

1. Introduction

Over the past 50 years, significant advancements have been made in the clinical management of severe to profound sensorineural hearing loss. Cochlear implants (CIs) are now widely considered to be an effective treatment option for both pediatric and adult patients. CIs are surgically implanted medical devices that use electrical stimulation of the auditory nerve to create the sensation of hearing. In the CI system, acoustic energy is detected by an external microphone and sent to a speech processor that is worn on the head. The speech processor is involved in extracting relevant acoustic features and determining electrode stimulation patterns. The signal is then transmitted via radio waves to an internal component, which transforms the input into electrical pulses to stimulate the auditory nerve (e.g., Loizou, 1998). Current electrode arrays are designed to provide stimulation for up to 22 intra-cochlear electrodes. The way in which electrodes are stimulated varies depending on the selected processing strategy. For example, in a continuous interleaved (CIS) processing strategy, acoustic information is bandpass-filtered based on the number of electrodes, and electrical pulses are generated using amplitudes proportional to the amount of energy contained in each bandpass channel. Electrodes are stimulated in a non-simultaneous fashion to minimize channel interaction (e.g., Loizou, 1998; Loizou, 2006). Conversely, an advanced combination encoder (ACE) processing strategy uses a subset of the available electrodes with stimulation occurring for frequency bands that contain the largest amplitudes (e.g., Hu & Loizou, 2008).

Listening through an implant is a markedly different experience than acoustic hearing. In the normal-hearing (NH) system, the cochlea contains thousands of hair cells that code fine-grained differences in frequency, which supports a high-resolution representation of auditory information in the environment. Conversely, most CI users only have up to 22 electrodes

available to represent the auditory signal, which results in a coarser signal representation. As a consequence, the implant does not afford access to the fine-grained spectral resolution that is available to NH listeners. Paradoxically, studies examining speech perception outcomes in CI users have demonstrated that listeners do not necessarily need access to the richness of spectral information present in the natural speech signal in order to achieve reliable speech comprehension – many CI users recognize speech transmitted through the implant, as do NH participants listening to spectrally degraded speech (e.g., Remez, Rubin, Pisoni, & Carrell, 1981; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). However, there is tremendous individual variability in speech perception outcomes among CI users. While some users show performance at or near the typical range, many others do not (e.g., Busby, Roberts, Tong, & Clark 1991; Dawson & Clark, 1997; Niparko et al., 2010). A variety of patient-related factors have been examined to determine their relevant contributions to speech perception outcomes. Research has found that many factors contribute to patient performance including age of implantation (e.g., Gantz et al., 1988; Parkin, Stewart, Dankowski, & Haas, 1990; Shea, Domio, & Orchik, 1990), pre- vs. post-lingual language status at time of implantation (e.g., Dawson et al., 1992), etiology of deafness (e.g., Battmer et al., 1995), and residual hearing following implantation (e.g., Gantz et al., 1988; van Dijk, van Olphen, Langereis, Mens, Brokx, & Smoorenburg, 1999; Friedland, Venick, & Niparko, 2003). In addition, other work has sought to link parameters of the device itself to outcome performance including number of active electrodes (e.g., Friesen, Shannon, Baskent, & Wang, 2001), coding strategy (e.g., Skinner, Holden, Whitford, Plant, Psarros, & Holden, 2002), and insertion depth (e.g., Finley & Skinner, 2009). Both patient- and device-related factors have been shown to partially account for the variance observed in patient outcomes following implantation. However, to date there is no comprehensive set of predictors

that can reliably account for differences in outcome performance observed with the device. That is, while a substantial portion our current knowledge base has justified the efficacy of the CI in order to demonstrate a benefit of the device at an implementation level (e.g., Pisoni, Kronenberger, Harris & Moberly, 2017), we have a more limited understanding as to why a CI works well for some patients, yet poorly for others. As a result, the current evidence base does not account for *why* patients with comparable language status, medical history, demographic background, and device may perform differently with the implant (e.g., Moberly, Bates, Harris, & Pisoni, 2016; Pisoni, Cleary, Geers, & Tobey, 1999).

