The Enigma of Poor Performance by Adults With Cochlear Implants

*Aaron C. Moberly, *Chelsea Bates, *Michael S. Harris, and †David B. Pisoni

*Department of Otolaryngology, The Ohio State University Wexner Medical Center, Ohio; and †Department of Psychological and Brain Sciences, Indiana University, Indiana

**Objective:** Considerable unexplained variability and large individual differences exist in speech recognition outcomes for postlingually deaf adults who use cochlear implants (CIs), and a sizeable fraction of CI users can be considered "poor performers." This article summarizes our current knowledge of poor CI performance, and provides suggestions to clinicians managing these patients.

**Method:** Studies are reviewed pertaining to speech recognition variability in adults with hearing loss. Findings are augmented by recent studies in our laboratories examining outcomes in postlingually deaf adults with CIs.

**Results:** In addition to conventional clinical predictors of CI performance (e.g., amount of residual hearing, duration of deafness), factors pertaining to both "bottom-up" auditory sensitivity to the spectro-temporal details of speech, and "top-down" linguistic knowledge and neurocognitive functions contribute to CI outcomes.

**Conclusions:** The broad array of factors that contribute to speech recognition performance in adult CI users suggests the potential both for novel diagnostic assessment batteries to explain poor performance, and also new rehabilitation strategies for patients who exhibit poor outcomes. Moreover, this broad array of factors determining outcome performance suggests the need to treat individual CI patients using a personalized rehabilitation approach.

**Key Words:** Adults—Cochlear implants—Sensorineural hearing loss—Speech perception.


It is well known in research and clinical settings that unexplained variability and large individual differences exist in speech recognition outcomes for adults with cochlear implants (CIs) (1–3). This is true even for adults with postlingual deafness whom we would expect to do well, given their previously normal language development. Studies typically focus on group mean performance in quiet or in noise (3), or occasionally consider factors that enable "star" performers to do exceptionally well (4). Unfortunately, there is very little in the literature regarding those patients on the other end of the spectrum: the poor performers. Depending on criteria used to define poor performance, 10 to 50% of adult CI users fall into this category (5). For example, 35 to 50% of CI users cannot use the telephone (6). Lenarz et al. (5) identified 13% of their adult CI users as poor performers, who were able to recognize less than 10% correct words in sentences in quiet. A fundamental gap in our knowledge currently exists regarding the underlying sources of poor performance, and this lack of knowledge directly leads to two major clinical problems: first, we cannot predict when a patient will do poorly with a CI; and, second, we cannot intervene appropriately for poorly performing patients.

Most often in clinical CI centers, a patient who is performing more poorly than generally expected (based on clinical intuition, since we cannot reliably predict outcomes) will undergo a limited diagnostic battery. This battery typically consists of imaging, usually computerized tomography (CT), to ensure the electrode array is in good position; remapping of the device by the audiologist to ensure appropriate stimulation parameters; and a hardware integrity check to confirm that the device itself is functioning normally. This limited battery often does not reveal any problems that can be addressed surgically or clinically. As a result, clinicians are then restricted to reassurance and recommending that the patient "keep working at it." In some settings, a struggling CI patient may be referred to a speech-language pathologist who focuses on aural rehabilitation, although this strategy is limited by a lack of evidence-based methodologies or support from insurance providers. Patients may elect to use one of several "one-size-fits-all" auditory training programs on a home computer, and often seek advice from other CI users or support groups and provided
THE ENIGMA OF POOR CI PERFORMANCE

Anecdotal support for strategies such as use of audiobooks and spending time in challenging listening environments (7). Ultimately, poorly performing patients are frustrated by difficulty understanding speech through their devices, and by their inability to meet outcome expectations. Based on these experiences, some patients even stop using their CIs.

As clinicians and researchers, it is imperative that we develop a better understanding of the sources of poor outcomes in this clinical population. Doing so should help us predict when a patient being evaluated for implantation is at risk for a poor outcome, identify the underlying problem for an individual patient with poor performance with a CI, and develop a personalized aural rehabilitation program for that patient, targeted at specific weaknesses.

