Audiological and Demographic Factors that Impact the Precision of Speech Categorization in Cochlear Implant Users

Sarah Colby
Dept. of Psychological and Brain Sciences
Dept. of Otolaryngology
University of Iowa

Michael Seedorff
Dept. of Biostatistics
University of Iowa
and
Alphabet, Inc.

Bob McMurray
Dept. of Psychological and Brain Sciences
Dept. of Otolaryngology
Dept. of Communication Sciences and Disorders
Dept. of Linguistics
University of Iowa

[This manuscript has been accepted for publication at *Ear and Hearing*]

Conflicts of Interest: The authors have no conflicts of interest to report.
Funding sources: This work was supported by NIH Grant P50 000242 awarded to Bruce Gantz and Bob McMurray and by DC 008089 awarded to Bob McMurray

Corresponding Author:
Sarah Colby
sarah-colby@uiowa.edu
G60, Psychological and Brain Sciences Building
University of Iowa
Iowa City, IA
52242

Abstract

Objectives: The ability to adapt to subtle variation in acoustic input is a necessary skill for successful speech perception. Cochlear implant (CI) users tend to show speech perception benefits from the maintenance of their residual acoustic hearing. However, previous studies often compare CI users in different listening conditions within-subjects (i.e., in their typical Acoustic + Electric configuration compared to Acoustic-only or Electric-only configurations) and comparisons among different groups of CI users do not always reflect an Acoustic + Electric benefit. Existing work suggests that CI users with residual acoustic hearing perform similarly to Electric-only listeners on phonetic voicing contrasts and unexpectedly poorer with fricative contrasts which have little energy in the range of the Acoustic + Electric listeners' acoustic hearing. To further investigate how residual acoustic hearing impacts sensitivity to phonetic ambiguity, we examined whether device configuration, age, and device experience influenced phonetic categorization in a large individual differences study.

Design: CI users with various device configurations (Electric-only N = 41; Acoustic + Electric N = 95) categorized tokens from five /b-p/ and five /s-ʃ/ minimal pair continua (e.g., bet-pet; sock-shock). We investigated age, device experience, and when applicable, residual acoustic hearing (pure tone hearing thresholds) as predictors of categorization. We also examined the relationship between phonetic categorization and clinical outcomes (CNC, AzBio) in a subset of our sample.

Results: Acoustic + Electric CI users were better able to categorize along the voicing contrast (steeper categorization slope) compared to Electric-only users, but there was no group-level difference for fricatives. There were differences within the subgroups for fricatives: bilateral users showed better categorization than unilateral users and bimodal users had better categorization than hybrid users. Age was a significant factor for voicing, while device experience was significant for fricatives. Critically, within the Acoustic + Electric group, hybrid CI users had shallower slopes than bimodal CI users.

Conclusions: Our findings suggest residual acoustic hearing is beneficial for categorizing stop voicing, but not frication. Age impacts the categorization of voicing, while device experience matters for fricatives. For CI users with ipsilateral residual acoustic hearing, those with better hearing thresholds may be over-relying on their acoustic hearing rather than extracting as much information as possible from their CI, and thus have shallower fricative categorization.

Introduction

Cochlear implants restore sound to profoundly deaf individuals by bypassing damaged hair cells in the cochlea to directly (electrically) stimulate the auditory nerve. However, the signal from the implant is degraded compared to the input of a normal hearing (NH) listener. This is largely because the entire frequency range is transmitted across a small number of electrodes (typically 12-22), though there are other sources of degradation (e.g., limited frequency range). Consequently, post-lingually deaf Cochlear Implant (CI) users must adapt speech perception systems that were previously developed for high-fidelity acoustic input to the novel patterns of electrical stimulation from the CI, which can take a year or more (Dorman et al., 2006).

In recent years, standards of cochlear implantation have changed to embrace bilateral implantation, and candidacy has expanded to include people with less than profound hearing loss. Consequently, there is wide variation in the hearing configurations used by many CI users, which vary based on the type and nature of the hearing loss, and the standards of care where they are treated. Some users will receive electric stimulation through a full-length implant, either in a single ear or in both (traditional *unilateral* and *bilateral* implants). Others receive both acoustic and electric (A+E) stimulation, in which some acoustic hearing is preserved to combine with the implant. Classically, A+E stimulation is achieved via *bimodal* configurations, in which the listener has a CI on one ear and uses a hearing aid to amplify any residual hearing in the contralateral ear (Wilson & Dorman, 2008). However, in recent years, newer (*hybrid*) implantation strategies leverage residual acoustic hearing in the implanted (ipsilateral) ear (Gantz et al., 2005; Woodson et al., 2009). These hybrid configurations use a CI that is constructed to maintain ipsilateral low-frequency acoustic hearing (and are sometimes shorter than a standard electrode).

These differences in device configuration lead to unique challenges for different types of CI users. While all CI users must adapt to novel stimulation through their implant, A+E users must also integrate electric and acoustic hearing, and this may differ depending on whether acoustic hearing arrives in the contra- or ipsilateral ear. Despite these challenges, most A+E users do benefit from the acoustic hearing (Dorman et al., 2008; Gantz et al., 2005; Turner et al., 2004, 2008). However, a recent study suggests that this may not be uniformly true across all speech sounds, particularly voiceless fricatives (McMurray et al., 2016); this rather special case could reveal limits to how A+E users adapt to their listening configuration. The goal of the present study is to examine this more closely with a large and variable sample of CI users.

Benefits and limits of A+E listening

Generally, there is a benefit for maintaining residual acoustic hearing in some form, as seen in gross outcome measures of word and sentence recognition. Substantial work compares A+E listeners in such measures with and without acoustic hearing. Monosyllabic word recognition is usually improved in combined A+E listening compared to unilateral electric-only (E-only) stimulation (Dorman et al., 2008), although some individuals do not show an acoustic benefit for word recognition in quiet (Dunn et al., 2005; Mok et al., 2006). Tests that directly compare types of listeners are more limited, but also show a benefit. Listeners in an A+E configuration often have better speech recognition in noise compared to E-only CI users (Dorman & Gifford, 2010). For example, when Hybrid CI are matched with E-only CI users on their speech in quiet performance, speech in noise performance showed benefit of up to around 9 dB SNR (Gantz et al., 2005; Turner et al., 2004). Combined A+E stimulation also benefits pitch discrimination and melody recognition in CI users. For example, Gfeller et al. (2007) showed that hybrid CI users performed similarly to NH when determining the direction of a pitch change while standard CI users performed worse. This was especially true at lower frequencies (i.e., where hybrid listeners are more likely to have residual acoustic hearing).

For some speech sounds however, A+E CI users do not always perform better than E-only users. McMurray et al. (2016) presented continua of stops and fricatives to CI users and NH listeners in an analogue of traditional phoneme identification tasks. These contrasts were chosen to highlight information that would primarily be transmitted by CIs (high-frequency spectral information for /s-ʃ/) or primarily by acoustic hearing (low-frequency voicing information for /b-p/). A+E and E-only CI users categorized stop voicing similarly, though both groups showed shallower categorization slopes than NH listeners (indicating noisier or more gradient categorization). Importantly, for fricatives, A+E CI users had unexpectedly *shallower* categorization slopes than E-only CI users; they appeared to perform worse. While the sample size was modest (N = 30), this was true for both hybrid and bimodal CI users. Residual acoustic hearing, in this case, was not advantageous for stop perception and seemed to be detrimental to fricative perception.

These results are surprising for two reasons. First, any simple reading of the extant literature would suggest that A+E users should have sharper categorization across both contrasts. Alternatively, they could show a benefit for voicing but not for fricatives. Zhang et al. (2010), for example, found that most of the benefit from combined A+E stimulation came from the low frequency region around the F0 of the voice. This should predict a benefit for a stop voicing continuum, where this acoustic range is critical. In contrast, fricatives—where the differences were observed—do not have any meaningful information in the range of a A+E listeners' residual acoustic hearing and should be entirely transmitted through the CI. This is particularly the case for bimodal users, who receive a full-length implant like unilateral users (so any differences cannot be attributed to differences in the CI). Thus, McMurray et al (2016) appears to suggest that A+E users may get less information from their implant than E-only listeners, or that they over-rely on acoustic hearing (which cannot help them with fricatives).

Evidence from Mok et al. (2006) supports the idea that bimodal users may not always take full advantage of their CI. They tested bimodal listeners with and without their hearing aid to compute a bimodal benefit (A+E minus E-only). They unexpectedly found that bimodal CI users with poorer residual hearing showed a greater bimodal benefit. Individuals with better residual hearing did not improve as much in the combined hearing aid + CI listening condition as those with poorer residual hearing. For these bimodal users, Mok et al. (2006) suggests that their residual mid- to high-frequency acoustic hearing may provide conflicting information that is hard to integrate with the CI. While not tested by Mok et al., the same premise could be true of hybrid CI users. For hybrid users, residual low-frequency hearing from their implanted ear could be difficult to integrate with information from their CI. This could explain why A+E users in McMurray et al. (2016) did not see a benefit during stop voicing categorization.

Second, the McMurray et al. (2016) result is surprising because it calls into question the broader assumption that acoustic hearing will always outperform electric stimulation. The evidence base for this assumption is not as strong as one might think. Most investigations of residual acoustic hearing in CI users do not compare performance of A+E users to different E-only CI users. Rather, these studies use within-subjects designs that compare performance in A+E listeners' day-to-day A+E configuration to performance in an E-only configuration (Dorman & Gifford, 2010; Dunn et al., 2005; Gantz et al., 2005; Mok et al., 2006; Turner et al., 2004). However, in such studies, the A+E condition is the configuration they use every day—the E-Only condition (where subjects often performed worse) is a less familiar listening configuration.

True comparisons between groups are rare (though see: Dorman & Gifford, 2010; Gantz et al., 2005; Turner et al., 2004). This is in part because it is difficult to properly match groups on things like age, device use, and most importantly pre-implantation audiological capacities. Moreover, a randomized control trial is not possible. Different kinds of listeners will be eligible for a hybrid or bimodal CI. Those with better residual hearing and thus, better neural health, can receive the shorter electrode to maintain their acoustic hearing, while someone with more substantial hearing loss (and likely also poorer neural health) generally receive a full electrode. Differences between groups may also be confounded with age—people who were implanted decades ago (and are therefore typically older) may have only had the option of an E-only configuration, or older individuals who received an A+E configuration years earlier may have lost their residual hearing in the meantime. These effects are also confounded with device experience: older adults, for example, may have worse residual hearing but longer experience with their CI depending on when they received it. These confounds make it difficult to construct well matched groups to execute a true between-subjects comparison.

Beyond these issues, there are also concerns about collapsing different A+E listening configurations. While McMurray et al. (2016) did not find differences between bimodal and hybrid listeners, these devices are often tuned differently. At the time of that study, hybrid implants at the University of Iowa were generally tuned to avoid overlap in frequency with the listener's residual acoustic hearing. As a result, low frequencies (below approximately 1000 Hz) may be transmitted only via acoustic hearing, and not via the CI. In contrast, bimodal listeners typically receive an implant tuned to the full spectrum (since their residual hearing will be from the contralateral ear). A larger sample with greater audiological and demographic diversity may be able to tease these factors apart.

The Present Study

The present study set out to address these questions by testing a large sample of CI users, including E-Only users (unilateral and bilateral), along with both bimodal and hybrid A+E listeners on a phonetic identification task similar to McMurray et al. (2016). This was intended to permit a multiple regression approach to better investigate the role of residual acoustic hearing for phonetic perception while controlling for age and device experience.