One account as to why some CI users show poor performance outcomes is the lack of availability of acoustic cues. However, research examining NH participants has found that individuals can recognize speech in the absence of acoustic cues that they routinely have access to. For example, listeners readily demonstrate comprehension of compressed speech (Licklider & Pollack, 1948), sine-wave speech (e.g., Remez, Rubin, Pisoni, & Carrell, 1981), and noise-vocoded speech (e.g., Shannon et al., 1995), suggesting that availability of all acoustic cues is not necessary to achieve comprehension – instead listeners dynamically learn to adapt to the input to which they are exposed. This is an important consideration in the context of CI rehabilitation as it demonstrates that listeners do not necessarily need access to the full spectrum of acoustic cues in natural speech because adaptation mechanisms can help facilitate comprehension for acoustically degraded input. Thus, an alternate approach to understanding performance outcomes in CI users might not be which acoustic cues are available per se, but rather how the individual learns to adapt to the input transmitted through the CI.

One proposed technique to improve listening through a CI is auditory training. In general, auditory training can be characterized as passive (i.e., incidental) or active training. In passive

auditory training, listeners learn to adapt to the input transmitted through the CI in an unsupervised and unstructured manner. Passive training is a relatively common recommendation in the clinical domain, particularly for adult patients, as it allows the patient to integrate use of the CI into their everyday life. However, passive adaptation may be a time-consuming process, requiring frequent exposure to routine sounds in order to accurately pair an auditory stimulus to meaning. One study illustrated this point by examining the role of passive learning on adaptation for CI users undergoing an update to their processing strategy – a manipulation that requires even experienced users to adjust the mappings between auditory input and linguistic representations (Fu, Shannon, & Galvin, 2002). In this study, the frequency mapping was modified for a group of experienced CI users, who were allowed a three-month time period to passively adapt to the new parameters. They found that the CI users showed only minimum adaptation (i.e., no or modest improvements in speech recognition compared to baseline measures) in the time period of passive adjustment, though performance did improve with continued experience. These results suggest that unsupervised learning may not promote peak performance, particularly in the case where the CI map is novel or is not optimally configured for the user.

In contrast, listeners who engage in active auditory training engage in listening tasks that target a specific goal, such as phoneme, word, or sentence comprehension. As such, learning can be facilitated using rich contexts and feedback to meet the goal. Active auditory training may be beneficial for all CI users; it can help poor performers by providing structured opportunities to learn meaningful linguistic information, and can aid high performers by providing learning opportunities under difficult listening conditions (e.g., poor signal-to-noise ratio, novel talkers; Fu & Galvin, 2008). Moreover, it is common for CI manufacturers to update processing

strategies and technology regularly, requiring even experienced CI users to adapt to novel input. However, few CI users participate in active auditory training as part of rehabilitation, with even fewer programs targeted at post-lingual deafened adults (Fu & Galvin, 2007; Prendergast & Kelley, 2002). As a consequence, most CI users adapt to the implant in a passive fashion, though growing research suggests that this form of adaptation may not be sufficient to maximize listeners' performance (e.g., Fu & Galvin, 2008).

Many critical challenges exist for implementing auditory training as part of standardized aural rehabilitation. First, there is significant variability in speech perception outcomes following training; some individuals demonstrate a benefit, while others do not (Fu et al., 2005; Wu, Yang, Lin, & Fu, 2007; Busby et al., 1991). As a consequence, it is unclear how recommendations might differ across patients, and whether training may only be beneficial for patients demonstrating poor performance relative to patients who perform at a high level. In addition, because there are no concrete clinical recommendations on what the structure of active training should be (i.e., tasks, training duration, outcome measures), it is likely that few clinicians prioritize training as part of aural rehabilitation given the lack of evidence-based recommendations. Another challenge is that gains are often only measured immediately following training, with few studies examining whether performance is maintained over time, which is a critical factor to establish for any rehabilitation program. This contributes to systemic billing challenges for audiologists who do provide training. Finally, even in cases where auditory training appears to promote improvements in speech understanding, the outcome variables (e.g., task accuracy) may not explain *why* the improvement occurred because the measures do not index mechanistic processing changes. Standard clinical assessments of speech perception outcomes include word (e.g., Maryland CNC lists) and sentence (e.g., AzBio lists)

comprehension in quiet and noise. Using these metrics alone, it is difficult to determine why one patient shows a greater benefit from training relative to a patient who has a similar profile yet gains little benefit from training. The use of other physiologic measures may better assess changes at lower processing levels that precede changes in behavioral responses, though the relationship between physiologic changes and mechanistic changes is an area of continued research.