The purpose of this article is to discuss the “enigma” of poor CI performance in postlingually deaf adults, and to suggest the roles that clinicians can play. Although not meant as an exhaustive review, we discuss the current state of knowledge regarding variability in performance and poor outcomes in CI users. We briefly discuss conventional clinical measures relating to speech recognition outcomes, and then relevant findings from studies of CI users and patients with lesser degrees of hearing loss, including recent work from our laboratories. We also discuss future work that needs to be undertaken, and, finally, we offer some practical recommendations for clinicians who are treating CI patients with poor speech recognition performance after implantation.

CONVENTIONAL CLINICAL PREDICTORS OF POOR PERFORMANCE

Only a few clinical predictors of speech recognition outcomes have been identified in postlingual adults with CIs. Greater amount of residual hearing before implantation and previous hearing aid use predict better speech recognition outcomes (8–12). Partial insertion of the electrode array, a history of meningitis, and congenital inner ear malformations negatively impact performance (13–14). A longer duration of moderate-to-profound hearing loss experienced by the patient has an important detrimental effect on outcomes (15–18). Performance has been found to be poorer for older adults using CIs (usually defined as over age 65 years) than younger adult CI users, but this finding is not universal (12,19–22). Unfortunately, these reported clinical predictors generally cannot be addressed therapeutically. Moreover, the underlying information processing and neural mechanisms by which these clinical factors affect recognition of speech are unclear.

UNDERSTANDING THE MECHANISMS UNDERLYING POOR OUTCOMES

Beyond these clinical predictors, a number of factors have been identified that may partially explain variability in speech recognition outcomes in postlingually deaf adult patients with CIs, and these factors likely contribute to poor performance for some individuals. While not entirely separate constructs, a useful way to conceptualize these factors is to group them into three broad domains: “auditory sensitivity,” “linguistic skills,” and “neurocognitive functions.” We will examine each of these domains independently, recognizing that abilities within each domain interact with skills within the other domains during the process of spoken language understanding. Moreover, there is good reason to suspect that different factors may contribute to poor performance for different patients.

“BOTTOM-UP” AUDITORY SENSITIVITY

A relatively common assumption in otology is that the variability in speech recognition among adult CI users, and the poor performance by some patients, arises as a direct consequence of variability in the degraded quality of the speech signals listeners receive through their devices. The electrode array of the CI is limited in its ability to provide highly detailed spectral (frequency-specific) information of speech for two main reasons: first, the electrode array cannot be inserted far enough into the cochlea to cover the entire apex, and attempting to do so is traumatic. As a result, the frequency range of the incoming auditory speech input must be allocated to electrodes that do not reach the apex, resulting in spectral mismatch between the acoustic input and the electrode locations inside the cochlea (23–24).

Second, the electrode array has a limited number of stimulating electrodes (usually around 20), but the effective number of independent channels of information presented through an implant is only around four to seven (25). This is a result of spread of excitation from adjacent electrodes, leading to overlapping regions of neural stimulation. Therefore, individuals with CIs hear speech that is both spectrally shifted and spectrally degraded (26–27). One method that has been used to examine spectral resolution is to obtain spectral ripple discrimination thresholds, which have been found to predict 25 to 30% of variability in speech reception thresholds and word recognition for words both in babble and in quiet conditions for adult CI users (28). Likewise, temporal resolution, assessed by amplitude modulation detection thresholds, has been found to explain variability in speech recognition (29).

Third, the electrode array proximity to the modulus is variable among patients, and it is likely that an electrode array that is more tightly coiled around the modulus would provide greater frequency resolution. As measured by electrically evoked compound action potential, there are data supporting the prediction that a shorter electrode-to-modulus distance correlates with higher speech recognition score (30). Additionally, there is some evidence that patients with perimodiolar arrays display better speech recognition than those with lateral-wall hugging electrodes (3).

