D., Bergeron, F., Beynon, A., Burke, E., Dillier, N., Dowell, |
| 781 | R., Fraysse, B., Gallégo, S., Govaerts, P. J., Green, K., Huber, A. M., Kleine-Punte, A., |
| 782 | Maat, B., Marx, M., Mawman, D., Mosnier, I., O'Connor, A. F., Lazard, D. S. (2012). |
| 783 | Factors affecting auditory performance of postlinguistically deaf adults using cochlear |
| 784 | implants: An update with 2251 patients. Audiology and Neurotology, 18(1), 36-47. |
| 785 | https://doi.org/10.1159/000343189 |
| 786 | Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. Cognition, |
| 787 | 106(2), 707-729. https://doi.org/10.1016/j.cognition.2007.04.005 |
| 788 | Bregman, M. R., & Creel, S. C. (2014). Gradient language dominance affects talker learning. |
| 789 | Cognition, 130(1), 85-95. https://doi.org/10.1016/j.cognition.2013.09.010 |
| 790 | Brown, C. A., Helms Tillery, K., Apoux, F., Doyle, N. M., & Bacon, S. P. (2016). Shifting |
| 791 | Fundamental Frequency in Simulated Electric-Acoustic Listening. Ear & Hearing, 37(1), |
| 792 | e18-e25. https://doi.org/10.1097/AUD.00000000000227 |
| 793 | Carmel, E., Kronenberg, J., Wolf, M., & Migirov, L. (2011). Telephone use among cochlear |
| 794 | implanted children. Acta Oto-Laryngologica, 131(2), 156-160. |
| 795 | https://doi.org/10.3109/00016489.2010.517784 |
| | |

- Ciocca, V., Francis, A. L., Aisha, R., & Wong, L. (2002). The perception of Cantonese lexical
 tones by early-deafened cochlear implantees. *The Journal of the Acoustical Society of America*, 111(5), 2250. https://doi.org/10.1121/1.1471897
- 799 Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. The

- 300 *Journal of the Acoustical Society of America*, *116*(6), 3647–3658.
- 801 https://doi.org/10.1121/1.1815131
- 802 Cleary, M., & Pisoni, D. B. (2002). Talker discrimination by prelingually deaf children with
- 803 cochlear implants: Preliminary results. Annals of Otology, Rhinology and Laryngology,
- 804 *111*(5 II), 113–118.
- 805 http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L3452231
- 806 6%5Cnhttp://sfx.library.uu.nl/utrecht?sid=EMBASE&issn=00034894&id=doi:&atitle=Talk
- 807 er+discrimination+by+prelingually+deaf+children+with+cochlear+implants%3A+Prelimin
- 808 ary+result
- 809 Cleary, M., Pisoni, D. B., & Kirk, K. I. (2005). Influence of Voice Similarity on Talker
- 810 Discrimination in Children With Normal Hearing and Children With Cochlear Implants.
- 811 *Journal of Speech, Language, and Hearing Research, 48*(1), 204–223.
- 812 https://doi.org/10.1044/1092-4388(2005/015)
- 813 Creel, S. C., & Bregman, M. R. (2011). How Talker Identity Relates to Language Processing.
- 814 *Linguistics and Language Compass*, 5(5), 190–204. https://doi.org/10.1111/j.1749-
- 815 818X.2011.00276.x
- 816 Creel, S. C., & Jimenez, S. R. (2012). Differences in talker recognition by preschoolers and

adults. *Journal of Experimental Child Psychology*, *113*(4), 487–509.

- 818 https://doi.org/10.1016/j.jecp.2012.07.007
- 819 Cullington, H. E., & Zeng, F.-G. (2011). Comparison of Bimodal and Bilateral Cochlear Implant
- 820 Users on Speech Recognition with Competing Talker, Music Perception, Affective Prosody
- 821 Discrimination, and Talker Identification. *Ear and Hearing*, *32*(1).
- 822 https://doi.org/10.1097/AUD.0b013e3181edfbd2

- Cullington, H. E., & Zeng, F. G. (2010). Comparison of bimodal and bilateral cochlear implant
 users. *Cochlear Implants International*, *11 Suppl 1*, 67–74.
- 825 https://doi.org/10.1179/146701010X12671177440262
- 826 Davidson, L. S., Firszt, J. B., Brenner, C., & Cadieux, J. H. (2015). Evaluation of Hearing Aid
- Frequency Response Fittings in Pediatric and Young Adult Bimodal Recipients. *Journal of the American Academy of Audiology*, 26(4), 393–407. https://doi.org/10.3766/jaaa.26.4.7
- 829 Davidson, L. S., Geers, A. E., Uchanski, R. M., & Firszt, J. B. (2019). Effects of Early Acoustic
- 830 Hearing on Speech Perception and Language for Pediatric Cochlear Implant Recipients.
- *Journal of Speech, Language, and Hearing Research, 62*(9), 3620–3637.
- 832 https://doi.org/10.1044/2019_JSLHR-H-18-0255
- B33 DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers'
 voices. *Science*, 208(4448), 1174–1176.
- Billier, N., Lai, W., & Bögli, H. (1994). A high spectral transmission coding strategy for a multielectrode cochlear implant. *Advances in Cochlear Implants*, 152–157.
- 837 Dorman, M. F., Gifford, R. H., Spahr, A. J., & McKarns, S. A. (2008). The benefits of
- combining acoustic and electric stimulation for the recognition of speech, voice and
 melodies. *Audiology and Neurotology*, *13*(2), 105–112. https://doi.org/10.1159/000111782
- 840 El Boghdady, N., Gaudrain, E., & Başkent, D. (2019). Does good perception of vocal
- 841 characteristics relate to better speech-on-speech intelligibility for cochlear implant users?
- 842 *The Journal of the Acoustical Society of America*, *145*(1), 417–439.
- 843 https://doi.org/10.1121/1.5087693
- El Boghdady, N., Langner, F., Gaudrain, E., Başkent, D., & Nogueira, W. (2021). Effect of
- 845 Spectral Contrast Enhancement on Speech-on-Speech Intelligibility and Voice Cue

- 846 Sensitivity in Cochlear Implant Users. *Ear and Hearing*, 271–289.
- 847 https://doi.org/10.1097/AUD.00000000000936
- 848 Fecher, N., Paquette-Smith, M., & Johnson, E. K. (2019). Resolving the (Apparent) Talker
- 849 Recognition Paradox in Developmental Speech Perception. *Infancy*, 24(4), 570–588.
- 850 https://doi.org/10.1111/infa.12290
- Fu, Q.-J., Chinchilla, S., & Galvin, J. J. (2004). The role of spectral and temporal cues in voice
- 852 gender discrimination by normal-hearing listeners and cochlear implant users. JARO -
- *Journal of the Association for Research in Otolaryngology*, *5*(3), 253–260.
- 854 https://doi.org/10.1007/s10162-004-4046-1
- Fu, Q.-J., Chinchilla, S., Nogaki, G., & Galvin, J. J. (2005). Voice gender identification by
- cochlear implant users: The role of spectral and temporal resolution. *The Journal of the Acoustical Society of America*, *118*(3), 1711–1718. https://doi.org/10.1121/1.1985024
- Fuller, C. D., Gaudrain, E., Clarke, J. N., Galvin, J. J., Fu, Q. J., Free, R. H., & Başkent, D.
- 859 (2014). Gender Categorization Is Abnormal in Cochlear Implant Users. *JARO Journal of*
- the Association for Research in Otolaryngology, 15(6), 1037–1048.
- 861 https://doi.org/10.1007/s10162-014-0483-7
- 862 Gaudrain, E., & Başkent, D. (2015). Factors limiting vocal-tract length discrimination in
- 863 cochlear implant simulations. *The Journal of the Acoustical Society of America*, 137(3),
- 864 1298–1308. https://doi.org/10.1121/1.4908235
- 865 Gaudrain, E., & Başkent, D. (2018). Discrimination of voice pitch and vocal-tract length in
- 866 cochlear implant users. *Ear and Hearing*, *39*(2), 226–237.
- 867 https://doi.org/10.1097/AUD.00000000000480
- 868 Geers, A. E., Davidson, L. S., Uchanski, R. M., & Nicholas, J. G. (2013). Interdependence of