Here we review current forms of active auditory training including bottom-up focused training, in which attention is directed towards fine-grained acoustic differences in the signal, and top-down focused training paradigms, in which attention is directed towards the global structure of a sentence and/or contextual cues. For the nature of this review, our focus is on examining task-specific differences across training types with the population of interest being post-lingual adults, unless otherwise stated. We adopt an interdisciplinary perspective to bridge findings from the clinical rehabilitation literature with findings from the psycholinguistics domain, highlighting the role of top-down lexical feedback on adaptation to acoustically degraded speech in NH listeners. In this review, we argue that active auditory training may be necessary to optimally target the plasticity mechanisms underlying improvements in speech perception for listeners adapting to a CI.

2. Current forms of active auditory training in aural rehabilitation

Formal auditory training is aimed at enhancing auditory skills and improving speech understanding through a variety of listening exercises (Sweetow & Palmer, 2005; Sweetow & Sabes, 2007; Rayes, Malky, & Vickers, 2019). It has been argued that auditory training should be a standard part of aural rehabilitation as a way to structure how the listener learns to understand the transmitted signal. However, few CI users engage in formal training, with even fewer

programs targeted at post-lingual deafened adults, which may be due to a failure to prioritize training as part of the aural rehabilitation process for these users (Fu & Galvin, 2007; Prendergast & Kelly, 2002). Thus, current training protocols are often targeted at the pediatric population who demonstrate greater degrees of plasticity. While neural structure and function is maximally plastic during the first few years of life, plasticity mechanisms are present throughout the lifespan (e.g., Schramm, Fitzpatrick, & Seguin, 2002; Tong, Busby, & Clark, 1988; Sharma, Gilley, Dorman, & Baldwin, 2007; Most, Shrem, & Duvdevani, 2010). Adults continue to learn new skills, but often require more exposure or practice to do so. This principle underlies flexibility in the CI candidacy criterion, which now allows pre-lingual or post-lingual deafened adults the opportunity to receive an implant. Theories of residual plasticity would predict that implanted adults could reach the performance of younger implanted individuals with sufficient and prolonged experience (Kral & Sharma, 2012). However, research has shown that late-implanted individuals, in this case pre-lingual deafened participants over the age of 12, who used their CI for an extended period of time (i.e., at least six months post-implantation) showed tremendous variability in open-set sentence recognition (Schramm, Fitzpatrick, & Segun, 2002). Of the 15 participants in their sample, 10 participants showed an improvement from pre- to post-implantation performance on open set word recognition. However, nine participants performed at or below 40% performance, a level considered to represent minimal benefit. Thus, not all of the patients showed an improvement post-implantation, and of those patients who did demonstrate improvement, most of the sample did not perform at a level consistent with an acceptable benefit. Thus, the results of Schramm et al. (2002) indicate that extended passive listening through a CI may not maximize speech understanding. Studies following newly implanted individuals suggest that most gains occur within the first 3–6 months following implantation

(Kessler, Loeb, & Barker, 1995; Spivak & Waltzman, 1990; Waltzman, Cohen, & Shapiro, 1986; Fu & Galvin, 2008), with less improvement observed after this time point (e.g., Dorman, Loizou, & Rainey, 1997; Pelizzone & Cosendai, & Tinembart, 1999; Helms et al., 2004). This has prompted discussion in the audiological field that CI users may need other rehabilitative support to fully acclimate to their device (e.g., Fu & Galvin, 2008).