We examined several factors that might influence phonetic categorization. First, we compared A+E and E-Only CI users and further broke down these groups into more specific device configurations (unilateral, bilateral, bimodal, hybrid, and single-sided deafness). In addition to the standard subgroups of A+E and E-Only listeners (hybrid, bimodal, unilateral, and bilateral CI users), we also included single-sided deafness listeners (SSD), who represent a rather extreme case of A+E listening. In these listeners, one ear has complete access to the full spectrum (i.e., has hearing thresholds within the normal range) while the other ear has a CI.

Second, to explain the variability within CI users, we examined age and device experience as predictors of categorization. All of the users had at least one year device experience and should be well-adapted to their CI, thus we did not expect large effects of experience. However, device experience may capture slower and more subtle differences past the initial adaptation period. Age was included because older adults tend to have poorer acoustic hearing thresholds, but age also impacts central auditory processing (e.g., Bidelman et al., 2014) and may thus especially impact timing cues for stop perception. It is also possible that older age impacts the ability to adapt to input from the CI and could thus hamper integration of acoustic and electric inputs.

Third, we examined degree of residual hearing (for A+E listeners). Our hypothesis was that individuals with better residual hearing over-rely on their acoustic hearing even when they could receive more information through their implant (as in Mok et al., 2006). If this is the case, individuals with poorer residual hearing in the low- to mid-frequency range will show better categorization than those with better residual hearing, an effect that should be highlighted by the fricatives, which functionally isolate the CI only. We specifically address this by asking how residual hearing thresholds (as measured by pure-tone audiometry) impact perception of acoustic cues in stops and fricatives in a group of A+E CI users.

Finally, we sought to relate phonetic categorization to standard clinical measures of speech perception. Where available, we obtained assessment scores from tests of word and sentence recognition. The ability to perceive fine-grained acoustic detail is likely beneficial to the process of recognizing words, so we expected listeners with steeper categorization to also perform better in standard clinical assessments. Steeper categorization of phonetic continua, in this case, reflects more definitive (i.e., less noisy) perception, which in turn could benefit the mapping of speech sounds to words.

Unlike prior studies, we did not test listeners in different listening configurations (e.g., with and without their hearing aid). This was for three reasons. First, the huge variety of device configurations made standardization of this difficult; for example, some hybrid listeners have residual acoustic hearing in both ears, while others use only one. Second, we wanted to ensure that all listeners performed the task in their most familiar configuration. Finally, our primary goal was a between-subjects comparison among people using their everyday listening configuration.

Materials and Methods

Participants

CI users were recruited through the Cochlear Implant Research Center at the Department of Otolaryngology, University of Iowa Hospitals and Clinics. 147 post-lingually deafened CI users completed this task as part of a longer visit to the lab. 11 participants were excluded because their data were incomplete, leaving 136 participants (69 female, 67 male). All recruitment and experimental protocols were approved by the University of Iowa Institutional Review Board.

The CI users included a range of device configurations. Because an individual's residual acoustic hearing could deteriorate over time, we used the participant's pure tone average (PTA) on the test day to determine which device configuration A+E subjects should be classified in. Low-frequency PTA was the average of thresholds at 0.25, 0.5, 1 and 1.5 kHz. Full PTA was the average of thresholds at .25, .5, 1, 1.5, 2, 4, and 8 kHz. We used 85 dB HL as a cut-off for recategorizing participants. Hybrid users with a low frequency PTA poorer than 85 dB on their ipsilateral side were recategorized depending on the status of their contralateral ear. If they were a bilateral hybrid user, they would remain categorized as a hybrid user, as long as one ear passed the PTA criterion. If the hybrid user had a contralateral hearing aid (and that ear still passed our PTA cut off), they would be categorized as bimodal. Based on this criterion, we recategorized 17 hybrids as bimodal CI users and 1 bilateral hybrid as a bilateral (traditional) CI user. Table 1 summarizes the demographic breakdown of each participant group. The average age of the included participants was 58.7 years old, with a mean device experience of 5.0 years.

A complete list of participant characteristics, including information about their hearing and their demographic details, is provided in the OSF repository associated with this project (https://osf.io/z5e3y/).

Table 1. Summary of participant demographics.

Group	N	Mean Age (SD)	Mean Device Experience (SD) in years
Electric-Only			
Unilateral	18	58.5 (11.2)	12.8 (9.1)
Bilateral	23	57.0 (13.2)	7.8 (5.2)
Acoustic + Electric			
Bimodal	43	60.4 (10.0)	3.4 (3.4)
Hybrid	25	62.5 (10.2)	2.0 (1.6)
Single-Sided Deafness (SSD)	27	54.3 (12.4)	2.9 (1.7)

Stimuli

We used five minimal pairs for each contrast (10 total). From each pair, we constructed an 8-step continuum. Fricative minimal pairs were *shack*-*sack*, *shave*-*save*, *shelf*-*self*, *ship*-*sip*, and *shock*-*sock*. Stop voicing minimal pairs were *beach*-*peach*, *bear*-*pear*, *bet*-*pet*, *bin*-*pin*, and *bug*-*pug*. We used the same auditory and visual stimuli as in McMurray et al. (2016). Briefly, auditory stimuli were spoken by a male speaker of American English. Fricative continua were created by extracting the long-term spectra of the fricatives, aligning them by their spectral means, and creating intermediate steps that varied in spectral shape. These were then shifted by spectral mean from /s/ to /ʃ/ (see Figure 2, McMurray et al., 2016). Stop voicing continua were created by successively cross-splicing the minimal pairs to make voice onset time (VOT) steps. That is, a section from the onset of the voiced token (e.g., *beach*) was cut and replaced by the onset of the voiceless token (e.g., *peach*).

Visual stimuli were images from a clipart database that had been selected from several options as the most prototypical image for a given target word. Images were edited for clarity and visual style to minimize distractions. All images received final approval by a senior member of the lab with extensive experimental experience.

Design

As in McMurray et al. (2016), listeners heard tokens from a stop-voicing and fricative-place continuum spanning two words (e.g., *beach/peach, sip/ship*), and selected the corresponding picture. While McMurray et al. (2016) also used eye-tracking as a more sensitive measure of sensitivity to within-category structure (and not just categorization), we used only the overt identification measure, as that was sufficient to show group differences in the prior study, and the number of participants we intended to test called for a shorter experiment than would be appropriate for eye-tracking.

Critically, in this 2AFC / speech continuum task "better" categorization is reflected in steeper categorization slopes. This pattern of data is often associated with categorical perception (Liberman et al., 1957), in which sharp categorization is thought to be the ideal and is achieved by ignoring within category variation. We make no such assumptions here; in fact, it is now well understood that gradient sensitivity to within category detail is the norm (Andruski et al., 1994; McMurray et al., 2002; Miller, 1997); and that higher within-category sensitivity can be linked to steeper categorization slopes (McMurray et al., 2018). However, in this case, we see steeper categorization not as indicative of an underlying mechanism of perception, but rather as indicative that listeners are simply more consistent in how they categorize the sounds. We return to this theoretical and methodological debate in the general discussion (and see McMurray, 2022 for further discussion).

Each continuum step was presented 5 times for a total of 400 trials (2 contrasts x 5 continua/contrast x 8 steps x 5 repetitions). The trials were randomized so that stimuli from different continua were interleaved. The visual display contained the two images representing the endpoints of the continuum for that trial and two images for a minimal pair from the other contrast. Thus, images for fricative minimal pairs served as the unrelated images for stop trials and vice versa. The pairing between fricative and stop continua was randomized across subjects, but was fixed within a subject's session (i.e., *shock*-*sock* might be consistently paired with *beach*-*peach* for one subject, but with *bet*-*pet* for another). Image positions were randomized across the four quadrants of the screen on each trial, such that all the possible combinations of target, competitor, and two distractors appeared equally often in each position.

Procedure

Participants were seated in a sound-attenuated booth in front of a 19” computer screen with a resolution of 1024×1280. Auditory stimuli played over loudspeakers approximately 1 meter from the participant. Participants completed the task using their daily device configurations. For SSD listeners, their NH ear was left unobstructed. Participants were first familiarized with the picture-word pairings that would appear in the experiment. Each image was presented in the middle of the computer screen with its label printed in capital letters below. Participants moved through the images at their own pace by pressing the space bar on a keyboard.

Next, the participant received instructions for the experimental task. On each trial, four images were displayed (two for a stop continuum, two for a fricative continuum) with a red circle in the center of the screen. After 500 msec, the circle turned blue, and the participant clicked on it to play a word over the loudspeakers. Participants were instructed to click on the picture that best matched what they heard. The entire task took approximately 30 minutes.

Data Processing

Responses were coded separately for each contrast. For the stops, /p/ responses were coded as 1 and /b/ responses were coded as 0. For the fricatives, /s/ responses were coded as 1 and /ʃ/ responses were coded as 0. We excluded any responses to one of the control items (i.e., selection of a fricative image for a stop auditory target).

One concern is that each participant (or item) may have a different category boundary along the continua; this could vary systematically across hearing configurations and could lead to averaging artifacts that look like an effect on slope (our core measure of phonetic categorization). Thus, to minimize differences in boundary placement (and reveal changes in the *slope* of the boundary), we relativized the continua steps to each participant's boundary prior to analysis. This allows the analysis to focus on how listeners are dealing with the ambiguity of the stimuli by setting everyone's category boundary to zero. Positive steps represent movement towards the /p/ or /s/ end of the continua and negative steps represent movement towards the /b/ or /ʃ/ end for stops and fricatives respectively.

To do this, we first fit a series of logistic regressions to each subjects' individual data using the *glm* function (family=binomial) in R (version 4.1.3; R Core Team, 2022). From this, we computed the boundary for each subject. This boundary value was then subtracted from each continuum step to relativize step to each listener's category boundary (a variable we termed relative step or rStep). If a listener's category boundary was estimated to be too extreme (i.e., fell outside of steps 2 through 7), their data were excluded from further analysis. This resulted in the exclusion of 10 subjects from both the stop-voicing and fricative analyses. We discuss these excluded subjects in Supplement 1.

Analyses

We conducted four sets of analyses that used slightly different subsets of the participants and predictors. First, we started with an initial analysis that examined only the effect of hearing configuration, in order to better understand the overall group differences. Second, we excluded the SSD listeners to conduct our primary analyses that examined demographic and hearing factors as moderators. The third analysis examined only the A+E group to identify the effect of residual acoustic hearing as a moderator. Finally, we related audiological outcomes (performance in two standard tests of speech perception) to individual differences in stop and fricative categorization. Data and the R script to recreate the analyses presented here are available in the OSF repository associated with this project (https://osf.io/z5e3y/).

Effect of Hearing Configuration. To analyze stop-voicing and fricative categorization as a function of hearing configuration, we ran separate logistic mixed effects models for each continuum. Both models used the same structure given in (1), in the notation of the glmer() function in R.

*response ~ rStep * hearing configuration +*
(1 + rStep || participant) + (1 + rStep || continuum) (1)

Here, rStep refers to the relative continuum step (after adjusting for the subject's own boundary). Hearing configuration was coded with a set of four orthogonal contrast codes, that were individually designed to test specific comparisons of interest: a) SSD users (+1) to the rest of the subjects (-0.25), b) E-only (-0.25) users to A+E (+0.25), c) Unilaterals (-0.5) to Bilaterals (+0.5, with all other groups set to 0), and d) Hybrids (-0.5) to Bimodals (+0.5, with all other groups set to 0). The interactions between rStep and the Hearing configuration contrasts were the crucial predictors in this model, as these can reveal if the slope of the categorization function (the effect of rStep) differed across hearing configurations. We included random intercepts and slopes for rStep by subjects and continua.