- 869 Linguistic and Indexical Speech Perception Skills in School-Age Children With Early
- 870 Cochlear Implantation. *Ear and Hearing*, *34*(5), 562–574.
- 871 https://doi.org/10.1097/AUD.0b013e31828d2bd6
- 872 Harzing, A. W. (2007). *Publish or Perish* (7.15.2643.7260).
- 873 https://harzing.com/resources/publish-or-perish
- Hay-McCutcheon, M. J., Peterson, N. R., Pisoni, D. B., Iler, K., Yang, X., Parton, J., Kirk, K. I.,
- 875 Yang, X., & Parton, J. (2018). Performance variability on perceptual discrimination tasks in
- 876 profoundly deaf adults with cochlear implants. *Journal of Communication Disorders*,

877 72(January 2017), 122–135. https://doi.org/10.1016/j.jcomdis.2018.01.005

- Hazrati, O., Ali, H., Hansen, J. H. L., & Tobey, E. (2015). Evaluation and analysis of whispered
 speech for cochlear implant users: Gender identification and intelligibility. *The Journal of the Acoustical Society of America*, *138*(1), 74–79. https://doi.org/10.1121/1.4922230
- Higgins, J. P. T., & Green, S. (2006). Highly sensitive search strategies for identifying reports of
- 882 randomized controlled trials in MEDLINE. In *Cochrane Handbook for Systematic Reviews*
- 883 of Interventions 4.2.6 (p. Appendix 5b). www.cochrane.org/resources/handbook/hbook.htm
- Hopkins, K., & Moore, B. C. J. (2009). The contribution of temporal fine structure to the
- 885 intelligibility of speech in steady and modulated noise. *The Journal of the Acoustical*886 Society of America, 125(1), 442–446. https://doi.org/10.1121/1.3037233
- Horng, M., Chen, H., Hsu, C., & Fu, Q. (2007). Telephone Speech Perception by MandarinSpeaking Cochlear Implantees. *Ear & Hearing*, *28*, 66S-69S.
- Huang, J., Sheffield, B., Lin, P., & Zeng, F. G. (2017). Electro-tactile stimulation enhances
- 890 cochlear implant speech recognition in noise. *Scientific Reports*, 7(1), 1–5.
- 891 https://doi.org/10.1038/s41598-017-02429-1

Johnson, K., Strand, E. A., & D'Imperio, M. (1999). Auditory-visual integration of talker gender
in vowel perception. *Journal of Phonetics*, 27(4), 359–384.

894 https://doi.org/10.1006/jpho.1999.0100

- 895 Kadam, M. A., Orena, A. J., Theodore, R. M., & Polka, L. (2016). Reading ability influences
- native and non-native voice recognition, even for unimpaired readers. *The Journal of the*
- 897 *Acoustical Society of America*, *139*(1), EL6–EL12. https://doi.org/10.1121/1.4937488
- Kovačić, D., & Balaban, E. (2009). Voice gender perception by cochlear implantees. *The Journal of the Acoustical Society of America*, *126*(2), 762–775.
- 900 https://doi.org/10.1121/1.3158855
- Wovačić, D., & Balaban, E. (2010). Hearing History Influences Voice Gender Perceptual
 Performance in Cochlear Implant Users. *Ear & Hearing*, *31*, 806–814.
- 903 Krull, V., Luo, X., & Iler Kirk, K. (2012). Talker-identification training using simulations of
- binaurally combined electric and acoustic hearing: Generalization to speech and emotion
- 905 recognition. *The Journal of the Acoustical Society of America*, *131*(4), 3069–3078.
- 906 https://doi.org/10.1121/1.3688533
- 907 Landwehr, M., Fürstenberg, D., Walger, M., von Wedel, H., & Meister, H. (2014). Effects of
- 908 various electrode configurations on music perception, intonation and speaker gender
- 909 identification. *Cochlear Implants International*, 15(1), 27–35.
- 910 https://doi.org/10.1179/1754762813Y.0000000037
- 911 Levi, S. V. (2019). Methodological considerations for interpreting the Language Familiarity
- 912 Effect in talker processing. *Wiley Interdisciplinary Reviews: Cognitive Science*, 10(2), 1–15.
- 913 https://doi.org/10.1002/wcs.1483
- Li, T., & Fu, Q. J. (2011). Voice gender discrimination provides a measure of more than pitch-

- 915 related perception in cochlear implant users. *International Journal of Audiology*, 50(8),
- 916 498–502. https://doi.org/10.3109/14992027.2011.576274
- 917 Manrique, M., Cervera-Paz, F. J., Huarte, A., & Molina, M. (2004). Advantages of cochlear
- 918 implantation in prelingual deaf children before 2 years of age when compared with later
- 919 implantation. *Laryngoscope*, *114*(8 I), 1462–1469. https://doi.org/10.1097/00005537920 200408000-00027
- 921 Massida, Z., Belin, P., James, C., Rouger, J., Fraysse, B., Barone, P., & Deguine, O. (2011).
- 922 Voice discrimination in cochlear-implanted deaf subjects. *Hearing Research*, 275(1–2),

923 120–129. https://doi.org/10.1016/j.heares.2010.12.010

Massida, Z., Marx, M., Belin, P., James, C., Fraysse, B., Barone, P., & Deguine, O. (2013).

925 Gender Categorization in Cochlear Implant Users. *Journal of Speech, Language, and*

926 *Hearing Research*, 56(5), 1389–1401. https://doi.org/10.1044/1092-4388(2013/12-0132)

927 Meister, H., Fürsen, K., Streicher, B., Lang-Roth, R., & Walger, M. (2016). The Use of Voice

- 928 Cues for Speaker Gender Recognition in Cochlear Implant Recipients. *Journal of Speech*,
- 929 Language, and Hearing Research, 59(3), 546–556. https://doi.org/10.1044/2015_JSLHR-H-
- 930 15-0128
- 931 Meister, H., Landwehr, M., Pyschny, V., Walger, M., & Wedel, H. von. (2009). The perception

932 of prosody and speaker gender in normal-hearing listeners and cochlear implant recipients.

933 International Journal of Audiology, 48(1), 38–48.

934 https://doi.org/10.1080/14992020802293539

- 935 Moore, B. C. J., & Carlyon, R. P. (2005). Perception of Pitch by People with Cochlear Hearing
- Users and by Cochlear Implant Users. In C. J. Plack, R. R. Fay, A. J. Oxenham, & A. N.
- 937 Popper (Eds.), *Pitch. Springer Handbook of Auditory Research, vol 24* (pp. 234–277).

- 938 Springer. https://doi.org/10.1007/0-387-28958-5_7
- 939 Morris, D. J., Christiansen, L., Uglebjerg, C., Brännström, K. J., & Falkenberg, E. (2015).
- 940 Parental comparison of the prosodic and paralinguistic ability of children with cochlear
- 941 implants and their normal hearing siblings. *Clinical Linguistics & Phonetics*, 29(11), 840–
- 942 851. https://doi.org/10.3109/02699206.2015.1055803
- 943 Muhler, R., Ziese, M., Rostalski, D., Mühler, R., Ziese, M., Rostalski, D., Muhler, R., Ziese, M.,
- 944 & Rostalski, D. (2009). Development of a Speaker Discrimination Test for Cochlear
- 945 Implant Users Based on the Oldenburg Logatome Corpus. *ORL*, 71(1), 14–20.
- 946 https://doi.org/10.1159/000165170
- Nagels, L., Gaudrain, E., Vickers, D., Hendriks, P., & Başkent, D. (2020). Development of voice
 perception is dissociated across gender cues in school-age children. *Scientific Reports*,
- 949 10(1), 1–11. https://doi.org/10.1038/s41598-020-61732-6
- 950 Nagels, L., Gaudrain, E., Vickers, D., Hendriks, P., & Başkent, D. (2021). School-age children
- 951 benefit from voice gender cue differences for the perception of speech in competing speech.
- 952 The Journal of the Acoustical Society of America, 149(5), 3328–3344.
- 953 https://doi.org/10.1121/10.0004791
- 954 Nogueira, W., Boghdady, N. El, Langner, F., Gaudrain, E., & Başkent, D. (2021). Effect of
- 955 Channel Interaction on Vocal Cue Perception in Cochlear Implant Users. *Trends in*
- 956 *Hearing*, 25. https://doi.org/10.1177/23312165211030166
- 957 Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception*958 *and Psychophysics*, 60(3), 355–376. https://doi.org/10.3758/BF03206860
- 959 Orena, A. J., & Werker, J. (2020). Infants use disambiguation to learn new voices. 1–26.
- 960 https://doi.org/10.31234/osf.io/czkhx