There are two primary approaches to active auditory training for rehabilitation: bottom-up focused training and top-down focused training. Both approaches have the same end goal of improving speech understanding for CI users, though they use different strategies to meet that goal. Bottom-up approaches, also called analytic training, focus on the building blocks of speech perception, including sensitivity to fine-grained acoustic details in the speech signal. Top-down approaches, also called synthetic training, focus on training the listener to use lexical and contextual cues to fill in perceptual gaps while processing speech (Fu & Galvin, 2007).

2.1. Bottom-up focused auditory training

The bottom-up approach encourages listeners to focus on the acoustic signal itself. True bottom-up training is designed to promote improvements in processing efficiency of the sensory signal (Fu & Galvin, 2007). This is particularly prominent for CI users as they experience a spectral mismatch in perception because the auditory input is filtered through a limited number of electrodes. While the amplitude envelope within each channel is relatively preserved, the spectral structure of the channel is poorly conveyed (e.g., Shannon et al., 1995; Nelson, Jin, Carney and Nelson, 2003; Moberly et al., 2014). Previous research has demonstrated that NH listeners rely on a variety of spectral and temporal cues for phonemic decisions including formants (e.g., Hillenbrand, Getty, Clark, Wheeler, 1995; Neel, 2008), formant transitions (e.g., Delattre, Liberman, Cooper, 1955; Stevens & Klatt, 1974; Stevens & Blumstein, 1978; Walley &

Carrell, 1983), spectral center (e.g., van Son & Pols, 1997), and periodicity (e.g., Cole & Cooper, 1975). A critical question is the degree to which CI users rely on the same acoustic cues that NH listeners use to achieve comprehension. Research has found evidence that some CI users demonstrate similar acoustic cue weighting strategies to that of NH listeners. For example, Iverson, Smith, and Evans (2006) found that vowel recognition in CI users was adversely affected when the speech signal was processed to remove formant movement or equate vowel duration, suggesting that these cues, as with NH listeners, are strongly weighted in perception for CI users. However, there is also evidence suggesting that CI users may weight acoustic cues differently than NH listeners. For example, Hedrick and Carney (1997) studied four post-lingual deafened CI users to determine the relative contributions of amplitude and formant transition information on consonant-vowel identification. They found that compared to a NH control group, the CI users consistently relied more on the relative amplitude cues while the NH listeners used both formant structure and amplitude information, suggesting a difference in perceptual weighting strategies between the groups. Similar findings were demonstrated by Moberly et al. (2014), who examined labeling patterns of the /ba/-/wa/ contrast in a sample of 20 post-lingual CI users. While NH listeners weighted the spectral structure over amplitude structure exclusively, CI users showed variability in which cues were prioritized. Some users relied on amplitude structure to a greater degree, in line with a theory of shifted perceptual strategies in which CI users weight the acoustic cues that are most saliently delivered. However, it was found that the CI users who adopted a strategy closer to NH listeners, in which spectral structure was weighted over amplitude cues, demonstrated the best word recognition performance overall.

It is thought that bottom-up training approaches promote changes in how acoustic information is weighted by providing structured training in which listeners engage in tasks that

direct attention towards fine-grained, meaningful acoustic cues. A limited number of studies have examined bottom-up auditory training in CI users and results have been equivocal. For example, Busby, Roberts, Tong, and Clark (1991) did not observe significant improvements in vowel perception for a small sample of pre-lingual deafened CI patients who completed 10 one-hour training sessions. Conversely, Dawson and Clark (1997) examined five CI patients following 10 weeks of bottom-up focused training sessions using explicit vowel training tasks that focused on rhyme generation, discrimination, and identification. They found that four out of the five participants showed some improvement in at least one of the trained areas, and one patient showed generalization to novel contexts. Similar results were reported by Fu, Galvin, Wang, and Nogaki (2005), who examined 10 CI patients who completed at-home training that was customized based on their baseline performance. Participants completed a three-alternative forced-choice discrimination task. On each trial of this task, participants heard two identical sounds and one outlier sound that differed maximally with respect to acoustic features. Over the course of the experiment, the outlier sound became increasingly similar to the identical sounds, making the experiment more difficult as it progressed. Participants also completed three-alternative forced-choice vowel, consonant, and nonsense syllable identification tasks that also increased in difficulty with improved performance. They found improvements in vowel and consonant recognition, with some participants also showing generalization to sentence recognition tasks.