Effect of Participant Characteristics. As we described in the introduction, a variety of demographic and audiological factors must be accounted for before we can be confident in differences among listening configurations. We thus ran two additional logistic mixed-effects models, again examining stop-voicing and fricatives separately. These models were based on the prior models and included age and device experience as main effects. As there were difficulties getting this model to converge, we dropped the four-way interaction. This model is given in (2).

*response ~ rStep * hearing configuration * age +*
*rStep * hearing configuration * device experience +* (2)
(1 + rStep || participant) + (1 + rStep || continuum)

We excluded the SSD listeners from these models to focus on individuals who had a more typical pattern of hearing loss. Hearing configuration was coded with the same three contrast codes as used above (minus the SSD specific code): a) E-only vs. A+E, b) Unilateral vs. Bilaterals; and c) Hybrids vs. Bimodals. We centered age and device experience. We included the three-way interactions between rStep, age, and hearing configuration and between rStep, device experience, and hearing configuration (along with their component two-way interactions). Again, the crucial interactions are with rStep as these would suggest that the slope of the categorization function is affected by the other predictors. We included random intercepts and slopes by rStep on participants and continua.

Effect of Residual Hearing. The third analysis asked whether categorization was related to the degree of preserved residual hearing. Thus, it was restricted to bimodal and hybrid CI users who have residual acoustic hearing. We again ran separate mixed-effects logistic regressions for stop voicing and fricatives using the model shown in (3).

response ~ *rStep * hearing configuration* better ear low PTA +*
*rStep *age + rStep * device experience +* (3)
(1 + rStep || *participant*) + (1 + *rStep* || *continuum*)

Hearing configuration was captured by a single contrast (hybrid [-0.5] vs. bimodal [+0.5] CI users). We used better ear low-frequency PTA to index residual acoustic hearing. For subjects who had residual hearing in both ears, we took the better (lower) of the two as their threshold. PTA was centered and included as a predictor along with rStep, age and device experience (both centered). All of these variables were allowed to interact with rStep. These models included random intercepts and slopes by rStep for participants and items.

Relationship to Clinical Outcomes. Finally, effects on phonetic categorization are likely to be small; this raises the question of whether they relate to everyday performance. Thus, our final analysis asked the degree to which performance in these speech categorization tasks is related to individual differences in standard clinical assessments of speech perception. We obtained several measures of word and sentence recognition for a large subset of these users: the Consonant-Nucleus-Consonant (CNC) test of word recognition (Peterson & Lehiste, 1962) and the AzBio test of sentence recognition (Spahr et al., 2012) in both quiet and noise (+5 dB

SNR). These tests were completed during participants' yearly audiological visit and were not administered to every participant. We were thus not able to obtain these scores for our full sample, so analyses were carried out on a subset of the data (CNC N = 109, AzBio in quiet N = 61, AzBio in noise N = 62).

Separate linear regressions predicted performance in each test (CNC, AzBio quiet, AzBio noise) as a function of categorization of each phonetic contrast (stops, fricatives). To quantify the slope of each contrast for each subject, we ran two logistic mixed effects models predicting either /p/ response for the stops or /s/ response for the fricatives. We included rStep as a main effect and random intercepts and slopes of rStep on subject and by continuum. We then used each subject's random effect of rStep as an estimate of their stop or fricative categorization (which we term *categorization slope*). The random slope of rStep for each subject reflects how much an individual varies around the group estimate of rStep, and thus provides a measure of their categorization slope. A positive random slope for a given participant reflects a steeper slope than the group average, while negative slopes reflect a shallower than average slope. Note that these models, which were intended solely to estimate each subjects' slope, did not include any of the fixed effects from the above models as this would have converted our measure of speech categorization to a relative measure (how much better was a listener doing than they would be predicted by their age, device experience, PTA, etc.).

The categorization slope from stop voicing and fricatives were moderately correlated ($r = 0.42$, $p < .001$), so we did not calculate a composite effect and instead ran separate regressions for each phonetic contrast. We included these slopes as predictors in the linear regressions, along with the Hearing configuration comparison between E-only (-0.5) and A+E (0.5) listeners, and age (centered). The model formula is given in (4).

*Accuracy ~ categorization slope * hearing configuration + age* (4)

Results

We present our results by phonetic contrast, starting with the stop-voicing contrast followed by the analyses of frication. Within each contrast, we conducted the three analyses described above relating categorization to various hearing and demographic variables. The analyses investigating the relationship to clinical outcomes are presented last.

Stop Voicing Categorization

Effect of Hearing Configuration. Figure 1 shows the mean responses for the voicing contrasts as a function of device configuration (Unilateral, Bilateral, Bimodal, Hybrid, and SSD). Panel A shows the raw data, Panel B shows the data after alignment by the subjects' boundary (the basis of the analyses), and Panel C shows the estimated effect of rStep for each individual subject grouped by device configuration. SSD users displayed the sharpest categorization. This is perhaps not surprising, as they have a completely NH ear. Within the more canonical CI users, A+E users appeared to have sharper categorization than E-only users (solid lines compared to dotted lines, Figure 1B).

Table 2 shows the results for the mixed-effect model investigating hearing configuration for stop-voicing categorization. We found the expected main effect of rStep (B = 1.22, z = 14.91,

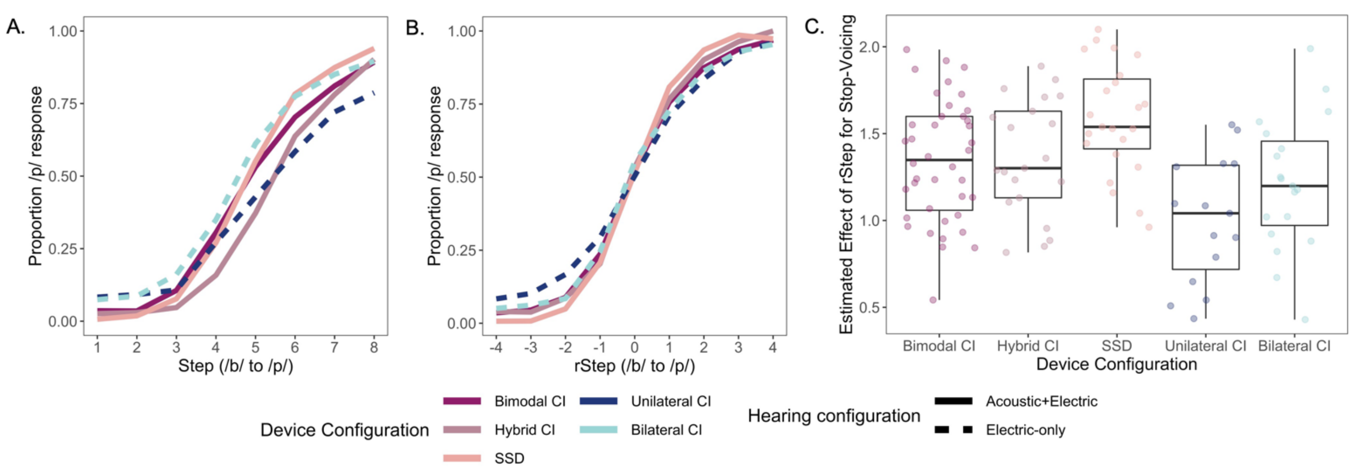

Figure 1. Proportion /p/ response to all /b-p/ continua for each listening configuration by (A) original continuum step and by (B) relative continuum step. C) Estimated Effect of rStep for each individual participant grouped by device configuration. This was estimated by summing the model's fixed effect of rStep, the interaction between rStep and hearing configuration, and each subject's random slope of rStep.

*Table 2. Summary of the mixed effects logistic regression (Equation 1) assessing stop voicing categorization, predicted by rStep and hearing configuration. * p < .05*

Estimate	**B**	**SE**	**z**	**p**	
rStep	1.22	0.08	14.91	< .001	*
Hearing config (SSD vs everyone else)	0.02	0.07	0.26	0.79	
Hearing config (A+E vs E-only)	0.008	0.14	0.06	0.95	
Hearing config (Bi- vs Unilateral)	-0.05	0.11	-0.46	0.64	
Hearing config (Bimodal vs Hybrid)	-0.06	0.09	-0.71	0.48	
rStep x Hearing config (SSD vs everyone else)	0.38	0.08	4.64	< .001	*
...........x Hearing config (A+E vs E-only)	0.44	0.18	2.49	0.01	*
...........x Hearing config (Bi- vs Unilateral)	0.22	0.14	1.55	0.12	
...........x Hearing config (Bimodal vs Hybrid)	-0.08	0.11	-0.69	0.49	

p < .001): as subjects moved from the negative rSteps to the positive rSteps, they were more likely to respond with /p/. Of particular interest, the interaction between rStep and Hearing configuration was significant for SSD subjects compared to everyone else (B = 0.38, z = 4.64, p < .001) and for A+E compared to E-only (B = 0.44, z = 2.49, p = .01). This means that SSD CI users had a steeper slope than the other subjects (a larger effect of rStep) and the A+E CI users had a steeper slope than E-only users. This suggests that the failure to observe the A+E benefit by McMurray et al. (2016) was likely due to reduced power.

Effect of Participant Characteristics. We next examined the same effects after

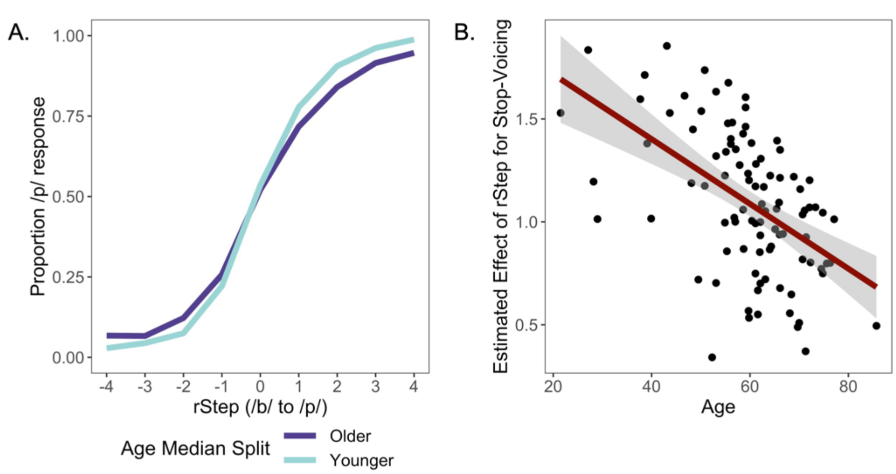

Figure 2. A) Proportion /p/ response by relative continuum step split by median age. B) Estimated effect of rStep by age.

accounting for demographic factors that may be confounded with performance (Table 3, Equation 2 for model). With increasing rSteps, listeners were more likely to respond with /p/ ($B = 1.11$, $z = 12.54$, $p < .001$). The critical interactions are with rStep. As with the initial model, the effect of rStep was significantly different between A+E and the E-only listeners ($B = 0.61$, $z = 2.82$, $p = .004$) with steeper categorization slopes for A+E listeners. There was also a significant interaction between rStep and age ($B = -0.37$, $z = -4.41$, $p < .001$; Figure 2A): as age increased, subjects showed shallower categorization slopes. Figure 2B shows the estimated effect of rStep

Table 3. Summary of a logistic mixed effects model (Equation 2) assessing stop voicing categorization, predicted by rStep, hearing configuration, age, and device experience.