- Perrachione, T. K. (2017). Speaker recognition across languages. *The Oxford Handbook of Voice Perception*. https://open.bu.edu/handle/2144/23877
- 963 Perrachione, T. K., Del Tufo, S. N., & Gabrieli, J. D. E. (2011). Human voice recognition
- depends on language ability. *Science*, *333*(6042), 595.
- 965 https://doi.org/10.1126/science.1207327
- 966 Perry, T. L., Ohde, R. N., & Ashmead, D. H. (2001). The acoustic bases for gender identification
- 967 from children's voices. The Journal of the Acoustical Society of America, 109(6), 2988–
- 968 2998. https://doi.org/10.1121/1.1370525
- 969 Peterson, G. E., & Barney, H. L. (1952). Control Methods Used in a Study of the Vowels.
- 970 *Journal of the Acoustical Society of America*, *24*(2), 175–184.
- 971 https://doi.org/10.1121/1.1906875
- 972 Peterson, N. R., Pisoni, D. B., & Miyamoto, R. T. (2010). Cochlear implants and spoken
- 973 language processing abilities: Review and assessment of the literature. *Restorative*
- 974 *Neurology and Neuroscience*, *28*(2), 237–250. https://doi.org/10.3233/RNN-2010-0535
- 975 Rigler, H., Farris-Trimble, A., Greiner, L., Walker, J., Tomblin, J. B., & McMurray, B. (2015).
- 976 The slow developmental time course of real-time spoken word recognition. *Developmental*977 *Psychology*, *51*(12), 1690–1703. https://doi.org/10.1037/dev0000044
- 8 Roberts, D. S., Lin, H. W., Herrmann, B. S., & Lee, D. J. (2013). Differential cochlear implant
 979 outcomes in older adults. *Laryngoscope*, *123*(8), 1952–1956.
- 980 https://doi.org/10.1002/lary.23676
- 981 Sidtis, D., & Kreiman, J. (2012). In the Beginning Was the Familiar Voice: Personally Familiar
- 982 Voices in the Evolutionary and Contemporary Biology of Communication. *Integrative*
- 983 Psychological and Behavioral Science, 46(2), 146–159. https://doi.org/10.1007/s12124-

984 011-9177-4

- 985 Sjoberg, K. M., Driscoll, V. D., Gfeller, K., Welhaven, A. E., Kirk, K. I., & Prusick, L. (2017).
- 986 The impact of electric hearing on children's timbre and pitch perception and talker
- 987 discrimination. *Cochlear Implants International*, *18*(1), 36–48.
- 988 https://doi.org/10.1080/14670100.2016.1263406
- 989 Skuk, V. G., Kirchen, L., Oberhoffner, T., Guntinas-Lichius, O., Dobel, C., & Schweinberger, S.
- 990 R. (2020). Parameter-Specific Morphing Reveals Contributions of Timbre and Fundamental
- 991 Frequency Cues to the Perception of Voice Gender and Age in Cochlear Implant Users.
- *Journal of Speech, Language, and Hearing Research, 63*(9), 3155–3175.
- 993 https://doi.org/10.1044/2020_JSLHR-20-00026
- 994 Skuk, V. G., & Schweinberger, S. R. (2014). Influences of fundamental frequency, formant
- frequencies, aperiodicity, and spectrum level on the perception of voice gender. *Journal of*
- 996 *Speech, Language, and Hearing Research*, 57(1), 285–296. https://doi.org/10.1044/1092-
- 997 4388(2013/12-0314)
- 998 Smith, D. R., & Patterson, R. D. (2005). The interaction of glottal-pulse rate and vocal-tract
- length in judgements of speaker size, sex, and age. *The Journal of the Acoustical Society of America*, 118(5), 3177–3186. https://doi.org/10.1121/1.2047107
- 1001 Spahr, A. J., & Dorman, M. F. (2004). Performance of Subjects Fit with the Advanced Bionics
- CII and Nucleus 3G Cochlear Implant Devices. Archives of Otolaryngology Head and
 Neck Surgery, 130(5), 624–628. https://doi.org/10.1001/archotol.130.5.624
- 1004 Spahr, A. J., & Dorman, M. F. (2003). A comparison of performance among patients with the
- 1005 CII Hi-Resolution, 3G and Tempo+ processors. 2003 Conference on Implantable Auditory
- 1006 Protheses, January, 5. http://www.zachary.com/creemer/three_group_comparison1.pdf

Spahr, A. J., Dorman, M. F., & Loiselle, L. H. (2007). Performance of patients using different
 cochlear implant systems: Effects of input dynamic range. *Ear and Hearing*, 28(2), 260–

1009 275. https://doi.org/10.1097/AUD.0b013e3180312607

- 1010 Stickney, G. S., Nie, K., Kong, Y. Y., Chen, H., & Zeng, F. G. (2004). Temporal fine structure:
- 1011 The missing component in speech processing algorithms. *International Congress Series*,

1012 *1273*(C), 23–26. https://doi.org/10.1016/j.ics.2004.09.017

- 1013 Tao, D., Deng, R., Jiang, Y., Galvin, J. J., Fu, Q. J., & Chen, B. (2015). Melodic pitch perception
- and lexical tone perception in mandarin-speaking cochlear implant users. *Ear and Hearing*,

1015 *36*(1), 102–110. https://doi.org/10.1097/AUD.00000000000086

- Theodore, R. M., Myers, E. B., & Lomibao, J. A. (2015). Talker-specific influences on phonetic
 category structure. *The Journal of the Acoustical Society of America*, *138*(2), 1068–1078.
 https://doi.org/10.1121/1.4927489
- 1019 van Heugten, M., Volkova, A., Trehub, S. E., & Schellenberg, E. G. (2014). Children's
- 1020 recognition of spectrally degraded cartoon voices. *Ear and Hearing*, *35*(1), 118–125.
- 1021 http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L5630837
- 1022 40%5Cnhttp://sfx.library.uu.nl/utrecht?sid=EMBASE&issn=15384667&id=doi:&atitle=Chi
- 1023 ldren%27s+recognition+of+spectrally+degraded+cartoon+voices.&stitle=Ear+Hear&title=
- 1024 Ear+and+h
- 1025 Vanags, T., Carroll, M., & Perfect, T. J. (2005). Verbal overshadowing: A sound theory in voice
 1026 recognition? *Applied Cognitive Psychology*, *19*(9), 1127–1144.
- 1027 https://doi.org/10.1002/acp.1160
- 1028 Vongpaisal, T., Trehub, S. E., Schellenberg, E. G., Lieshout, P. Van, & Papsin, B. C. (2010).
- 1029 Children with Cochlear Implants recognize their Mother's voice. Ear & Hearing, 31, 555–

1030 566.

- 1031 Vongphoe, M., & Zeng, F.-G. (2005). Speaker recognition with temporal cues in acoustic and
 1032 electric hearing. *The Journal of the Acoustical Society of America*, *118*(2), 1055–1061.
- 1033 https://doi.org/10.1121/1.1944507
- Wilkinson, E. P., Abdel-Hamid, O., Galvin, J. J., Jiang, H., & Fu, Q. J. (2013). Voice conversion
 in cochlear implantation. *Laryngoscope*, *123*(SUPPL. 3). https://doi.org/10.1002/lary.23744
- 1036 Zaltz, Y., Goldsworthy, R. L., Kishon-Rabin, L., & Eisenberg, L. S. (2018). Voice
- 1037 Discrimination by Adults with Cochlear Implants: the Benefits of Early Implantation for
- 1038 Vocal-Tract Length Perception. JARO Journal of the Association for Research in
- 1039 *Otolaryngology*, *19*(2), 193–209. https://doi.org/10.1007/s10162-017-0653-5
- 1040 Zhang, T., Dorman, M. F., Fu, Q. J., & Spahr, A. J. (2012). Auditory Training in Patients With
- 1041 Unilateral Cochlear Implant and Contralateral Acoustic Stimulation. *Ear and Hearing*,