Overall, it is difficult to conclude whether strict bottom-up focused training promotes significant improvements in speech comprehension as these studies suffer from limited sample sizes, document wide variability in benefits across patients, and show limited generalization. However, the remaining low-level acoustic cues available to CI users may be more amenable to

auditory training and may play a greater role for attending to speech in acoustically simple environments, like listening in quiet (e.g., Fu & Galvin, 2007). A key limitation with the bottom-up approach is that the tasks used during training may be considered less functionally relevant than top-down approaches. Specifically, training focused on recognition of single phonemes is not typically a task required to achieve comprehension under everyday listening conditions. Instead, most listeners require comprehension of a continuous stream of sentences in everyday environments and have access to broad contextual cues. Spoken language comprehension occurs through interactive processes at both the low-level acoustic level and higher-order linguistic level. Future research is needed to explicate the degree to which bottom-up focused training generalizes to more functionally relevant communication tasks, such as single words or sentences.

2.2. Top-down focused auditory training

Top-down based approaches to auditory training emphasize the use of lexical and contextual cues to decode the speech signal. This type of training targets skills that support auditory processing and language skills. While bottom-up approaches first direct attention towards the fine-grained structure of the input, top-down approaches first direct attention towards the global structure of the signal, which is thought to closely mirror how NH listeners learn language (Nittrouer & Caldwell-Tarr, 2016). Unlike analytic approaches, the focus of the top-down approach is on using contextual cues to facilitate comprehension. Materials and training tasks for top-down based auditory training include transcribing words or sentences under clear or challenging listening conditions, such as in background noise or with multiple talkers.

One common form of top-down auditory training is connected discourse tracking (DeFilippo & Scott, 1978). This type of training requires a talker (sender) and listener (receiver)

who create a set of speech materials for the talker to read. The talker reads the prepared text and the receiver must repeat back exactly what was said, with access to visual and/or auditory information. If the listener makes an error, then the talker may paraphrase the materials or prompt the listener with additional cues (Levitt, Waltzman, Shapiro, Cohen, 1986). Clinicians assess the number of words correct per minute and can compare performance in auditory only, visual only, or audiovisual conditions to determine individual contributions of each domain on comprehension. Research utilizing connected discourse tracking has found variable results on its efficacy. Levitt et al. (1986) examined performance of five CI patients using Connected Discourse Tracking over a 10-week training period. They report that all five subjects showed significant improvements in tracking rate, though the magnitude of learning differed considerably, and none of the participants reached the threshold considered within normal listening levels. The use of connected discourse tracking has been a topic of debate in the audiology domain because the paradigm has many uncontrolled variables (e.g., talker differences, listener differences, text materials, repeated presentations) and the guidelines to compare performance within subjects are not well established (Tye-Murray & Tyler, 1988). However, use of top-down focused training paradigms may hold promise for aural rehabilitation, as they are commonly used in the psycholinguistics literature for adaptation to degraded speech input for NH listeners. In the following section, we review the role of lexical information on speech perception for NH listeners, and highlight findings demonstrating that contextual cues play an influential role for adaptation to atypical speech input.

3. Lexical influences on speech perception

3.1. Lexical effects in NH users

Everyday listening environments present a challenge for even the most skilled NH individuals because listening often occurs in suboptimal environments and talkers show a high degree of variability in the acoustic realization of speech sounds (e.g., Delattre, Liberman, Cooper, 1955; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Peterson & Barney, 1952; Jusczyk, 1997; Theodore, Miller, & DeSteno, 2009; Newman et al., 2001). These findings have been formalized as the “lack of invariance” problem for speech perception, which reflects the fact that there is no one-to-one mapping between the acoustic signal and a given speech sound. However, structured variability is not inherently negative as it can give listeners information about the talker, for example, which listeners can use to dynamically modify the structure of phonetic category representations (e.g., Allen, Miller, DeSteno, 2003; Theodore, Miller, & DeSteno, 2009; Clayards, Tannenhaus, Aslin, & Jacobs, 2008; Miller, 1994; Drouin, Theodore, & Myers, 2016).