Effect	**B**	**SE**	**z**	**p**	
rStep	1.11	0.09	12.54	< .001	*
Hearing configuration (A+E vs E-only)	-0.07	017	-0.41	0.68	
Hearing configuration (Bi- vs Unilateral)	0.06	0.13	0.49	0.63	
Hearing configuration (Bimodal vs Hybrid)	0.04	0.11	0.39	0.69	
Age	-0.004	0.07	-0.06	0.95	
Device Experience	-0.10	0.08	-1.34	0.18	
Age x Hearing configuration (A+E vs E-only)	-0.46	0.27	-1.75	0.08	
x Hearing configuration (Bi- vs Unilateral)	-0.02	0.19	-0.09	0.93	
x Hearing configuration (Bimodal vs Hybrid)	0.11	0.18	0.62	0.53	
Device Experience x Hearing configuration (A+E vs E-only)	-0.23	0.31	-0.74	0.46	
x Hearing configuration (Bi- vs Unilateral)	-0.33	0.20	-1.60	0.11	
x Hearing configuration (Bimodal vs Hybrid)	0.44	0.23	1.87	0.06	
rStep x Hearing configuration (A+E vs E-only)	0.61	0.22	2.82	0.004	*
.... x Hearing configuration (Bi- vs Unilateral)	0.09	0.17	0.53	0.59	
.... x Hearing configuration (Bimodal vs Hybrid)	-0.17	0.14	-1.24	0.21	
.... x Age	-0.37	0.08	-4.41	< .001	*
.... x Device Experience	0.06	0.09	0.59	0.56	
rStep x Age x Hearing configuration (A+E vs E-only)	0.52	0.34	1.55	0.12	
....x Hearing configuration (Bi- vs Unilateral)	0.15	0.25	0.60	0.55	
....x Hearing configuration (Bimodal vs Hybrid)	0.14	0.23	0.63	0.53	
rStep x Device Exp. x Hearing configuration (A+E vs E-only)	-0.07	0.39	-0.19	0.85	
....x Hearing configuration (Bi- vs Unilateral)	0.31	0.26	1.20	0.23	
....x Hearing configuration (Bimodal vs Hybrid)	-0.21	0.3	-0.72	0.47	

for each subject, calculated as the sum of the fixed effect of rStep, the rStep × age interaction and the participants' random effect of rStep. It is evident that older CI users show a smaller effect of rStep.

Effect of Residual Acoustic Hearing. For the model investigating residual hearing for stop-voicing categorization (Table 4), we again found the expected main effect of rStep (B = 1.27, z = 14.71, $p < .001$). There was a significant interaction between Hearing configuration (Hybrid vs. Bimodal) and PTA (B = 0.38, z = 2.37, p = .01), suggesting that the Hybrid CI users have a larger effect of PTA overall. Note that, because this is not an interaction with rStep, this represents a difference in the placement of the category boundary (as all effects not including rStep should be interpreted relative to rStep = 0). That is, the Hybrid listeners with poorer PTA are more likely to have a boundary shifted towards the /p/-end of the continuum.

We are again most interested in the interactions with rStep. The only significant interaction with rStep was age (B = -0.25, z = -2.15, p = .03). This is consistent with our previous analysis: older CI users have shallower slopes while categorizing stop voicing compared to younger CI users. The lack of an interaction of rStep and PTA is consistent with

Table 4. Summary of a logistic mixed effects model (Equation 3) assessing stop voicing categorization, predicted by rStep, residual acoustic hearing, age, and device experience for A+E CI users.

Effect	**B**	**SE**	**z**	**p**	
rStep	1.27	0.09	14.71	< 0.001	*
Hearing configuration (Hybrid vs Bimodal)	-0.03	0.08	-0.32	0.75	
Age	-0.07	0.08	-0.83	0.41	
Device Experience	0.001	0.08	0.01	0.98	
Better ear PTA	-0.12	0.08	-1.51	0.13	
Hearing configuration (Hybrid vs Bimodal) x Better ear PTA	0.38	0.16	2.37	0.01	*
rStep x Hearing configuration (Hybrid vs Bimodal)	-0.12	0.11	-1.07	0.29	
.... x Age	-0.25	0.11	-2.15	0.03	*
.... x Device Experience	-0.04	0.12	-0.32	0.75	
.... x Better ear PTA	0.06	0.11	0.57	0.57	
rStep x Hearing config. (Hybrid vs Bimodal) x Better ear PTA	-0.32	0.21	-1.49	0.14	

Table 5. Summary of a mixed effects regression assessing fricative categorization (Equation 1), predicted by rStep and hearing configuration.

Estimate	B	SE	Z	p	
rStep	1.23	0.08	14.53	< .001	*
Hearing config (SSD vs everyone else)	-0.03	0.06	-0.49	0.62	
Hearing config (A+E vs E-only)	0.10	0.14	0.71	0.48	
Hearing config (Bi- vs Unilateral)	-0.17	0.11	-1.55	0.12	
Hearing config (Bimodal vs Hybrid)	0.11	0.09	1.19	0.23	
rStep x Hearing config (SSD vs everyone else)	0.38	0.09	4.13	< .001	*
.... x Hearing config (A+E vs E-only)	0.03	0.21	0.14	0.89	
.... x Hearing config (Bi- vs Unilateral)	0.42	0.16	2.53	0.01	*
.... x Hearing config (Bimodal vs Hybrid)	0.24	0.13	1.82	0.07	

the idea that as long as listeners have some residual hearing, they can reap the benefits of acoustic hearing, but it may not matter how much (at least for voicing categorization).

Stop-Voicing Summary. Our first analysis suggests that SSD users perform better than other CI users in stop categorization—even A+E CI configurations (hybrid and bimodal). This is unsurprising given that they have a fully NH ear. Moreover, unlike McMurray et al. (2016), we also found that other A+E CI users (bimodal and hybrid) showed steeper stop-voicing categorization than E-only users. Critically, we confirmed the benefit of acoustic hearing for stop-voicing – even after controlling for confounding factors like age and device experience. Finally, these analyses point to the importance of age for stop-voicing categorization. We consistently found that increasing age impacts the categorization slope for stop voicing. We did not find any evidence that the *amount* of residual hearing impacts stop-voicing categorization, suggesting that the presence of residual acoustic information is enough to provide some benefit.

Fricative Categorization

Effect of Hearing Configuration. As with voicing, we started by examining the overall effect of hearing configuration for fricative categorization (Table 5, Figure 3). We again found the expected main effect of rStep ($B = 1.23$, $z = 14.53$, $p < .001$), subjects responded more with

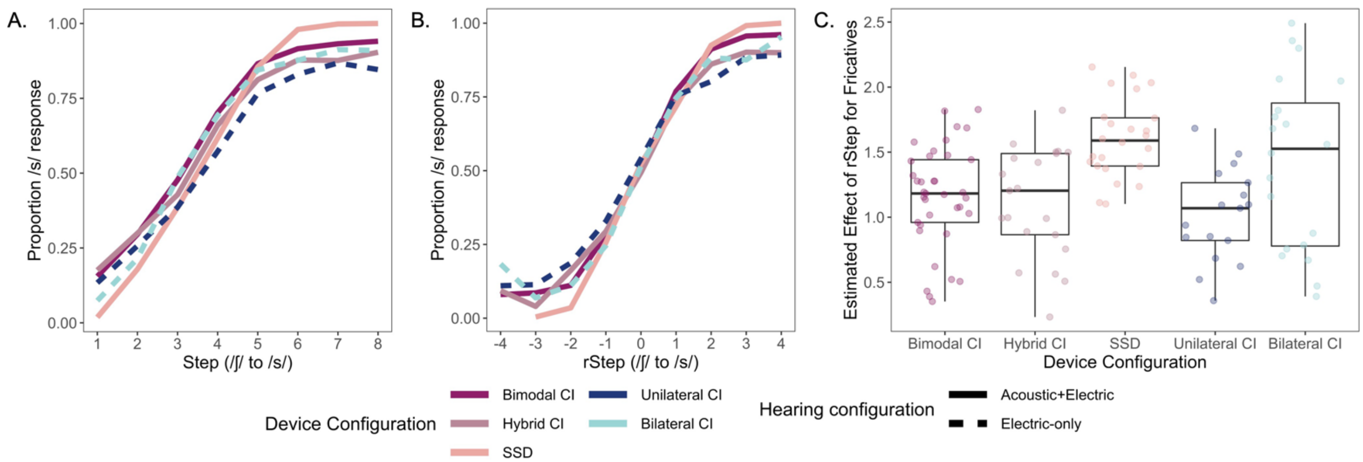

Figure 3. Proportion /s/ response to all /ʃ-s/ continua for each listening configuration by (A) original continuum step and by (B) relative continuum step. C) Estimated Effect of rStep for each individual participant grouped by device configuration.

/s/ as they moved from the negative to the positive ends of the continuum. The crucial effects were the interaction between rStep and Hearing configuration. Not surprisingly, SSD subjects had a steeper slope than other CI users (B = 0.38, z = 4.13, $p < .001$). There was no significant difference between A+E and E-only users, but within the E-only group, the bilateral users had a steeper slope than unilateral (B = 0.42, z = 2.53, p = .01).

Effect of Participant Characteristics. Table 6 presents the summary of the model examining the effect of hearing configuration after accounting for demographic factors. The expected main effect of rStep was significant (B = 1.01, z = 10.78, $p < .001$); again, subjects were more likely to respond /s/ as the stimulus moved from the negative to the positive rSteps. There was also a significant main effect of age (B = -0.13, z = -2.11, p = .03)— younger listeners had a boundary shifted towards /s/ compared to older listeners.

Turning to the critical interactions with rStep, we found a significant interaction between rStep and the Bilateral vs. Unilateral hearing configuration comparison (B = 0.39, z = 1.99, p = .04), suggesting that Bilateral CI users have steeper fricative categorization. Both age and device experience also interacted with rStep for the A+E vs. E-only listener comparison. Younger E-only listeners had steeper slopes than older E-only listeners, while A+E listeners do

Table 6. Summary of a logistic mixed effects model assessing fricative categorization (Equation 2), predicted by rStep, hearing configuration, age, and device experience.