1042 *33*(6), e70–e79. https://doi.org/10.1097/AUD.0b013e318259e5dd

- 1043
- 1044
- 1045
- 1046
- 1047
- 1048
- 1049
- 1050

1051

| Study | Goal | Participant characteristics | Task characteristics | Major findings | Quality Assessment |
|------------------------------|--|--|--|---|-----------------------|
| Abdeltawwab et al. (2016) | Examining whether there is a bimodal advantage in speech processing for CI users | CI group (n = 19) <u>Age</u> : M = 36.4y (SD = 14.3y; R = 18y - 68y) <u>OoD</u> : M = 28.7y* (SD = 14.0y*; R = 9y - 59y) <u>DoD</u> : M = 4.8y (SD = 2.7, R = 1y - 12y) <u>AoI</u> : M = 33.5y* (SD = 14.0y*; R = 14y - 64y) <u>LoU</u> : M = 2.9y* (SD = 0.8y*; R = 2y - 4y) <u>HC</u> : Bimodal | <u>Task</u> : 2AFC and AX discrimination <u>Stimuli</u> : Sentences <u>Talkers</u> : Men and women | CI group could discriminate male from female voices, but had difficulty discriminating within female and male voices No statistical difference in talker discrimination performance between bimodal and CI-only modes Some evidence that certain types of bimodal stimulation is favorable for talker discrimination | Strong |
| Carmel et al. (2011) | Comparing telephone use in children with CIs and NH children | CI group (n = 38) <u>Age</u> : M = 9y (R = 5y - 17y) <u>AoI</u> : M = 5y44d (R = 1y - 15y110d) <u>LoU</u> : M = 56m (R = 19m - 89m) <u>HC</u> : 46.0% Bimodal NH group (n = 27) <u>Age</u> : M = 7y (R = 5y - 17y) | <u>Task</u> : Parental survey | Parents of CI group reported that their children have more difficulty distinguishing between voices of men, women, and children, compared to parents of NH children. CI group reported having more difficulty distinguishing between voices of strangers and familiar people, and this difficulty was more pronounced for teenage participants | Moderate |
| Cleary & Pisoni (2002) | Examining the talker discrimination abilities of children with CIs | CI group (n = 44) <u>Age</u> : M = 8.8y (SD = 0.5y; R = 7.9y - 9.9y) <u>OoD</u> : M = 2.5m (SD = 7m; R = 0m - 36m) <u>DoD</u> : M = 2.9y (SD = 1.1y; R = 0.6y - 5.2y) <u>LoU</u> : M = 5.6y (SD = 0.7y; R = 4.1y - 6.9y) NH group (n = 21) <u>Age</u> : M = 5y6m (SD = 2m; R = 5y3m - 5y8m) | <u>Task</u> : AX discrimination <u>Stimuli</u> : Sentences <u>Talkers</u> : 3 women | CI group performed above chance in the talker discrimination task NH group outperformed CI group in trials where the sentence varied across talkers | Strong |
| Davidson et al. (2015) | Comparing effects of three different hearing aid frequencies on speech processing for bimodal CI users | CI group (n = 14) <u>Age</u> : M = 12y (SD = 4.3y*; R = 7y - 21y) <u>DoD</u> : M = 4.6y* (SD = 2.1y*; R = 1y - 9y*) <u>AoI</u> : M = 7.3y* (SD = 4.5y*; R = 3y - 20y*) <u>LoU</u> : M = 4.9y* (SD = 2.1y*; R = 1y - 9y*) <u>HC</u> : Bimodal | <u>Task</u> : AX discrimination <u>Stimuli</u> : Sentences <u>Talkers</u> : 8 men, 8 women <u>Conditions</u> : 3 HA settings (Wideband, restricted high- frequency, NLFC) and CI-only | • No difference between bimodal performance in any of the three bimodal conditions (wideband, restricted high frequency, or nonlinear frequency compression), compared to CI-only | Moderate |
| Davidson et al. (2019) | Examining role of acoustic experience for language development in children with CIs | CI group (n = 117) <u>Age</u> : M = 7.0y (SD = 1.3y, R = 4.8y - 9.4y) <u>OoD</u> : Prelingual (congenital or acquired before 20m) <u>AoI</u> : M _{FIRST} = 2.1y (SD = 1.1y, R = 0.67y - 4.9y) M _{SECOND} = 2.6y (SD = 1.4y, R = 0.7y - 6.7y) <u>HC</u> : 29 bimodal, 65 sequential bilateral, 23 simultaneous bilateral | <u>Task</u> : AX discrimination <u>Stimuli</u> : Sentences <u>Talkers</u> : 8 women | CI group performed above chance in the talker discrimination task Children with lower unaided PTA benefitted more from longer HA use Longer durations of acoustic experience were correlated with greater scores, however children with profound HL did not benefit from prolonged acoustic experience | Strong |

Table 1. Study design characteristics for studies assessing talker discrimination.

| Dorman et al. (2008) | Examining the contribution of low- frequency acoustic hearing to performance in speech perception tasks | CI group with EAS (n = 15) <u>LoU</u> : R = 5m - 7y <u>HC</u> : Bimodal CI representative group with E-only (n = 54) <u>HC</u> : Unilateral CI high-performing group with E-only (n = 65) <u>HC</u> : Unilateral | <u>Task</u> : AX discrimination <u>Stimuli</u> : Words <u>Talkers</u> : 5 men, 5 women | • No difference in within-gender or between- gender voice discrimination ability between EAS and E-only CI groups | Moderate |
|---------------------------------|--|--|---|---|----------|
| Geers et al. (2013) | Examining the linguistic and indexical skills of children with CIs | $\begin{array}{l} \textbf{CI group } (n = 60) \\ \underline{Age}: M = 10.5y \ (SD = 0.8y; R = 9.1y - 12.8y) \\ \underline{OoD}: M = 10m^* \ (SD = 8m^*; R = 1m - 2y6m) \\ \underline{Aol}: M = 1y10m^* \ (SD = 7m^*; R = 1y - 3y2m) \\ \underline{LoU}: M = 8y6m \ (SD = 0y11m; R = 6y2m - 11y2m) \\ \underline{HC}: 29 \ bilateral, 31 \ unilateral \\ \textbf{NH group } (n = 30) \\ \underline{Age}: M = 10.4y \ (SD = 0.8y) \end{array}$ | <u>Task</u> : AX discrimination <u>Stimuli</u> : Sentences <u>Talkers</u> : 8 men, 8 women | Majority of the CI group were able to discriminate male from female voices, but had difficulty discriminating across female voices Performance in the talker discrimination task was related to performance in linguistic processing tasks Earlier age of implantation was correlated with better voice discrimination skill and better indexical skills in general | Strong |
| Hay-McCutcheon et al. (2018) | Comparing talker and regional accent discrimination between unilateral and bimodal CI users | $CI group (n = 16)$ $Age: M = 57.3y (SD = 12.5y; R = 30.4y - 81.8y)$ $OoD: Postlingual$ $AoI: M = 52.9y^{*} (SD = 13.6y^{*}; R = 28.3y - 78.8y)$ $LoU: M = 4.4y (SD = 3.6; R = 1.2y - 15.6y)$ $HC: Unilateral$ $CI group (n = 19)$ $Age: M = 63.6y (SD = 10.3; R = 38.5y - 74.2y)$ $OoD: Postlingual$ $AoI: M = 60.4y^{*} (SD = 9.3y^{*}; R = 37.9y - 70.3y)$ $LoU: M = 3.2y (SD = 2.7; R = 0.5y - 9.2y^{*})$ $HC: Bimodal$ | Task 1: AX discrimination <u>Stimuli</u> : Sentences <u>Talkers</u> : 4 men and 4 women <u>Task 2</u> : AX discrimination <u>Stimuli</u> : Sentences <u>Talkers</u> : 8 men and 8 women <u>Conditions</u> : 4 men and 4 women had northern American dialects, and the other 4 men and 4 women had southern American dialects | Variable performance across participants, with some benefit to bimodal participants for talker discrimination Listeners may have used F0 and speaking rate to discriminate between talkers, and consistent acoustic-phonetic features to discriminate between regional accents | Moderate |
| Mühler et al. (2009) | Proposing new talker discrimination test for CI users in German | CI group (n = 12) <u>Age</u> : M = 51.8y* (SD = 16.5y*; R = 24y - 79y) <u>OoD</u> : R = 9y - 67y [#] <u>DoD</u> : M = 7.4y* (SD = 4.6y*; R = 2y - 15y) <u>AoI</u> : M = 49.3y* (SD = 17.6y*; R = 14y - 78y) <u>LoU</u> : M = 2.5y* (SD = 2.8y*; R = 0.5y - 10y) NH group (n = 10) <u>Age</u> : R = 28y - 58y | <u>Task</u> : AX discrimination <u>Stimuli</u> : Monosyllabic non-words <u>Talkers</u> : 5 men, 5 women | CI group performed above chance in the talker discrimination task CI group discrimination scores correlated to difference in F0 between talkers | Moderate |