Lastly, the health of the spiral ganglion cells that are stimulated explains some variability in speech recognition, based on intraoperative electrocochleography.
(ECOG) measurements that serve as a biomarker of peripheral auditory system integrity (31–32). In particular, the ECOG “total response” (an estimate of neural survival) has been found to predict 47% of the variability in consonant-nucleus-consonant (CNC) scores for adults with CIs (32). An area ripe for investigation is the potential influence of peripheral auditory system declines on higher-order, cortical functions: a longer duration of deafness before CI may lead to declines in dendritic, spiral ganglion, and auditory nerve function, but it may also result in detrimental auditory cortex plastic changes that do not automatically reverse following restoration of peripheral input through a CI (33).

**“TOP-DOWN” LINGUISTIC SKILLS**

Although auditory sensitivity factors are historically the most commonly investigated sources of outcome variability for CI users, recent studies in adults with less severe hearing loss have provided clues to top-down sources of variability. Broadly, these top-down factors relate to the linguistic skills and neurocognitive information processing functions of the listener.

Top-down abilities, and their interactions with the incoming speech signal, are widely accepted as crucial in all models of spoken word recognition (34–39). It is generally thought that the transmission of an undistorted signal leads primarily (though not exclusively) to a bottom-up hearing strategy with fast and implicit decoding of linguistic content and sequential lexical access. However, a distorted speech signal, whether due to masking by noise, a hearing impairment, or decoding through a CI (or, experimentally, using vocoded speech), requires top-down mechanisms for explicit decoding of the linguistic content (40). In those cases, successful recognition of speech requires use of lexical and contextual knowledge (41). This linguistic knowledge can include phonological knowledge (sensitivity to the sounds and sound patterns of the language), lexical knowledge (a large vocabulary and familiarity with how the sounds of the language are typically combined), semantic knowledge (an understanding of relationships among words and their meanings), and grammatical skills (knowledge of how phrases and sentences are put together).

Sensitivity to the phonological structure of speech has been found to predict open-set speech recognition in CI users (42). In a recent study, we demonstrated that a single measure of phonemic awareness (final consonant choice) predicted 40% of variability in word recognition in quiet (43). In general, CI users demonstrate deficits in tasks that explicitly require phonological access (e.g., nonword repetition) (Moberly et al., under review). Moreover, knowledge of the phonotactic probabilities of a language—the frequencies with which phonological segments and sequences of segments legally occur—has been found to influence how adult CI users recognize spoken words (44).

When hearing a spoken word under degraded listening conditions, better lexical knowledge should decrease the ambiguity of any given phoneme as a result of the listener’s previous experience with the possible sequences of phonemes within each word (45). When it comes to recognition of words within a sentence, greater lexical knowledge may support better use of lexical connectivity among words (46–49). For example, a larger receptive vocabulary assessed using the Peabody Picture Vocabulary Test (50) has been found to correlate with speech intelligibility scores of sentences in noise for listeners with normal hearing (51). However, for postlingually adult listeners with CIs, word and/or sentence recognition scores have not been found to correlate with scores of receptive vocabulary (52), expressive vocabulary (43), or word familiarity (Moberly et al., under review), suggesting possible differences in linguistic coding and integration processes.

Top-down processes routinely take on larger roles as linguistic context from the speech input increases, or as the speech signal becomes more degraded (53–54). The recognition of individual words automatically triggers semantic information from long-term memory that is helpful in sentence recognition (55–56). Moreover, the semantic and grammatical constraints imposed by the words surrounding a target word should support recognition, as listeners apply their linguistic knowledge to make inferences regarding what is being said and what is likely to occur next. The listener’s knowledge and previous developmental experience helps to segment the continuous acoustic stream into phonemes (individual sound units of the language), syllables, and words (57). For example, there is evidence that better grammatical knowledge correlates with higher recognition scores for words in sentences in adults with CIs (Moberly et al., under review).

Taken together, the above studies suggest that linguistic skills—phonological, lexical, semantic, and grammatical knowledge—contribute to speech recognition outcome variability in adult CI users. Thus, it is likely that poor performance in some adult patients with CIs could be related to declines or deficits in these foundational language skills.

**NEUROCOGNITIVE SKILLS**

Cognition and the information processing skills underlying perception, attention, and memory are increasingly being recognized as important in explaining variability in the speech recognition abilities of adults with lesser degrees of hearing loss, suggesting its impact in adult CI users (58–59). Active and effortful processing of degraded speech clearly places additional information processing demands on a listener’s limited cognitive resources (41,60).