Effect	B	SE	z	p	
rStep	1.01	0.09	10.78	< .001	*
Hearing configuration (A+E vs E-only)	0.16	0.15	1.06	0.29	
Hearing configuration (Bi- vs Unilateral)	-0.18	0.11	-1.63	0.10	
Hearing configuration (Bimodal vs Hybrid)	0.11	0.09	1.13	0.26	
Age	-0.13	0.06	-2.11	0.03	*
Device Experience	0.04	0.07	0.65	0.52	
Age x Hearing configuration (A+E vs E-only)	0.0005	0.25	0.002	0.99	
...........x Hearing configuration (Bi- vs Unilateral)	0.09	0.17	0.49	0.62	
...........x Hearing configuration (Bimodal vs Hybrid)	-0.19	0.19	-1.02	0.31	
Device Experience x Hearing configuration (A+E vs E-only)	-0.21	0.28	-0.78	0.43	
...........x Hearing configuration (Bi- vs Unilateral)	0.13	0.18	0.74	0.46	
...........x Hearing configuration (Bimodal vs Hybrid)	0.003	0.21	0.02	0.98	
rStep x Hearing configuration (A+E vs E-only)	0.35	0.26	1.34	0.18	
...........x Hearing configuration (Bi- vs Unilateral)	0.39	0.2	1.99	0.04	*
...........x Hearing configuration (Bimodal vs Hybrid)	0.20	0.17	1.16	0.25	
...........x Age	-0.11	0.11	-0.99	0.32	
...........x Device Experience	0.22	0.12	1.79	0.07	
rStep x Age x Hearing configuration (A+E vs E-only)	1.14	0.43	2.65	0.008	*
.... x Hearing configuration (Bi- vs Unilateral)	-0.18	0.3	-0.62	0.53	
.... x Hearing configuration (Bimodal vs Hybrid)	-0.41	0.31	-1.30	0.19	
rStep x Device Exp. x Hearing configuration (A+E vs E-only)	-1.11	0.49	-2.28	0.02	*
.... x Hearing configuration (Bi- vs Unilateral)	0.31	0.32	0.98	0.34	
.... x Hearing configuration (Bimodal vs Hybrid)	-0.46	0.37	1.26	0.21	

not show a large effect of age on their slopes (B = 1.14, z = 2.65, p = .008; Figure 4). E-only listeners with longer device experience also had steeper slopes than E-only with less experience, while A+E listeners did not show a large effect of device experience on their slopes (B = -1.11, z = -2.28, p = .02; Figure 5).

Effect of Residual Acoustic Hearing. Finally, we investigated the role of residual acoustic hearing for fricative categorization (Table 7). This showed the expected main effect of rStep (B = 1.12, z = 11.68, p < .001). There was also a main effect of age (B = -0.16, z = -2.14,

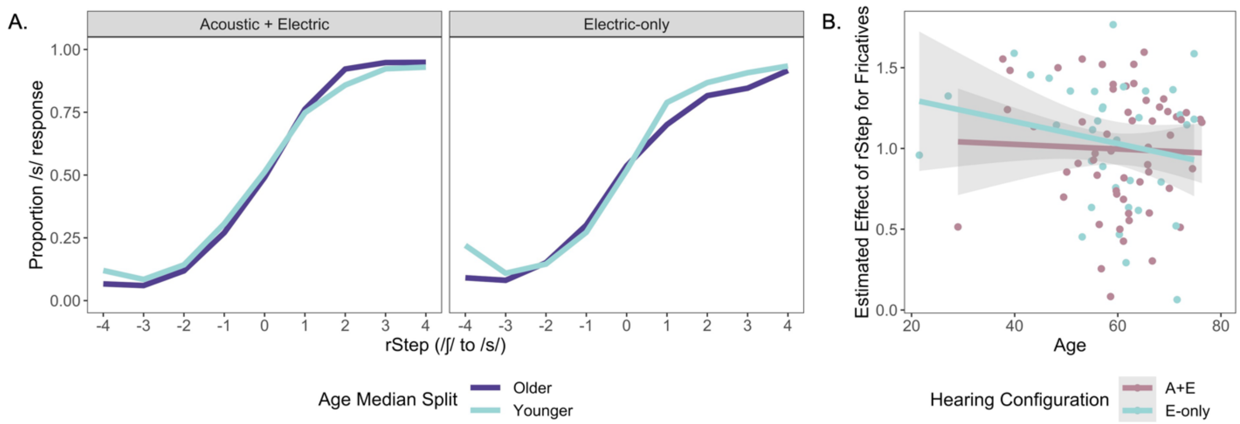

Figure 4. A) Proportion /s/ response by relative continuum step for each hearing configuration (Acoustic+Electric vs. Electric-only) split by median age. B) Estimated effect of rStep by age and hearing configuration

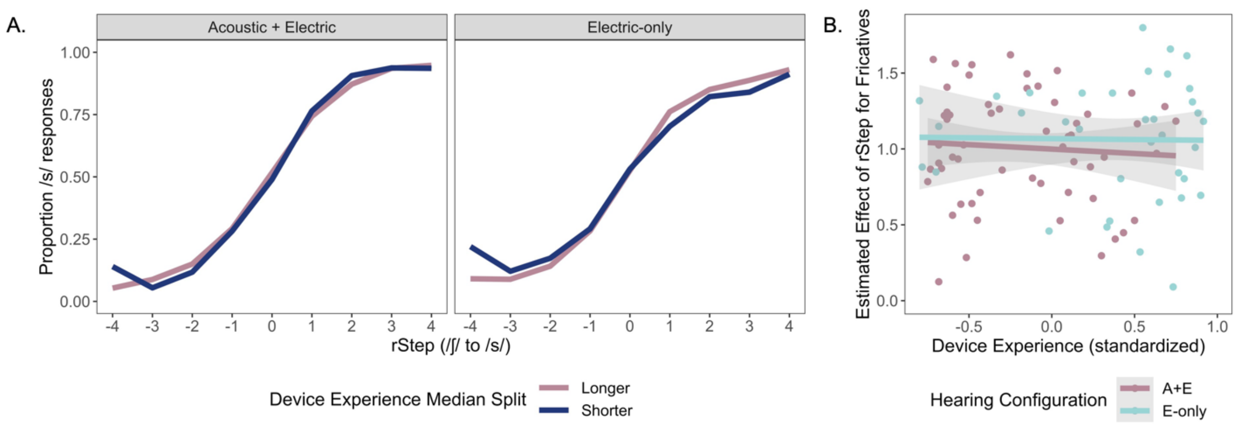

Figure 5. A) Proportion /s/ response by relative continuum step for each hearing configuration (Acoustic+Electric vs. Electric-only) split by median device experience. B) Estimated effect of rStep by device experience and hearing configuration

Table 7. Summary of a logistic mixed effects model assessing fricative categorization (Equation 3), predicted by rStep, residual acoustic hearing, age, and device experience for A+E CI users.

Effect	**B**	**SE**	**z**	**p**	
rStep	1.12	0.1	11.68	< 0.001	*
Hearing configuration (Hybrid vs Bimodal)	0.012	0.08	1.51	0.13	
Age	-0.16	0.07	-2.14	0.03	*
Device Experience	-0.05	0.08	-0.61	0.53	
Better ear PTA	-0.07	0.08	-0.92	0.36	
Hearing configuration (Hybrid vs Bimodal) x Better ear PTA	-0.18	0.16	-1.11	0.27	
rStep x Hearing configuration (Hybrid vs Bimodal)	0.32	0.13	2.54	0.01	*
.... x Age	0.05	0.11	0.47	0.64	
.... x Device Experience	-0.23	0.12	-1.88	0.06	
.... x Better ear PTA	-0.09	0.12	-0.78	0.43	
rStep x Hearing config. (Hybrid vs Bimodal) x Better ear PTA	-0.37	0.25	-1.50	0.13	

there was a significant interaction with rStep. Bimodal listeners had steeper slopes than hybrid listeners (rStep x Hearing configuration: B = 0.32, z = 2.54, p = .01).

There was no significant interaction between rStep and low frequency PTA in this model, however, we wanted to further investigate what might be impacting hybrid listeners' shallower slopes. Thus, we ran an additional exploratory model with just these listeners. This logistic mixed-effects regression predicted /s/ responses with fixed effects of rStep, age, device experience, and ipsilateral low-frequency PTA, as well as the interactions of the latter three variables with rStep. Random intercepts and slopes for rStep by subject and continuum were included. This model is reported in Table 8. We again found a main effect of rStep (B = 0.97, z = 10.28, p < .001). We also found an interaction between rStep and ipsilateral PTA (B = 0.54, z = 3.23, p = .001), suggesting that as PTA increases, so does the effect of rStep (Figure 6). In other words, because a larger PTA implies higher thresholds and worse residual hearing, this suggests that hybrid CI users with poorer acoustic hearing have steeper slopes than those with better residual hearing in their ipsilateral ear.

Fricative Summary. In fricatives, we did not find a large group difference based on the availability of acoustic input (A+E vs. E-only). However, we did find differences within the subgroup comparisons (SSDs had steeper slopes than other CI users, bilateral CI users had steeper slopes than unilateral, and in the standalone A+E analysis, bimodal CI users had

Table 8. Summary of a logistic mixed effects model assessing fricative categorization predicted by rStep, residual acoustic hearing, age, and device experience for hybrid CI users.

Effect	B	SE	z	p	
rStep	0.97	0.09	10.28	< 0.001	*
Age	-0.007	0.12	-0.06	0.95	
Device Experience	0.03	0.12	0.20	0.84	
Ipsilateral PTA	-0.14	0.12	-1.16	0.25	
rStep x Age	0.21	0.16	1.35	0.18	
...........x Device Experience	-0.16	0.17	-0.96	0.34	
...........x Ipsilateral PTA	0.54	0.17	3.23	0.001	*

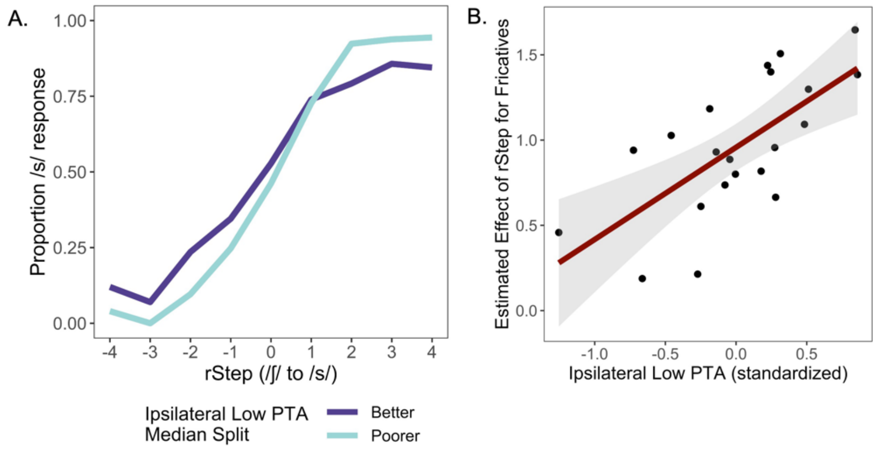

Figure 6. A) Proportion /s/ response by relative continuum step for the Hybrid CI users split by median ipsilateral low PTA. B) Estimated effect of rStep by ipsilateral low PTA.

steeper slopes than hybrids). This is not entirely consistent with McMurray et al.'s prior work. Minimally, it supports the baseline assumption that acoustic hearing would not contribute substantially to fricative performance.

At the same time, however, there were a number of sources of evidence that A+E configurations show a complex response with fricatives. We found group differences between the listeners with residual hearing. Bimodal listeners had steeper slopes than hybrid listeners. In an additional analysis focusing on the hybrid listeners, we found that individuals with better ipsilateral residual hearing had shallower slopes. Moreover, an analysis of the excluded subjects (who failed to show meaningful fricative categorization at all) in Supplement 1 supports this: for fricatives, 9 out of 10 excluded subjects used an A+E configuration, whereas for stop voicing, A+E listeners comprised 5 of 10 excluded subjects. Thus, maintaining acoustic hearing on the ipsilateral side is not beneficial for fricative perception. It may result in CI users over-relying on their acoustic hearing (where there is no meaningful information for fricatives) instead of taking full advantage of their CI.

When investigating demographic factors, we found that device experience and age affect perception of fricatives for E-only CI users. Younger E-only listeners had steeper slopes, as do E-only listeners with longer device experience. A+E listeners did not show large effects of these demographic variables.