| Sjoberg et al. (2017) | Examining relationship between timbre, pitch, and talker perception in children with CIs and NH children | CI group (n = 30) <u>Age</u> : M = 12.1y (R: 7y - 17y) <u>OoD</u> : M = 0.8y (SD = 1.2y; R = 0y - 4y) <u>AoI</u> : M = 4.8y (SD = 1.2y; R = 0.9y - 16.1y) <u>LoU</u> : M = 92.4m* (SD = 52.7m*; R = 18m - 192m) <u>HC</u> : 15 unilateral, 15 bilateral NH group (n = 35) <u>Age</u> : M = 10.4y (R = 8y - 14y) | <u>Task</u> : AX discrimination <u>Stimuli</u> : Sentences for CI group, noise-vocoded sentences for NH group <u>Talkers</u> : 5 men, 5 women | No difference in discrimination ability between CI group and CI simulations for NH group CI group performance on talker discrimination task was related to pitch perception, speech recognition and prior experience with music | Moderate |
|--------------------------|--|--|---|--|----------|
| Spahr & Dorman (2003) | Comparing speech processing performance across different CI devices | CI group with CII Bionic ear system $(n = 10)$ <u>Age</u> : M = 53.3y <u>DoD</u> : M = 11.9y <u>LoU</u> : M = 1.5y CI group with 3G system $(n = 10)$ <u>Age</u> : M = 54.1y <u>DoD</u> : M = 14.4y <u>LoU</u> : M = 1.6y CI group with Tempo+ system $(n = 10)$ <u>Age</u> : M = 53.8y <u>DoD</u> : M = 12.3y <u>LoU</u> : M = 1.8y | <u>Task</u> : AX discrimination <u>Stimuli</u> : Words <u>Talkers</u> : 5 men 5 women | • No difference in discrimination ability between CII Bionic ear, 3G and Tempo+ systems | Weak |
| Spahr & Dorman (2004) | Comparing speech processing performance across different CI devices | CI group with CII Bionic ear system (n = 15) <u>Age</u> : M = 56.1y (SD = 12.2y; R = 36.3y - 83.4y) <u>DoD</u> : M = 17.5y (SD = 19.7y; R = 0.2y - 53.9y) <u>LoU</u> : M = 1.5y (SD = 0.4y; R = 1.0y - 2.5y) CI group with 3G system (n = 15) <u>Age</u> : M = 56.7y (SD = 13.5y; R = 28.9y - 75.7y) <u>DoD</u> : M = 14.1y (SD = 10.4y; R = 1.1y - 32.8y) <u>LoU</u> : M = 2.0y (SD = 1.5y; R = 0.2y - 5.0y) | <u>Task</u> : AX discrimination <u>Stimuli</u> : Words <u>Talkers</u> : 5 men, 5 women | • No difference in discrimination ability between CII Bionic ear and 3G systems | Strong |

| Spahr et al. (2007) | Comparing speech processing performance across different CI devices | CI group with CII Bionic ear system (n = 26) <u>Age</u> : M = 55.0y <u>DoD</u> : M = 13.1y <u>LoU</u> : M = 1.5y <u>HC</u> : Unilateral | <u>Task</u> : AX discrimination <u>Stimuli</u> : Words <u>Talkers</u> : 5 men, 5 women | • No difference in discrimination ability between CII Bionic ear, 3G and Tempo+ systems | Strong |
|---------------------|--|---|--|--|--------|
| | | CI group with 3G system (n = 32) Age: $M = 50.5y$ DoD: $M = 9.9y$ LoU: $M = 2.1y$ HC: Unilateral | | | |
| | | CI group with Tempo+ system (n = 18) <u>Age</u> : $M = 52.2y$ <u>DoD</u> : $M = 12.6y$ <u>LoU</u> : $M = 2.2y$ <u>HC</u> : Unilateral | | | |

Indicates missing data

* Value calculated using data presented in table from original paper

NOTE: Carmel et al. (2011) contains both talker discrimination and identification tasks and is summarized in this table. Similarly, Abdeltawwab et al. (2016) contains both talker discrimination and gender identification tasks and is summarized in this table.

List of Abbreviations: AFC: Alternate Forced Choice, AoI: Age of Implantation, CI: Cochlear implant, E-only: Electric-only Stimulation, EAS: Electric-Acoustic Stimulation, DoD: Duration of Deafness, HC: Hearing Configuration, LoU: Length of Use, M: Mean, NH: Normal hearing, OoD: Onset of Deafness, R: Range, SD: Standard Deviation, VTL: Vocal tract length

| T 11 A G 1 | 1 . | 1 , . ,. | C / 1. | | , • |
|-----------------------|-----------------|--------------------------------------|-------------|-----------------|---------------------|
| Table / Mudy | docian ci | haractoristics | tov studios | accoccing voico | norcontion |
| 1 a 0 0 2.0 i u u v 0 | i c s i z n c i | uuuuuuuuuuuuuuuuuuuuuuuuuuuuuuuuuuuu | or sinuics | ussessing voice | DerceDuon. |
| | | | | | r · · · · · · · · · |

| Study | Goal | Participant characteristics | Task characteristics | Major findings | Quality Assessment |
|------------------------------|---|---|--|--|-----------------------|
| Barone et al. (2016) | Comparing influence of visual information on voice-gender categorization in CI and NH subjects | CI group (n = 14) <u>Age</u> : M = 61.7y (SD = 14y; R = 43y - 86y) <u>OoD</u> : Postlingual <u>DoD</u> : M = 11.5y (SD = 9y; R = 0.09y - 32y) <u>AoI</u> : M = 54.1y* (SD = 17.3y*; R = 25y - 85y) <u>LoU</u> : M = 7.6y (SD = 9y; R = 1y - 18y) NH group (n = 32) Age: M = 25y (SD = 7y) | <u>Task</u> : 2AFC <u>Stimuli</u> : hVd word <u>Talkers</u> : 1 man, 1 woman, each acoustically manipulated into a continuum of 11 voices <u>Conditions</u> : Audio stimuli presented alongside visual stimuli of either a man or a woman | CI group strongly influenced by visual information when performing auditory categorization No effect of visual information found in NH listeners, even when signal mimics a CI processor | Moderate |
| Cleary et al. (2005) | Comparing the talker discrimination abilities of children with CIs and NH children | CI group (n = 18) Age: M = 8y1m (SD = 23.9m*; R = 5y6m - 12y7m) OoD: M = 4.9m* (SD = 10.6m*; R = 0y - 3y) DoD: M = 30.7m* (SD = 13.2m*, R = 7.0m - 63.0m) AoI: M = 3y (SD = 18.5m*; R = 1y5m - 6y1m) LoU: M = 5y1m (SD = 16.2m*; R = 2y9m - 7y6m) NH group (n = 24) | <u>Task</u> : AX discrimination <u>Stimuli</u> : Fixed or varied sentences <u>Talkers</u> : 1 woman acoustically manipulated to a 13-point continuum | NH group outperformed CI group, but a subset of CI group performed similarly to NH group CI group displayed more variable individual performance than NH group | Strong |
| | | <u>Age</u> : $M = 65.6m$ (SD = 2.2m) | | | |
| Dillier et al. (1994) | Examining different signal processing strategies for talker discrimination and phoneme recognition performance | CI group $(n = 5)$ <u>Age</u> : $R = 16y - 50y$ <u>OoD</u> : Postlingual <u>DoD</u> : $R = 1 - 10y$ <u>LoU</u> : $R = 12m - >10y$ | <u>Task</u> : 2AFC ⁺ <u>Stimuli</u> : Sentences <u>Talkers</u> : 2 men, 2 women <u>Conditions</u> : Different CI processing strategies | Participants performed at or below chance with continuous interleaved sampling, but performed much better with all other strategies High-spectral transmission (HST) was best strategy as it preserved speech quality features and had high continuous stimulus rate | Strong |
| El Boghdady et al. (2019) | Investigate whether sensitivity to F0 and VTL is related to speech-on-speech intelligibility and comprehension | CI group (n = 18) <u>Age</u> : M = 60.8y (SD = 12.4y; R = 33y - 76y) <u>DoD</u> : M = 29.1y* (SD = 19.3y*; R = 0.2y - 61.6y*) <u>LoU</u> : M = 6.8y* (SD = 5y*; R = 0.6y - 16.3y*) <u>HC</u> : 13 unilateral, 2 bilateral, 3 bimodal | <u>Task</u> : Adaptive 3AFC oddball <u>Stimuli:</u> CVCVCV pseudowords <u>Talkers:</u> 1 woman <u>Conditions:</u> F0 manipulation, VTL manipulation | CI users who were more sensitive to F0 (smaller JND) had greater benefit in speech-on-speech intelligibility CI users who were more sensitive to VTL (smaller JND) performed poorer with greater differences in VTL during speech-on-speech intelligibility and comprehension CI users who are more sensitive to both F0 and VTL performed better in both speech-on-speech intelligibility and comprehension | Strong |
| El Boghdady et al. (2021) | Investigate the role of spectral contrast enhancement for F0 and VTL perception and speech-on-speech intelligibility | CI group (n = 14) <u>Age</u> : M = 63y (SD = 13.3y; R = 39 - 81y) <u>LoU</u> : M = 9.9y* (SD = 4.2y*; R = 2.8y - 16.4y*) <u>DoD</u> : M = 11.1y* (SD = 17.8y*; R = 0.3 - 46.7y*) | <u>Task</u> : Adaptive 3AFC oddball <u>Stimuli</u> : CVCVCV pseudowords <u>Talkers</u> : 1 woman <u>Conditions</u> : F0 manipulation, VTL manipulation, or both | No changes in JND between processing strategies (SCE versus ACE) Benefit of SCE for speech-on-speech intelligibility not due to increased sensitivity to F0 and VTL | Strong |