One neurocognitive process in particular—working memory (WM)—has been suggested to play a critical role in compensating for the loss of fine spectro-temporal details in the speech input received by hearing impaired individuals and CI users. Using non-auditory tasks of reading span and visual digit- or letter-monitoring tasks,
measures of working memory capacity have been found to predict 10 to 30% of variability in speech recognition in noise for hearing aid users, though results are not always consistent (61–66); but see (67–68). Converging evidence suggests that a direct relationship exists between increasing background noise (or decreasing signal-to-noise ratio) and reliance on working memory during speech recognition tasks (69–70). Adult listeners' immediate benefit from context in speech recognition relates to their ability to keep and update a semantic representation of the sentence content in WM, providing evidence that WM underlies efficient speech processing and semantic integration of the spoken message (71). Working memory capacity (WMC) has also been associated with release from informational masking by semantically related information (72).

Although verbal WM has received little attention in adult CI users, several studies have examined this cognitive ability in pediatric CI users. It is clear that, in this population, phonological WM plays an important role in speech and language outcomes (73–76). When it comes to adults with CIs, Lyxell et al. (42) found a small but significant correlation between reading span scores before implantation and speech recognition abilities after 12 months of implant experience. Tao and colleagues (77) found significant correlations of speech recognition (Mandarin disyllable recognition) and digit span scores (forward and backward) in adult CI users. A recent study in our lab examined WM in 30 postlingually deaf adult CI users and 30 NH peers (Moberly et al., under review). Performance on WM tasks that explicitly required phonological sensitivity (a task of serial recall of rhyming words and a nonword repetition task) was significantly poorer for CI users than NH peers, and scores on those tasks predicted 14 to 18% of the variability in recognition of words in sentences by CI users. In a follow-up study, we demonstrated that WMC, using an auditory Listening Span task, and grammatical skills both predicted recognition of words in sentences, but in a combined regression analysis, neither contributed independently of the other, providing further evidence that better WM abilities enable more effective use of linguistic (e.g., grammatical) skills during the process of recognizing sentences presented through a CI.

Another neurocognitive skill, perceptual organization (or perceptual closure), refers to the process of using degraded sensory input to create a meaningful perceptual form (78–80). For speech, perceptual organization is the principle of treating the multiple sensory elements that compose the signal as a coherent percept (81). Relations have been found between visual perceptual organization skills and speech recognition in older adults with hearing loss (82). Perceptual closure was also assessed in hearing impaired listeners by George et al. (83), using a Text Reception Threshold test. Participants were asked to read degraded sentences on a computer screen with varying degrees of visual masking using vertical bars. Accuracy threshold scores predicted 10 to 15% of variability in speech detection thresholds in noise. Zekveld et al. (80) found that Text Reception Threshold scores predicted 30% of the variability in auditory speech reception threshold in speech-shaped noise for NH listeners. Similarly, using a Fragmented Sentence Test in which portions of printed letters had been deleted, Watson et al. (84) reported correlations with speech recognition in white noise for NH college students. We are currently investigating whether accuracy on a degraded visual Fragmented Sentence Test will predict sentence recognition in adult CI users.

An additional neurocognitive skill that is related to speech recognition outcome variability in adult CI users is inhibitory control, the ability to inhibit irrelevant stimulus information (e.g., noise), or to inhibit lexical competitors during recognition of words. Sommers and Danielson (85) provided support for age-related inhibitory deficits as a mechanism contributing to poorer word recognition in older adults. We recently examined inhibitory control abilities in adult CI users, using a computerized Stroop task (Moberly et al., under review). During this task, the participant viewed a computer monitor on which two types of visual stimuli were presented: color words that matched the color of the word (congruent condition) or color words that were presented in a different color (incongruent condition). Response times for the incongruent condition, indexing speed of inhibition, were found to be negatively correlated with recognition scores for words in sentences in speech-shaped noise. This finding suggests that inhibitory control processes play a role in speech recognition for adult CI users.