Relationship to Clinical Outcomes

Finally, we related the categorization slopes from the models above to standardized outcomes of speech perception. This analysis predicts speech perception outcomes from categorization slope, hearing configuration and age with separate regressions for each of the two categorization slopes.

For ***CNC word recognition in quiet*** (Table 9), Figure 7 shows accuracy against the estimated effect of rStep for stop voicing (A) and fricatives (B) where a larger effect of rStep corresponds to a steeper response slope. For CNC word recognition, only the interaction between hearing configuration and either categorization slope was significant (Hearing config. x Stop categorization slope: B = -8.21, t(104) = -3.07, p = .003; Hearing config. x Fricative categorization slope: B = -9.36, t(104) = -4.94, p < 0=.001). For both stop-voicing and fricatives, E-only listeners with steeper slopes performed better at CNC word recognition compared to those with shallower slopes, while the opposite was true for A+E listeners.

Table 9. Summary of linear regressions investigating CNC scores (Equation 4), predicted by the random effect of stops or fricatives, hearing configuration, age, and device experience.

A) CNC score x Stops					
Effect	**B**	**SE**	**t**	**p**	
Intercept	69.89	2.82	24.8	< 0.001	*
Categorization slope (Stop voicing)	2.32	1.41	1.64	0.10	
Hearing configuration (E-only vs A+E)	-9.2	5.69	-1.62	0.11	
Age	5.05	4.76	1.06	0.29	
Hearing config. (E-only vs A+E) x Cat. Slope (Stop voicing)	-8.21	2.67	-3.07	0.003	*
B) CNC score x Fricatives					
Intercept	67.35	2.34	28.78	< 0.001	*
Categorization slope (Fricative)	1.49	0.96	1.57	0.12	
Hearing configuration (E-only vs A+E)	-3.80	4.65	-0.82	0.42	
Age	5.83	4.37	1.33	0.19	
Hearing config. (E-only vs A+E) x Cat. Slope (Fricative)	-9.36	1.89	-4.94	< 0.001	*

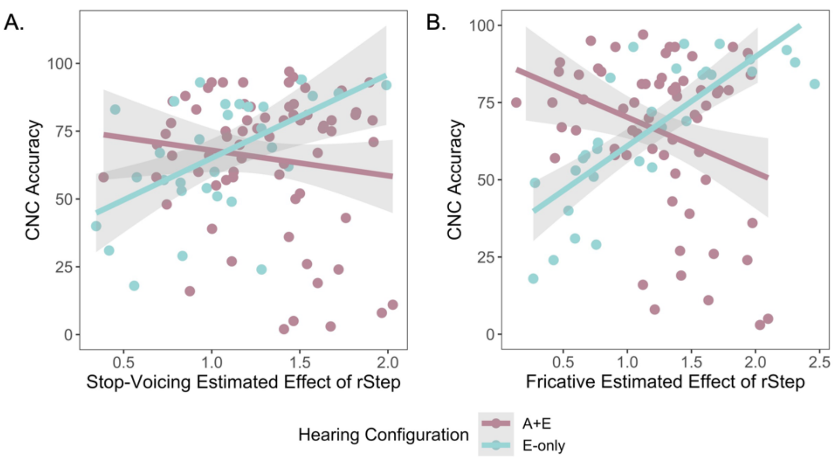

Figure 7. Accuracy on the CNC word recognition task for A+E and E-only CI users by the estimated effect of rStep for A) stop voicing and B) fricatives.

The same relationship was seen for ***AzBio in quiet*** (Table 10; Figure 8). E-only listeners with a larger effect of rStep for both stops (B = -12.63, t(56) = -4.30, p < .001) and fricatives (B = -13.18, t(56) = -4.25, p < .001) scored higher at sentence recognition in quiet compared to those with shallower slopes, while the opposite was true for A+E listeners.

Finally, for ***AzBio in noise*** (Table 11, Figure 9), there was a main effect of the stop categorization slope (B = 6.44, t(57) = 4.39, p < .001) and of hearing configuration (B = 11.83,

Table 10. Summary of linear regressions investigating AzBio in quiet (Equation 4), predicted by the random effect of stops or fricatives, hearing configuration, age, and device experience.

A) AzBio in quiet x Stops					
Effect	**B**	**SE**	**t**	**p**	
Intercept	81.89	3.22	25.45	< 0.001	*
Categorization slope (Stop voicing)	1.32	1.62	0.81	0.42	
Hearing config. (E-only vs A+E)	-12.21	6.26	-1.95	0.06	
Age	6.12	5.9	1.04	0.30	
Hearing config. (E-only vs A+E) x Cat. Slope (Stop)	-12.63	2.94	-4.30	< 0.001	*
B) AzBio in quiet x Fricatives					
Intercept	80.73	3.10	26.02	< 0.001	*
Categorization slope (Fricative)	-0.7	1.55	-0.45	0.66	
Hearing config. (E-only vs A+E)	-2.22	6.14	-0.36	0.72	
Age	5.19	5.38	0.96	0.34	
Hearing config. (E-only vs A+E) x Cat. Slope (Fricative)	-13.18	3.10	-4.25	< 0.001	*

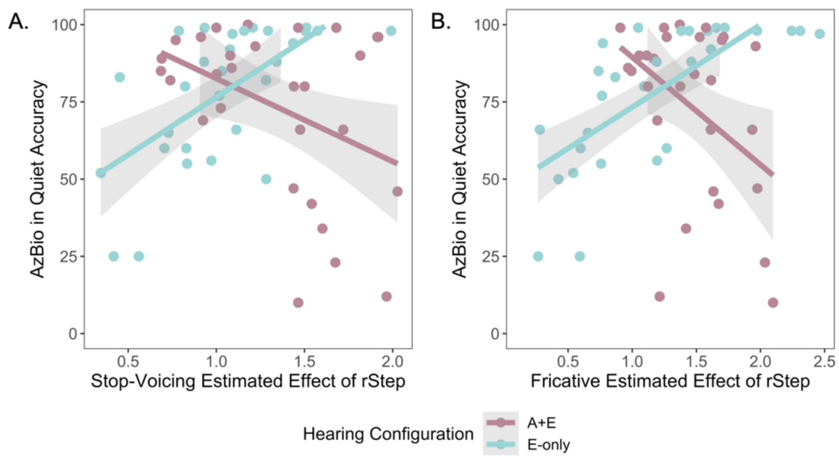

Figure 8. Accuracy for AzBio in quiet for A+E and E-only CI users by the estimated effect of rStep for A) stop voicing and B) fricatives.

$t(57) = 2.36$, $p = .02$), suggesting that the CI users with steeper stop slopes also performed better at sentence recognition in noise and A+E listeners performed better than E-only listeners. In the fricative model, the same main effects were significant (fricative categorization slope: $B = 3.38$, $t(57) = 3.47$, $p = .001$; hearing configuration: $B = 21.95$, $t(57) = 4.45$, $p < .001$). There was also an interaction between hearing configuration and fricative categorization slope ($B = -5.36$, $t(57) = -2.75$, $p = .008$), suggesting that the E-only listeners have a larger effect of fricative

Table 11. Summary of linear regressions investigating AzBio in noise (Equation 4), predicted by the random effect of stops or fricatives, hearing configuration, age, and device experience.

A) AzBio in noise x Stops					
Effect	**B**	**SE**	**t**	**p**	
Intercept	51.48	2.48	20.73	< 0.001	*
Categorization slope (Stop voicing)	6.44	1.46	4.39	< 0.001	*
Hearing config. (E-only vs A+E)	11.83	5.03	2.36	0.02	*
Age	3.39	4.74	0.72	0.48	
Hearing config. (E-only vs A+E) x Cat. Slope (Stop voicing)	-5.26	2.83	-1.86	0.07	
B) AzBio in noise x Fricatives					
Intercept	48.45	2.49	19.48	< 0.001	*
Categorization slope (Fricative)	3.39	0.97	3.47	0.001	*
Hearing config. (E-only vs A+E)	21.95	4.93	4.45	< 0.001	*
Age	0.41	4.93	0.08	0.93	
Hearing config. (E-only vs A+E) x Cat. Slope (Fricative)	-5.36	1.95	-2.75	0.008	*

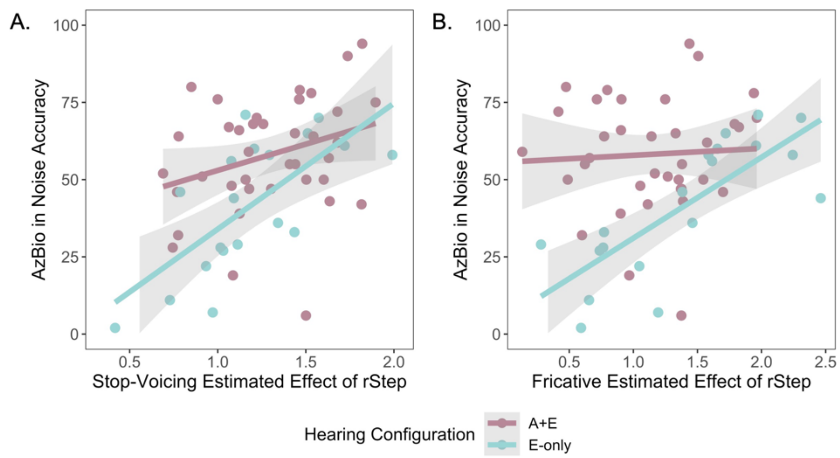

Figure 9. Accuracy for AzBio in noise (+5 dB SNR) for A+E and E-only CI users by the estimated effect of rStep for A) stop voicing and B) fricatives.

random slope.

Summary. These analyses provide evidence that the resolution of phonetic categorization matters for standard clinical assessments, particularly for E-only CI users. E-Only listeners with steeper slopes (i.e., clearer categorization) perform better on standard clinical measures of word and sentence recognition. On the other hand, the A+E listeners show a less clear relationship between their phonetic perception and clinical outcomes. A+E listeners with steeper (i.e., clearer, more defined) slopes in both fricative and stop perception do not necessarily perform better in word and sentence recognition tasks. The lack of a clear effect of categorization slope for both the word and sentence recognition tasks in quiet suggests that when presented with unambiguous tokens, CI users are able to perform well regardless of their attention to fine-grained acoustic ambiguity.

Discussion

The present study examined a large heterogenous group of CI users—both electric-only CI users and those with residual acoustic hearing—to unpack the relationship between acoustic hearing and fine-grained phonetic perception. This led to a number of key findings.

The A+E Benefit

For both stop-voicing and fricative categorization, SSD subjects had steeper categorization than the other groups of CI users. This is unsurprising but it confirms that having one NH ear outperforms listeners with reduced residual acoustic hearing in one or both ears. However, as we discuss, our findings with hybrid listeners suggest that an overreliance on acoustic input in some circumstances could prevent listeners from fully adapting to their implant. It is possible that because some listeners have better residual hearing, they rely on the more familiar acoustic signal, even when it would be beneficial to adapt to the CI input. Thus, it may be fruitful for future work to compare electric-only listeners to SSD CI users with only their CI in similar tasks.

Within the more canonical CI users (those with bilateral hearing loss), we found mixed evidence for a A+E benefit. A+E CI users performed better than E-only users with stop voicing. This benefit was observed even controlling for demographic factors. Unlike prior studies that largely compared A+E listeners with and without their acoustic hearing, this between-subject comparison suggests a robust benefit. The cue to stop voicing is transmitted at the low frequencies of hybrid and bimodal CI user's residual hearing, thus these listeners are benefitting from the maintenance of their acoustic hearing. Importantly, this benefit was not moderated by the amount of low frequency acoustic hearing (PTA). Thus, it may be present even with minimal acoustic hearing (as long as there is some).