| Fu et al. (2004) | Comparing contribution of spectral and temporal cues for voice gender identification between NH and CI users | CI group (n = 11) <u>Age</u> : M = 56.7y* (SD = 10y*; R = 49y - 70y) <u>OoD</u> : M = 46y* [#] (SD = 17.9y* [#] ; R = 7y - 69y* [#]) <u>LoU</u> : M = 6.22y* (SD = 4.3y*; R = 2y - 13y) NH group (n = 6) Age: R = 22y - 30y | <u>Task:</u> 2AFC <u>Stimuli:</u> hVd words <u>Talkers:</u> 5 men, 5 women | CI group performed similarly to NH group when listening to 4-8 spectral channels with greater amounts of temporal information The number of channels and the cutoff frequency of envelope extraction played an important role for gender voice recognition | Moderate |
|------------------------------|---|--|--|--|----------|
| Fu et al. (2005) | Comparing contribution of spectral and temporal cues for voice gender identification | CI group (n = 10) <u>Age</u> : M = 58.4y* (SD = 13.4y*; R = 35y - 74y) <u>OOD</u> : M = 33.8y* (SD = 19.6y*; R = 0y - 65y) <u>AOI</u> : M = 51.2y* (SD = 13.5y*; R = 32y - 67y*) <u>LOU</u> : M = 7.2y* (SD = 4.3y*; R = 2y - 14y) NH group (n = 10) Age: R = 22y - 38y | Task: 2AFC <u>Stimuli</u> : hVd words <u>Talkers</u> : 10 men, 10 women <u>Conditions</u> : 2 talker groups (non- overlapping group with 5 female voices with highest F0 and 5 male voices with lowest F0; remaining voices in overlapping group) | CI group performed similarly to NH group exposed to 4- and 8-channel sine-wave speech Temporal cues were enough for CI group to identify gender when F0 was distinct When F0 overlapped, CI group could not effectively attend to cues when discriminating talker gender | Moderate |
| Fuller et al. (2014) | Examining the contribution of F0 and VTL cues to gender categorization by CI users | CI group $(n = 19)$ <u>Age</u> : M = 64.6y (R = 28y - 78y) <u>OoD</u> : Postlingual <u>LoU</u> : M = 4.6y (SD = 3.0y*; R = 1y - 12y) <u>HC</u> : 1 bilateral, 18 unilateral NH group $(n = 19)$ Acc: M = 22 by (R = 10y - 28y) | Task: 2AFC <u>Stimuli</u> : Words <u>Talkers</u> : 1 woman, acoustically manipulated to a continuum of 30 male and female voices | CI group relied more on F0 and less on VTL cues compared to NH group Some evidence that the type of speech processor affected gender categorization | Strong |
| Gaudrain & Başkent (2018) | Examining the perceptibility of F0 and VTL cues by CI users | <u>Age:</u> $M = 22.19 (R = 19y - 28y)$ CI group (n = 11) <u>Age:</u> $M = 60.19 (SD = 9.2y^*; R = 47y - 74y)$ <u>Aol:</u> $M = 51.8y^* (SD = 9.4y^*; R = 38y - 67y)$ <u>LoU</u> : $M = 8.3y^* (SD = 4.3y^*; R = 1y - 13y)$ <u>HC</u> : 1 unilateral with residual acoustic hearing, 10 unilateral only | <u>Task</u> : Adaptive 3AFC oddball <u>Stimuli</u> : CVCVCV pseudowords <u>Talkers</u> : 1 woman, acoustically manipulated <u>Conditions</u> : F0 only, VTL only, both | CI group had larger just noticeable difference (JND) in both F0 and VTL CI group relied on F0 to discriminate gender Only 1 CI user had a VTL JND smaller than normal difference between male and female voices | Strong |
| Hazrati et al. (2015) | Investigating speaker- gender identification and speech intelligibility for CI users in neutral and whispered speech | CI group (n = 6) <u>Age</u> : M = 62.8y (SD = 9.3y*; R = 55y - 80y) <u>OOD</u> : Postlingual <u>AOI</u> : M = 58.3y* (SD = 7.8y*; R = 50.2y - 72y) <u>LOU</u> : M = 4.6y* (SD = 1.8y*; R = 3y - 8y) | <u>Task</u> : 2AFC <u>Stimuli:</u> Spontaneous and read sentences <u>Talkers</u> : 10 men, 10 women | CI group performed better when presented with neutral versus whispered speech Speech intelligibility decreased for CI group when presented with whispered speech | Moderate |
| Horng et al. (2007) | Examining CI users' abilities to process speech via the telephone | $\begin{array}{l} \textbf{CI group } (n = 15) \\ \underline{Age}: M = 10.9y \ (SD = 2.3y^*; R = 8y - 16y) \\ \underline{OOD}: \ Prelingual \\ \underline{AoI}: M = 5.5y^* \ (SD = 2.7y^*; R = 2y - 11y) \\ \underline{LoU}: M = 5.5y \ (SD = 1.5y^*; R = 3y - 7y) \end{array}$ | <u>Task</u> : 2AFC <u>Stimuli</u> : Single-vowel syllables <u>Talkers</u> : 1 man, 1 woman <u>Conditions</u> : Broadband speech (not modified), telephone speech (filtered broadband speech) | • CI group had worse gender identification performance for telephone speech, compared to broadband speech | Moderate |
| Kovačić & Balaban (2009) | Examining the quality of information provided by CI devices for voice gender discrimination | CI group $(n = 41)$ <u>Age</u> : M = 12.3y (SD = 3.2y; R = 5.3y - 18.8y) <u>HC</u> : Unilateral NH group $(n = 15)$ <u>Age</u> : M = 9.3y (R = 6.7y - 10.6y) | Task: (1) Fixed single-interval 2AFC (2) Adaptive two-interval 2AFC Stimuli: 2 second long speech sample Talkers: 20 men, 20 women | • No difference was found in quality of information provided by the devices of subjects who performed well at gender identification and those who did not | Strong |