**IMPLICATIONS: WHAT NEEDS TO BE DONE?**

Although findings from the studies reviewed above begin to provide some insights into sources of variability in speech recognition outcomes for adults with hearing loss, these top-down processes have not been examined extensively in adult CI users. One of our goals is to develop and test a more comprehensive model of speech recognition in adult CI users to explain the interacting roles of auditory sensitivity, linguistic skills, and neurocognitive functions. As a field, a first step will be to re-conceptualize the way we look at cochlear implantation; we are not simply restoring auditory function, we are not simply restoring auditory function through a CI to an otherwise normal peripheral and central auditory system. A helpful way to approach this idea is using a "connectome" framework (86). A connectome is a network of neural projections and synaptic connections that shape an individual's global communication functions. As a child, development of a connectome is highly dependent on early sensory experience and activities. Likewise, postlingual adults with hearing loss should be considered as having a connectome disease: the sensory loss likely leads to downstream neurocognitive effects, which, in turn, have implications for adaptation to the CI. Application of a connectome model to patients with hearing loss suggests that outcomes following cochlear implantation will not be confined only to the auditory system.
itself, but also that the neurocognitive effects of prolonged hearing loss will affect outcomes in other related domains as well.

Second, there are clearly areas of investigation that have barely been touched relating to outcomes for adults with CIs, but have demonstrated relations in pediatric CI users. For example, greater maternal sensitivity to communication needs of the child predicts better speech recognition performance (87); interpersonal family dynamics have not been explored at all in adult CI users. Similarly, the role of intensive postoperative aural rehabilitation has been emphasized in the pediatric population but has primarily been limited in adult CI users to patient-driven computerized auditory training (88). Moreover, patients’ personality and “grit”—their perseverance and drive toward long-term goals (89)—likely contribute to ultimate outcomes, as demonstrated in a recent pilot study in our lab (7).

Third, new clinical assessment batteries need to be developed to examine these additional sources of variability in patients with CIs. We hypothesize that individual CI users may display poor speech recognition performance for different reasons, and these differences may suggest divergent therapeutic approaches. For patients with auditory sensitivity through their implants that is too poor to access essential speech structure, remapping of their devices or alternative signal processing strategies may be needed. For patients with poor linguistic skills, language training may be helpful. For CI users with poor neurocognitive processing, perceptual training which incorporates high cognitive demands may improve speech perception in degraded listening conditions directly, by training WM and attention, and indirectly strengthening cortical-subcortical sound-to-meaning relationships (90). Additionally, training methods may provide the opportunity to teach adults to effectively use compensatory mechanisms to cope with the processing demands of complex listening environments (91).

Lastly, a number of recommendations can be made for surgeons and audiologists. (1) It is important to recognize that poor performance for a given patient may not be entirely attributed to problems with bottom-up auditory sensitivity and audibility (i.e., electrode array placement, status of the auditory nerve); rather, it may be that top-down linguistic and neurocognitive factors are contributing, and formal aural rehabilitation in conjunction with an experienced speech-language pathologist may be helpful. (2) Anecdotal evidence from focused patient interviews suggests the potential benefit of patient-driven strategies like use of support groups or audio books to assist with rehabilitation (7). (3) It should be kept in mind that some patients may require more than 2 years of CI experience before reaching a plateau in performance (5,92). (4) Clinicians should talk with their colleagues regarding their poor performers. Although every clinician would prefer to discuss the CI users who are “stars,” it is likely that more open discussion regarding our poor performers will help us to develop more effective intervention strategies for individual patients (5). Understand that the broad array of factors contributing to CI outcomes suggests that a personalized rehabilitation approach likely needs to be developed for each patient. This thinking is in line with general advances across medical fields to develop and implement “personalized medicine” that is tailored to individual patients.

CONCLUSIONS

Poor speech recognition performance among adult CI users is, unfortunately, relatively common. Bottom-up auditory sensitivity, along with top-down linguistic and neurocognitive skills, contribute to variability in outcomes and likely differentially explain poor performance for many patients. By understanding these sources of variability and their interactions, and developing novel targeted intervention strategies that help patients remediate and compensate for them, we should be able to optimize speech and language outcomes for more patients and therefore drastically reduce the numbers of poor performers whom we see in our clinics.

REFERENCES