In contrast, for fricative categorization, there was no clear benefit for A+E listeners, both as a whole and when demographic factors were accounted for. On the one hand, this was expected—the cues to fricative place of articulation are largely in high-frequency spectral regions. Consequently, the low-frequency residual acoustic hearing retained by hybrid CI users was not expected to be beneficial. On the other hand, this was unexpected given the widespread view that A+E stimulation is universally better. Importantly, we do not fully replicate earlier work showing *worse* performance on fricatives for A+E listeners (McMurray et al., 2016).

As we observe here, there is substantial variability in performance in this task which is in part due to demographic and hearing factors (e.g., length of device use, age). Thus, the smaller sample of A+E listeners in the earlier study may not have been a close match to the electric only users in that study. However, at the same time there is evidence that A+E configurations respond differently to fricatives.

Other Effects of Device Type on Fricative Categorization

For the fricatives, we did find some differences between sub-groups of CI users: bilateral CI users performed better than unilateral. Note that fricatives are often quite a bit lower in amplitude than vowels and stop consonants. In this case, binaural hearing offers a well-known benefit of about 6 dB for detecting quiet sounds (Shaw et al., 1947). This could be particularly helpful for lower amplitude fricatives.

Within the A+E group, we also found evidence for differences. First, most of the excluded subjects (9 out of 10) for fricative categorization were A+E listeners suggesting this configuration may not support good fricative categorization (in the extreme). Second, we found differences in fricative categorization between the bimodal and hybrid listeners. In particular, the hybrid listeners had shallower slopes. Third, within the hybrid listeners, this seemed to be driven by ipsilateral PTA. We suggest that hybrid CI users with better residual hearing may be over-relying on their acoustic input and not taking full advantage of their CI. Whether that is the result of reduced adaptation to the electrical stimulation of their CI (i.e., because they have better residual hearing, these listeners continue to rely on their acoustic input, rather than getting used to the novel input of electrical stimulation) or reduced attention paid to the CI input (as a result of poorer integration between the CI and acoustic input) will require further investigation.

Demographic and Other Effects

Beyond the effects of hearing configuration, we uncovered a number of other relevant effects. First, we found a strong effect of age on stop-voicing categorization: older CI users had shallower stop categorization regardless of their device configuration. Though we did not test NH listeners, this may reflect normal aging processes. One possibility is that this reflects an effect of cognitive aging on the general perceptual/cognitive processes of categorization. We also found that older participants have a boundary shifted towards /ʃ/ when categorizing the fricatives. This provides additional support that age-related perceptual changes are impacting categorization, as age-related hearing loss especially impacts high frequencies sounds (like /s/), which could bias older adults' perception of tokens that fall along an /s/-/ʃ/ continuum towards /ʃ/. However, given that the age x slope interaction for the fricatives is largely driven by E-only listeners, it is possible that maintaining residual acoustic hearing offsets some of the effects of age.

Additionally, this may reflect some more specific perceptual processes that underlie voicing perception. Age affects the use of timing cues, like stop voicing (Toscano & Lansing, 2017), and older adults with age-typical hearing have more asynchronous neural responses to stop voicing cues (i.e., VOT), which is thought to reflect poorer time-locking to the acoustic signal (Bidelman et al., 2014). It follows that older CI users would display the same influence of age on the perception of stop voicing as this seems to be driven by an aging process separate from peripheral hearing ability.

For fricative categorization, we also found that longer device experience predicted steeper categorization, especially in the E-only listeners. It takes time to learn to distinguish the different turbulent, noisy fricative sounds through electric stimulation. Bearing in mind that all the participants in this study had at least 1 year of device experience, this suggests an even more protracted period of adaptation to finding the cues that comprise sibilant fricatives. One possibility is that the A+E listeners' over reliance on the acoustic input may lead them to reach an early plateau in adapting to their CI, while the E-only listeners continue to improve with further experience. This seems to be especially true for sounds that are well encoded through electric stimulation alone as we found this relationship only for fricatives and not for stop voicing.

Relation Between Categorization and Standard Outcomes

Finally, we found evidence for a relationship between categorization slope and performance on standard clinical measures of speech perception, especially in noise. The E-only CI users who have steep categorization slopes also perform well in word and sentence recognition tasks. This validates the utility of strong speech categorization skills as an under-pinning of success with standard CI approaches.

However, these relationships were less clear for the A+E listeners. In fact, for both speech perception measures in quiet, A+E listeners with poorer categorization had better speech perception! This was unexpected, but several potential explanations are worth further consideration. First, it may be that underutilization of the CI leads listeners to down weight individual segmental cues all together in favor of broader envelope or prosodic cues (to sentences). This could disrupt the relationship between phoneme categorization and word or sentence recognition. The dissociation between categorization slope and recognition accuracy for A+E listeners in quiet may provide further evidence for the unconventional relationship between these listeners' hearing ability and their speech processing.

Second, it is now well known that both NH listeners (Andruski et al., 1994; McMurray et al., 2002; Miller, 1997) and CI users (McMurray et al., 2016) are not striving for a discrete representation of phonemes, but a gradient one that may enable them to be more flexible (c.f., Kapnoula et al., 2021; McMurray et al., 2009). In this case, the shallower categorization slope of the A+E listeners may mean something completely different than that of the Electric-only listeners. In the former group, their more fine-grained skills enable them to achieve this gradient representation, whereas in the latter, the shallower slope reflects inconsistency and/or noise. As we describe, this will require more sophisticated measures to detect.

For the sentence recognition in noise task, we found an expected effect for both CI groups. Listeners with steeper categorization slopes also performed better at sentence recognition in noise. It should be noted that this may not be a causal relationship by which better categorization directly contributes to sentence recognition. The perceptual skills that improve phonetic categorization likely also underly talker streaming skills that help with sentence in noise perception. Thus, in this case, improved categorization may simply reflect better perceptual skills, which help with sentence perception (for other reasons).

Categorical Perception

Lastly it is important to consider the fundamental mechanisms of categorization that CI users may be deploying. The 2AFC speech continua task used here is classically interpreted under the assumption that categorization should be reasonably sharp and that any deviation from that is suboptimal and will appear as a shallower slope (see Apfelbaum et al., 2022 for a critique).

This view is classically linked to categorical perception (Liberman, 1957) which further posits that categorization warps the perception of continuous cues like VOT or frication spectra. As a model of perception this has not held up to empirical scrutiny (Massaro & Cohen, 1983; Schouten et al., 2003; Toscano et al., 2010); indeed, it is now thought that a more gradient mode of perception may be more functional for listeners by allowing them to preserve flexibility (Kapnoula et al., 2021; McMurray et al., 2009; Miller, 1997).

How does this more modern understanding impact our interpretation of the present study? It is important to point out that even if perception is gradient, some listeners may nonetheless be more inconsistent or noisy than others. Indeed, that is the most likely cause of the electric vs. A+E differences observed in stop-voicing categorization. Importantly, a crucial limitation of this paradigm is that a shallow slope could arise from a situation in which the listener is striving for a steep categorization but is inconsistent near the boundary. However, It could also arise from a situation in which a listener is trying to be gradient to cope with or reflect uncertainty in the system (Clayards et al., 2008). This could be the case for fricatives -- A+E listeners are more uncertain because they don't have their acoustic hearing to support them. This could also then explain the reversed relationship between speech categorization and outcomes in the A+E listeners – for them a shallower slope is a good thing.

Several alternative paradigms may help. For example, eye-tracking in the visual world paradigm can examine commitment to each option (/b/ vs. /p/) conditioned on the final response (e.g., all the trials where the listener chose /b/). This has shown that CI users are in fact similarly sensitive to fine-grained differences in VOT (relative to NH listeners), despite showing shallower categorization slopes. This suggests these shallow slopes may be due to noise. Second, newer continuous rating scale tasks may be able to disentangle these. In these tasks, listeners make a continuous rating as to the degree to which a speech token is a /b/ or /p/. Here, the variation around the mean can be informative: listeners who are truly gradient may show individual trial responses tightly clustered around the mean, whereas listeners who are noisier may show much more variation (see Apfelbaum et al., 2022 for review; Kapnoula et al., 2017).

Future work may be able to harness these theoretical and methodological advances to better understand the way that A+E hearing alters speech categorization. Critically, fricatives, which uniquely do not contain substantial information in the range of the acoustic hearing may be a fertile ground for such investigations.

Conclusions

Listeners must be able to cope with phonetic ambiguity in the speech signal to successfully map sounds to words. CI users must adapt to novel electric stimulation, and for those with residual acoustic hearing, they must integrate acoustic and electric stimulation. Our results suggest that CI users with residual acoustic hearing can more clearly resolve phonetic ambiguity than listeners who receive only electric stimulation in some situations, such as stop-voicing. However, other times, they may rely on acoustic hearing when it is not appropriate to do so, such as in fricatives. Ensuring that CI users with residual acoustic hearing adapt as much as possible to their CI and are able to integrate acoustic and electric stimulation may lead to improved language outcomes for these listeners.

Acknowledgements

The authors would like to thank Kristin Rooff, Sarah Plock, Francis Smith, and Ashley Farris-Trimble for help in developing and conceptualizing the task and collecting the data.

A preprint of this manuscript is available at https://psyarxiv.com/5tuje.

Author contribution: S.C. analyzed the data and wrote the manuscript. M.S. designed the experiment and collected the data. B.M. developed the experiment and provided critical revision. The authors discussed results and implications for findings throughout the manuscript writing process.

References

Andruski, J. E., Blumstein, S. E., & Burton, M. (1994). The effect of subphonetic differences on lexical access. *Cognition*, *52*(3), 163–187. https://doi.org/10.1016/0010-0277(94)90042-6

Apfelbaum, K. S., Kutlu, E., McMurray, B., & Kapnoula, E. C. (2022). *Don't Force It! Gradient speech categorization calls for continuous categorization tasks*. https://psyarxiv.com/7w93f/

Bidelman, G. M., Villafuerte, J. W., Moreno, S., & Alain, C. (2014). Age-related changes in the subcortical-cortical encoding and categorical perception of speech. *Neurobiology of Aging*, *35*(11), 2526–2540. https://doi.org/10.1016/j.neurobiolaging.2014.05.006

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition*, *108*(3), 804–809. https://doi.org/10.1016/j.cognition.2008.04.004

Dorman, M. F., & Gifford, R. H. (2010). Combining acoustic and electric stimulation in the service of speech recognition. *International Journal of Audiology*, *49*(12), 912–919. https://doi.org/10.3109/14992027.2010.509113

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

Dorman, M. F., Spahr, A. J., Ryan, A., & Wilson, S. (2006). Speech perception by adults with multichannel cochlear implants. In S. B. Waltzman & T. Roland (Eds.), *Cochlear Implants* (2nd Editio, Vol. 1, Issue 1, pp. 194–203).