| Kovačić & Balaban (2010) | Examining the hearing history factors that affect CI users' gender identification ability | CI group (n = 41) <u>Age</u> : M = 12.3y (SD = 3.2y; R = 5.3y - 18.8y) <u>DoD</u> : M = 6.2y (SD = 3.2y; R = 0y8m - 12y5m) <u>Aol</u> : M = 8.5y* (SD = 3.3y*; R = 2y1m - 15y4m) <u>LoU</u> : M = 3.8y (SD = 1y; R = 0y7m - 7y) <u>HC</u> : Unilateral NH group (n = 15) <u>Age</u> : M = 9.3y (SD = 1.3y; R = 6.7y - 10.6y) | <u>Task</u> : (1) Fixed single-interval 2AFC (2) Adaptive two-interval 2AFC <u>Stimuli</u> : Two-second speech samples <u>Talkers</u> : 20 men, 20 women | Duration of auditory deprivation and age of CI activation appeared to be related to gender identification Longer duration of deafness before implantation negatively impacted gender identification | Moderate |
|-----------------------------|---|--|--|---|----------|
| Landwehr et al. (2014) | Examining the effects of various electrode configurations and coding strategies on various speech processing skills | CI group (n = 6) <u>Age</u> : M = 51.3y* (SD = 26.1y*; R = 18y - 75y) <u>OoD</u> : Postlingual <u>LoU</u> : M = 10.3m* (SD = 6.9m*; R = 3m - 22m) | <u>Task</u> : 2AFC <u>Stimuli</u> : Sentences <u>Talkers</u> : 3 men, 3 women <u>Conditions</u> : 6 different electrode configurations | • No significant difference in gender identification across different processing strategies | Moderate |
| Li & Fu (2011) | Examining whether voice gender discrimination is a useful indicator of the spectral and temporal processing abilities of CI users | CI group (n = 11) Age: M = 60.6y* (SD = 18.2y*; R = 25y - 79y) OoD: Postlingual AoI: M = 52.9y* (SD = 16.6y*; R = 23y - 71y) LoU: M = 7.7y* (SD = 6.3y*; R = 1y - 19y) HC: 7 unilateral, 4 bilateral, 3 bimodal# | <u>Task</u> : 2AFC <u>Stimuli</u> : hVd words <u>Talkers</u> : 2 sets of 5 men and 5 women <u>Conditions</u> : 2 sets of talkers with varying intergender F0 differences; processed and unprocessed stimuli | • Performance in voice gender discrimination was correlated with tasks that assessed spectral and temporal processing | Moderate |
| Massida et al. (2013) | Examining voice gender discrimination in CI users, and how it changes as a function of CI experience | CI Transversal group (n = 32) <u>Age</u> : M = 54.5y (SD = 15y; R = 21y - 81y*) <u>OoD</u> : Postlingual <u>DoD</u> : R = 2y - >40y [#] <u>HC</u> : 16 Bimodal [%] , 17 Unilateral CI Follow-up group (n = 10) <u>Age</u> : M = 51.9y (SD = 16y; R = 21y - 81y) <u>OoD</u> : Postlingual <u>DoD</u> : R = 3y - >40y [#] <u>HC</u> : 4 Bimodal; 6 Unilateral NH group (n = 14) <u>Age</u> : M = 24.6y (SD = 2.9y) | Task: 2AFC Stimuli: hVd words Talkers: 1 man, 1 woman, each synthesized into a continuum of 11 voices | CI users performed poorly in differentiating between male and female voices, as well as vocal and non-vocal sounds This deficit did not improve with CI experience | Moderate |
| Meister et al. (2009) | Comparing the effect of changes in F0 on the perception of prosody and gender between NH adults and CI users | CI group (n = 12) Age: M = 58y (SD = 13.3y; R = 38-75y) AoI: M = 55.4y* (SD = 13.5y; R = 36.8y - 73.5y) LoU: M = 2.9y* (SD = 1.9y*; R = 0.5y - 6.7y) HC: Unilateral NH group (n = 12) Age: M = 47y (R = 34y - 68y) | Task 1: 2AFC Stimuli: Sentences Talkers: 3 men, 3 women Task 2: 2AFC Stimuli: Sentences Talkers: 1 woman, F0 contour and VTL acoustically manipulated Conditions: question vs statement, sentence stress, speaker gender | CI group had difficulty with one male voice that had highest F0 (closest to female voices) Categorization of F0 continuum did not differ between NH and CI participants Significant correlation between natural utterance and modified stimuli results suggest that CI group used F0 contour of speakers to identify gender | Moderate |

| Meister et al. (2016) | Examining the influence of F0 and VTL modifications on gender identification | CI group (n = 16) <u>Age</u> : M = 53.8y (SD = 17.4y*; R = 21y - 73y) <u>OoD</u> : Postlingual <u>LoU</u> : M = 55.4m* (SD = 53.7m*; R = 6m - 192m) <u>HC</u> : 9 Bimodal, 5 Bilateral, 2 Unilateral NH group (n = 12) Age: M = 56y (R = 26y - 74y) | <u>Task</u> : Gender rating scale <u>Stimuli</u> : Words and sentences (forward and reversed) <u>Talkers</u> : 1 woman, acoustically manipulated to a continuum of voices <u>Conditions</u> : F0 manipulation, VTL manipulation, both | Compared to NH group, CI group relied more on F0 cues over VTL cues in rating the gender of a voice CI group can make use of both cues only when presented with sentences, and not with words | Moderate |
|----------------------------|---|---|--|--|----------|
| Nogueira et al. (2021) | Examining the effect of electrode interaction on F0, VTL, and speech-on-speech perception | CI group (n = 12) <u>Age</u> : M = 50.3y* (SD = 12.4y*; R = $20y - 70y$) <u>LoU</u> : M = 6.9y* (SD = $3.2y*$; R = $2.5y - 11.4y$) | <u>Task 1</u> : 3AFC adaptive oddball <u>Stimuli</u> : CVCVCV syllables <u>Talkers</u> : 1 woman, acoustically manipulated <u>Conditions</u> : 3 acoustic manipulations (F0, VTL, or both); 3 coding strategies (sequential, paired, triplet) | • Increasing electrode interaction reduces sensitivity to F0 and combined F0 + VTL cues | Strong |
| Skuk et al. (2021) | Assessing the relative importance of F0 and timbre for voice gender and age perception | CI group (n = 28) <u>Age</u> : M = 59y (SD = 19.7y; R = $12y - 85y$) <u>DoD</u> : M = 23.8y (SD = 19.2y; R = $0y - 59y$) <u>LoU</u> : M = 3.1y (SD = 3y; R = $0y - 11y$) <u>HC</u> : 15 bilateral, 10 bimodal, 3 unilateral NH group (n= 19) <u>Age</u> : M = 42.9y (SD = 19.6y; R = 15 - 85y) | Task: 2AFC Stimuli: VCV nonwords Experiment 1 Talkers: 4 women, 4 men Conditions: F0 manipulation, timbre manipulation, or both Experiment 2 Talkers: 8 young, 8 old Conditions: F0 manipulation, timbre manipulation, time manipulation, or full manipulation (all) | For gender perception, CI group largely relies on F0, no evidence for efficient use of timbre For age perception, CI group shows an ability to use timbre Best performing CI users show pattern similar to NH listeners, worst performing CI users are below chance at perceiving age | Strong |
| Wilkinson et al. (2013) | Evaluating voice conversion algorithms for improving speech processing in CI users | CI group (n = 6) <u>Age</u> : M = 64.8y* (SD = $11.8y*$; R = $47y - 79y$) <u>OOD</u> : Postlingual <u>AOI</u> : M = $52.8y*$ (SD = $8.7y*$; R = $42y - 64y$) <u>LOU</u> : M = $12y$ (SD = $6.7y*$; R = $3y - 21y$) <u>HC</u> : 2 Bilateral, 4 Unilateral | Task: 2AFC Stimuli: Sentences Talkers: 1 man, 1 woman Conditions: Different VTL processing algorithms used to transform man's voice to woman's voice and woman's voice to man's | Gender identification was significantly better with original speech compared to transformed speech, only when source F0 was included in transformed speech Spectral envelope and pitch information were important for gender identification | Moderate |
| Zaltz et al. (2018) | Comparing talker discrimination abilities of pre-lingually deafened CI users and NH adults based on F0 and VTL | CI group (n = 18) <u>Age</u> : M = 25.1y (SD = 3.9y; R = 19y7m - 35y3m) <u>OoD</u> : 16 congenitally deaf, 2 deafened <1.5y <u>AoI</u> : M = 9.7y* (SD = 9.7y*; R = 2y3m - 33y4m) <u>LoU</u> : M = 14.6y* (SD = 5.9y*; R = 1.5y - 21.8y) <u>HC</u> : 7 bilateral, 11 unilateral users NH group (n = 9) <u>Age</u> : M = 23.8y (SD = 2.4y) | <u>Task</u> : AX discrimination <u>Stimuli</u> : Sentences <u>Talkers</u> : 1 woman, acoustically manipulated to a 13-point continuum | Age of implantation was related to participants' ability to discriminate voices based on VTL cues, but not on F0 cues A moderate relationship was found between VTL discrimination and speech recognition scores VTL perception was poorer in CI users who were implanted late compared to CI users who were implanted early | Moderate |

| Zhang et al. (2012) | Examining whether | CI group $(n = 7)$ | Task: 2AFC | No improvement in gender identification post- | Moderate |
|---------------------|-------------------------|---|-------------------------|---|----------|
| | auditory training | <u>Age</u> : $M = 63.1y^*$ (SD = 8.8y*; R = 51y - 78y) | Stimuli: hVd words | training | |
| | improves the benefit | <u>OoD</u> : $M = 26.8y^*$ (SD = 20.8y*; R = 2y - 50y) | Talkers: 5 men, 5 women | | |
| | of EAS in speech | <u>DoD</u> : $M = 36.3y^*$ (SD = 23.9y*; R = 16y - 76y) | | | |
| | processing for CI users | <u>AoI</u> : $M = 59.1y^*$ (SD = 9.1y*; $R = 49y - 76y$) | | | |
| | | <u>LoU</u> : $4.3y$ (SD = $3.4y$; R = $2y - 10y^*$) | | | |
| | | HC: Bimodal | | | |

+ Presumed from in-text description of "female/ male distinction sentence test"

[#] Missing or unknown data from original paper

[%] In-text information is inconsistent with data presented in respective table from original paper

* Value calculated using data presented in table from original paper

NOTE: Abdeltawwab et al. (2016) contains both talker discrimination and gender identification tasks, and is summarized in Table 1.

List of Abbreviations: ACE: Advanced combination encoder, AFC: Alternate Forced Choice, AoI: Age of Implantation, CI: Cochlear implant, EAS: Electric-Acoustic Stimulation, DoD: Duration of Deafness, HC: Hearing Configuration, JND: Just Noticeable Difference, LoU: Length of Use, M: Mean, NH: Normal hearing, OoD: Onset of Deafness, R: Range, SCE: Spectral contrast enhancement, SD: Standard Deviation, VTL: Vocal tract length

| T 11 2 C 1 | 1 • 1 | · · · · · | , 1. | • , | 11 • 1 | ·· · · |
|---|----------------------------|--------------------|--------------|------------|------------|------------|
| Table & Mudi | $, d\rho c \alpha n c h a$ | ractoristics to | r chidioc ac | ccoccinot | alkov idon | titication |
| 1 a 0 0 0.0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | uesizn enu | i ucies issues j | i sinuics us | socssing n | iner iden | uncanon |
| | 0 | | | | | ./ |

| Study | Goal | Participant characteristics | Task characteristics | Major findings | Quality Assessment |
|------------------------------|---|---|---|--|-----------------------|
| Cullington & Zeng (2010) | Comparing aspects of speech recognition in bimodal and bilateral CI users | CI group (n = 26) <u>Age</u> : Adult <u>OoD</u> : Postlingual <u>HC</u> : 13 bilateral, 13 bimodal | <u>Task</u> : 10AFC <u>Stimuli</u> : hVd words <u>Talkers</u> : 2 boys, 2 girls, 3 men, 3 women | Bimodal CI group performed better than bilateral CI group on exact talker identification and category identification (woman, man, child), but not to a statistically significant level No correlation found between talker identification and other pitch-related tasks | Moderate |
| Cullington & Zeng (2011) | Comparing bimodal and bilateral CI users on various speech tasks | Bilateral CI group (n = 13) Age: M = 56y (SD = 12.7y*; R = 38y - 75y) OoD: Postlingual DoD: M = 7.5y (SD = 8.4y*; R = 0.3y - 24y) LoU: M = 7.2y Bimodal CI group (n = 13) Age: M = 63y (SD = 17.7y*; R = 42y - 87y) OoD: Postlingual DoD: M = 16.5y (SD = 17.0y*; R = 1y - 50y) LoU: M = 2.6y | <u>Task</u> : 10 AFC <u>Stimuli</u> : hVd words <u>Talkers</u> : 2 boys, 2 girls, 3 men, 3 women | No difference found in performance between bilateral and bimodal CI groups Positive correlation found between talker identification and affective prosody identification | Strong |
| Morris et al. (2015) | Comparing children with CIs to their NH siblings on various dimensions of speech processing | CI group (n = 18) <u>Age</u> : M = 7.5y (SD = 2y9m; R = 2y - 15.67y) <u>OOD</u> : Prelingual <u>AOI</u> : M = 2.5y (SD = 2.6y; R = 0.75y - 5.08y) <u>LOU</u> : M = 4.9y* (SD = 4.2y*; R = 0.3y - 12.8y) <u>HC</u> : 16 bilateral, 2 unilateral NH group (n = 18) <u>Age</u> : M = 6y5m (SD = 2y8m) | <u>Task</u> : 11-item parent questionnaire with one item about talker identification | CI group rated as being worse at voice identification compared to their NH siblings Ratings were not related to demographic factors (age, age at implantation, length of use) | Moderate |
| Stickney et al. (2004) | Examining new speech processing algorithm for CIs which includes temporal fine structure | CI group (n = 10) NH group (n = 5) | <u>Task</u> : 3AFC <u>Stimuli</u> : Vowels <u>Talkers</u> : 2 boys, 2 girls, 3 men, 3 women | • CI group presented with unprocessed speech perform similarly to NH group presented with 4- channel vocoded amplitude modulation information | Weak |
| van Heugten et al. (2014) | Comparing ability of NH children and children with CIs to recognize cartoon voices | CI group (n = 15) <u>Age</u> : M = 5.5y (SD = 0.7y; R = 4y - 7y) <u>AoI</u> : M = 1.5y* (SD = 0.8y*; R = 0.8y - 3.4y) <u>LoU</u> : M = 4.0y* (SD = 0.8y*; R = 2.4y - 5.3y) <u>HC</u> : Bilateral | <u>Task</u> : 3AFC <u>Stimuli</u> : Sentences <u>Talkers</u> : 12 cartoon characters | • NH group slightly better than CI group at recognizing unprocessed cartoon voices | Moderate |
| | | NH group (n = 15) <u>Age</u> : M = 4.7y (SD = 0.3 y; R = 4y - 5y) | | | |

| Vongpaisal et al. (2010) | Investigating whether children with CIs can identify their mother's voice among other talkers (Experiment 1) and among women's voices (Experiment 2) | Experiment 1 CI group (n = 21) Age: M = 8.9y (SD = 3.0y; R = 4.7y - 14.3y) OoD: 20 prelingual or congenital, 1 progressive HL. AoJ: Bilateral M = 2.4y* (SD = 1.3y*; R = 0.7y - 3.5y); Unilateral M = 3.1y* (SD = 2.0y*; R = 1.0y - 8.7y) LoU: Bilateral M = 6.5 y (SD = 1.4y; R = 4.3y - 7.6y); Unilateral M = 5.7y (SD = 2.4y; R = 2.9y - 8.8y) HC: 5 bilateral, 16 unilateral NH group (n = 16) Age: M = 5.6y (SD = 0.3y) Experiment 2 CI group (n = 19 from Exp 1) Age: M = 9.3y* (SD = 2.9y; R = 5.3y - 14.3y) AoI: Bilateral M = 2.2y* (SD = 0.9y*; R = 0.7y - 3.5y); Unilateral M = 3.9y* (SD = 2.1y*; R = 1.1y - 8.0y) LoU: Bilateral M = 5.6y* (SD = 1.8y*; R = 3.3y - 8.5y); Unilateral M = 6.9y* (SD = 2.7y*; R = 1.0y - 10.7y) HC: 9 bilateral, 10 unilateral | Experiment 1 Task: 4AFC Stimuli: 5 questions, 5 statements in CDS Talkers: Mother and 3 actors (man, woman, child) Conditions: Natural & imitation (actors imitate mother's voice) Experiment 2 Task: 2AFC Stimuli: 9 utterances from Exp 1 Talkers: Mother and 3 additional women Conditions: Natural & imitation (actors imitate mother's voice) | Experiment 1: CI group performed worse than NH group overall, but performed better with natural productions than imitated productions Experiment 2: CI group identified their mother's voice better with natural over imitated productions; despite smaller F0 differences between voices, participants performed similarly to Exp 1 | Moderate |
|-----------------------------|--|--|--|---|----------|
| | | NH group (n = 11 from Exp 1) | | | |

| Vongphoe & Zeng (2005) | Examining the role of temporal cues on vowel and voice recognition by CI users and NH adults | $\begin{array}{l} \textbf{CI group} \ (n=10) \\ \underline{Age}: \ M=65.9y^{*} \ (SD=7.3y^{*}; \ R=49y \ -74y) \\ \underline{OoD}: \ M=37.4y^{*} \ (SD=12.9y^{*}; \ R=9y \ -57y) \\ \underline{DoD}: \ M=29.3y^{*} \ (SD=9.2y^{*}; \ R=17y \ -40y) \\ \underline{LoU}: \geq 1y \end{array}$ | <u>Task</u> : 10 AFC <u>Stimuli</u> : hVd words <u>Talkers</u> : 2 boys, 2 girls, 3 men, 3 women | CI group performed poorly compared to NH participants A slowly varying form of frequency modulation is important for speaker recognition | Moderate |
|---------------------------|--|---|---|---|----------|
| | | NH group $(n = 6)$ <u>Age</u> : R = 18y - 32y | | | |

* Value calculated using data presented in table from original paper NOTE: Carmel et al. (2011) contains both discrimination and identification results and is summarized in the Talker Discrimination Table (Table 1).

List of Abbreviations: AFC: Alternate Forced Choice, AoI: Age of Implantation, CI: Cochlear implant, CDS: Child directed speech, DoD: Duration of Deafness, HC: Hearing Configuration, LoU: Length of Use, M: Mean, NH: Normal hearing, OoD: Onset of Deafness, R: Range, SD: Standard Deviation