Dunn, C. C., Tyler, R. S., & Wilt, S. A. (2005). Benefit of wearing a hearing aid on the unimplanted ear in adult users of a cochlear implant. *Journal of Speech, Language, and Hearing Research*, *48*(3), 668–680. https://doi.org/10.1044/1092-4388(2005/046)

Gantz, B. J., Turner, C., Gfeller, K. E., & Lowder, M. W. (2005). Preservation of hearing in cochlear implant surgery: Advantages of combined electrical and acoustical speech processing. *Laryngoscope*, *115*(5), 796–802. https://doi.org/10.1097/01.MLG.0000157695.07536.D2

Gfeller, K., Turner, C., Oleson, J., Zhang, X., Gantz, B., Froman, R., & Olszewski, C. (2007). Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise. *Ear and Hearing*, *28*(3), 412–423. https://doi.org/10.1097/AUD.0b013e3180479318

Kapnoula, E. C., Edwards, J., & McMurray, B. (2021). Gradient activation of speech categories facilitates listeners' recovery from lexical garden paths, but not perception of speech-in-noise. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(4), 578–595. https://doi.org/10.1037/xhp0000900

Kapnoula, E. C., Winn, M. B., Kong, E. J., Edwards, J., & McMurray, B. (2017). Evaluating the sources and functions of gradiency in phoneme categorization: An individual differences approach. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(9), 1594–1611. https://doi.org/10.1037/xhp0000410

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*(5), 358–368. https://doi.org/10.1037/h0044417

Massaro, D. W., & Cohen, M. M. (1983). Categorical or continuous speech perception: A new test. *Speech Communication*, *2*(1), 15–35. https://doi.org/10.1016/0167-6393(83)90061-4

McMurray, B. (2022). *The myth of categorical perception*. https://psyarxiv.com/dq7ej/

McMurray, B., Danelz, A., Rigler, H., & Seedorff, M. (2018). Speech categorization develops slowly through adolescence. *Developmental Psychology*, *54*(8), 1472–1491. https://doi.org/10.1037/dev0000542

McMurray, B., Farris-Trimble, A., Seedorff, M., & Rigler, H. (2016). The effect of residual acoustic hearing and adaptation to uncertainty on speech perception in cochlear implant users: Evidence from eye-tracking. *Ear and Hearing*, *37*(1), e37–e51. https://doi.org/10.1097/AUD.0000000000000207

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, *86*, B33–B42.

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2009). Within-category VOT affects recovery from “lexical” garden-paths: Evidence against phoneme-level inhibition. *Journal of Memory and Language*, *60*(1), 65–91. https://doi.org/10.1016/j.jml.2008.07.002

Miller, J. L. (1997). Internal Structure of Phonetic Categories. *Language and Cognitive Processes*, *12*(5–6), 865–870. https://doi.org/10.1080/016909697386754

Mok, M., Grayden, D., Dowell, R. C., & Lawrence, D. (2006). Speech perception for adults who use hearing aids in conjunction with cochlear implants in opposite ears. *Journal of Speech, Language, and Hearing Research*, *49*(2), 338–351. https://doi.org/10.1044/1092-4388(2006/027)

Peterson, G. E., & Lehiste, I. (1962). Revised CNC Lists for Auditory Tests. *Journal of Speech and Hearing Disorders*, *27*(1), 62–70. https://doi.org/10.1044/jshd.2701.62

R Core Team. (2022). *R: A language environment for statistical computing* (4.1.3). R Foundation for Statistical Computing.

Schouten, B., Gerrits, E., & Van Hessen, A. (2003). The end of categorical perception as we know it. *Speech Communication*, *41*(1), 71–80. https://doi.org/10.1016/S0167-6393(02)00094-8

Shaw, W. A., Newman, E. B., & Hirsh, I. J. (1947). The Difference between Monaural and Binaural Thresholds. *The Journal of the Acoustical Society of America*, *19*(4), 734–734. https://doi.org/10.1121/1.1916604

Spahr, A. J., Dorman, M. F., Litvak, L. M., Van Wie, S., Gifford, R. H., Loizou, P. C., Loiselle, L. M., Oakes, T., & Cook, S. (2012). Development and Validation of the AzBio Sentence Lists. *Ear & Hearing*, *33*(1), 112–117. https://doi.org/10.1097/AUD.0b013e31822c2549

Toscano, J. C., & Lansing, C. R. (2017). Age-Related Changes in Temporal and Spectral Cue Weights in Speech. *Language and Speech*, 002383091773711. https://doi.org/10.1177/0023830917737112

Toscano, J. C., McMurray, B., Dennhardt, J., & Luck, S. J. (2010). Continuous Perception and Graded Categorization: Electrophysiological Evidence for a Linear Relationship Between the Acoustic Signal and Perceptual Encoding of Speech. *Psychological Science*, *21*(10), 1532–1540. https://doi.org/10.1177/0956797610384142

Turner, C. W., Gantz, B. J., Vidal, C., Behrens, A., & Henry, B. A. (2004). Speech recognition in noise for cochlear implant listeners: Benefits of residual acoustic hearing. *The Journal of the Acoustical Society of America*, *115*(4), 1729–1735. https://doi.org/10.1121/1.1687425

Turner, C. W., Reiss, L. A. J., & Gantz, B. J. (2008). Combined acoustic and electric hearing: Preserving residual acoustic hearing. *Hearing Research*, *242*(1–2), 164–171. https://doi.org/10.1016/j.heares.2007.11.008

Wilson, B. S., & Dorman, M. F. (2008). Cochlear implants: A remarkable past and a brilliant future. *Hearing Research*, *242*(1–2), 3–21. https://doi.org/10.1016/j.heares.2008.06.005

Woodson, E. A., Reiss, L. A. J., Turner, C. W., Gfeller, K., & Gantz, B. J. (2009). The hybrid cochlear implant: A review. *Cochlear Implants and Hearing Preservation*, *67*, 125–134. https://doi.org/10.1159/000262604

Zhang, T., Dorman, M. F., & Spahr, A. J. (2010). Information from the voice fundamental frequency (F0) region accounts for the majority of the benefit when acoustic stimulation is added to electric stimulation. *Ear and Hearing*, *31*(1), 63–69.

Audiological and Demographic Factors that Impact the Precision of Speech Categorization in Cochlear Implant Users

ONLINE SUPPLEMENT

Supplementary Analysis of Excluded Subjects.

Several subjects were excluded on the basis of having an outlying estimated category boundary. A descriptive analysis of these subjects is carried out here.

Figure S1 and Table S1 contain the subjects excluded from the stop-voicing analysis. These subjects are an equal mix of Electric-only listeners (n=5) and Acoustic + Electric listeners (n=5). They have a diverse range of device experience (M = 8.3 years, SD = 10.5) and are middle- to older-aged (M = 63.8 years, SD = 7.9). For the most part, these subjects show an overwhelming bias to respond /b/ for the majority of the continua. Only at the most extreme end (step 7 & 8) do they begin to respond /p/ (although still not consistently). The one exception to this pattern is the one bilateral listener, who shows a flat response around chance (50% /p/ response) across all steps. This suggests this listener struggled with this task and may have been guessing.

Figure S2 and Table 2 contain the subjects excluded from the fricative analysis. These subjects are mostly Acoustic + Electric listeners (n = 9), with only one Electric-only listener. They have a wide range of device experience (although not as broad as those excluded from the stop voicing analysis, M = 2.0 years, SD = 1.4) and are middle- to older-aged (M = 66.3 years, SD = 11.1). These subjects largely show an /s/ response bias although it is not as extreme as those excluded from the stop-voicing analysis. Given that our main analysis also suggests that listeners with residual acoustic hearing may struggle with fricatives, it is of note that the majority of these subjects have residual acoustic hearing.

Unfortunately, due to missing scores within an already small subset of the data, we are unable to examine the relationship between these excluded subjects' categorization and their accuracy on clinical measures. The scores that were available are included in Tables S1 and S2.

Table S1. Characteristics of subjects excluded from the stop voicing analysis.

Participant ID	Hearing Configuration	Age (years)	Device Experience (years)	CNC Accuracy	AzBio (5 dB SNR)
672	Bimodal	56.0	0.98		
864	Bimodal	65.8	1.0		
839	Hybrid	76.0	1.0	72	21
862	Hybrid	66.7	0.99	83	65
722	SSD	60.0	4.2	4	
1086	Bilateral	71.5	14.1		
692	Unilateral	54.5	6.6	26	1
773	Unilateral	57.11	21.0	89	40
784	Unilateral	57.0	2.0	19	
1093	Unilateral	73.4	31.2	55	

Table S2. Characteristics of subjects excluded from the fricative analysis.

Participant ID	Hearing Configuration	Age (years)	Device Experience (years)	CNC Accuracy	AzBio (5 dB SNR)
600	Bimodal	50.8	0.92	93	75
721	Bimodal	56.3	3.9	76	25
724	Bimodal	58.1	1.1	24	14
873	Bimodal	71.1	0.96	73	
875	Bimodal	75.6	4.1	55	
1078	Bimodal	66.1	1.0		
1079	Bimodal	66.1	2.0	2	
853	Hybrid	85.7	1.0	48	28
855	Hybrid	77.1	0.98	76	68
1006	Unilateral	56.1	4.1	76	

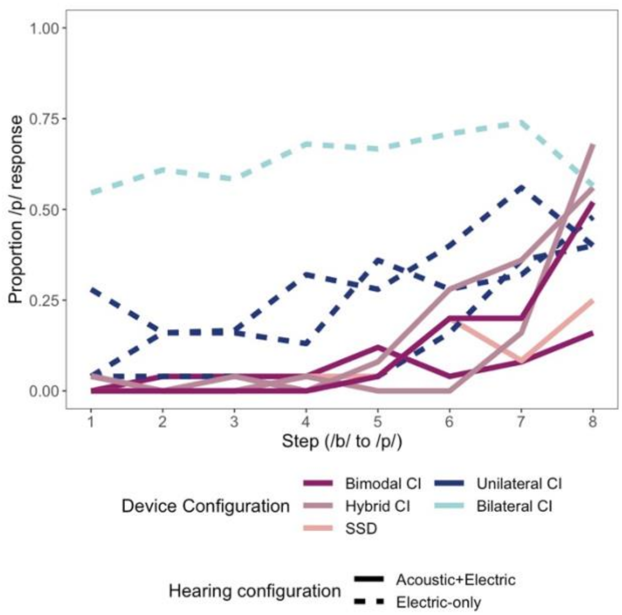

Figure S1. Proportion /p/ response for the excluded subjects by continuum step.

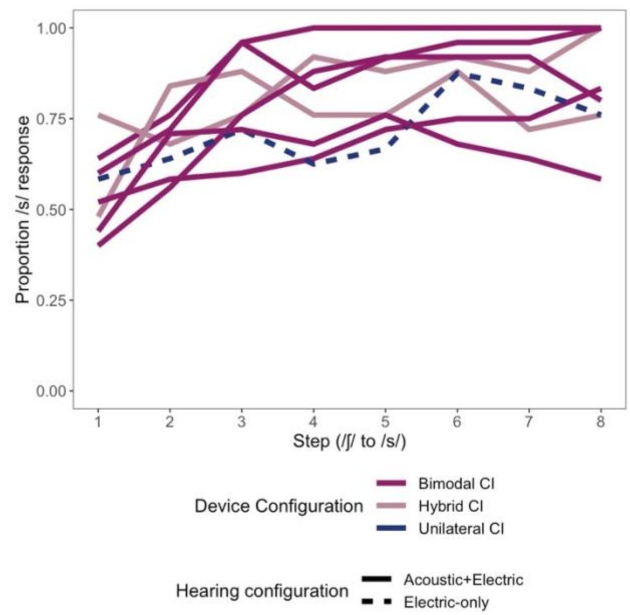

Figure S2. Proportion /s/ response for the excluded subjects by continuum step.