Seminal research has shown that listeners use lexical information to guide speech perception, which is particularly useful when the speech signal is variable, ambiguous, or degraded. Phonemic restoration is one demonstration of a lexical effect on perception, in which listeners fill in missing or degraded input using contextual knowledge. Warren (1970) first demonstrated that when a cough or tone replaced the initial /s/ in a word like *legislatures*, listeners could not reliably identify when the gap occurred and many participants reported that the word was fully intact, suggesting an apparent restoration of the missing phoneme, which has also been observed in listeners with hearing loss (Başkent, Eiler, & Edwards, 2010). The lexical effect has also been observed with perception of ambiguous speech input. For example, Ganong (1980) exposed listeners to stimuli varying along a voice-onset-time continuum where one end of the continuum yielded perception of a lexical item (e.g., *kiss* or *gift*), while the other end did not

(e.g., *giss* or *kift*). He found that listeners were biased to perceive more of the speech continuum as a /g/ when in the context of /ift/, where only the /g/ interpretation yields a real English word. However, when the same voice-onset-time variants were placed in the context of /is/, listeners perceived more of the continuum as /k/, as only the /k/ interpretation yields a lexical interpretation.

Collectively, these findings demonstrate how speech perception for NH listeners reflects interactions at both the phonemic and lexical levels – all of which occurs with relative ease and within milliseconds as the speech signal unfolds in time. Indeed, many studies on the time course of lexical access have shown that access to even a single phoneme creates lexical competition, supporting activation of lexical candidates consistent with the phonemic input (e.g., Allopenna, Magnuson, & Tannehaus, 1998). For example, Allopenna et al. (1998) used a visual world eye-tracking paradigm in which listeners heard a target word (e.g., *beaker*) while simultaneously viewing an array of pictures containing a cohort competitor (e.g., *beetle*), a rhyme competitor (e.g., *speaker*), and an unrelated word (e.g., *dolphin*). As the first few hundred milliseconds of the target word unfolded over time, the cohort competitor showed the greatest initial competition. As the end of the target word was processed, the rhyme competitor emerged as the dominant competitor. Such patterns have been modeled using interactive accounts of spoken word recognition, such as TRACE, in which lower-level perceptual processing is altered based on top-down lexical feedback that unfolds over time (e.g., McClelland & Elman, 1986; Allopenna et al., 1998).

3.2. Lexical effects in CI users

An important consideration is how lexical information may be weighted differently depending on the specific listening context. When the quality of the signal is good (e.g., clear

speech, minimal background noise), bottom-up information may contribute more to processing. However, when the stimulus quality is poor (e.g., degraded input, background noise), top-down feedback may play a more dominant role in speech processing (e.g., Norris et al., 2003; Fu & Galvin, 2007). Under this view, top-down feedback might contribute more to speech processing for CI users because the transmitted signal contains limited spectral information. Strikingly, previous research has shown atypical patterns of lexical influences on speech perception for CI users. For example, recent work on lexical access in CI users using a visual world paradigm has demonstrated significant delays in lexical activation which may be more pronounced for pre-lingual CI users (Farris-Trimble, McMurray, Cigrand, & Tomblin, 2014; McMurray, Farris-Trimble, & Rigler, 2017). Delays in lexical access at the word level may have detrimental, cascading effects at the sentence level because the opportunity to build sentence structure is impaired as the delay lengthens. Indeed, one study found that pediatric CI users do not use sentence context to facilitate word recognition; instead, they appear to process sentences as a string of unrelated words (Conway, Deocampo, Walk, Anaya, & Pisoni, 2014). Future research is needed to examine how auditory training for CI users may be used to foster a perceptual system that uses lexical cues to improve processing efficiency. In the next section, we review the role of both lexically-oriented and non-linguistic training paradigms on improving perception of degraded speech signals for naïve NH listeners, which lay the foundation to promote optimal translation to the clinical population.

4. Factors that promote perceptual learning for degraded speech signals

A common occurrence in everyday listening conditions is the experience of poor speech understanding when initially parsing an acoustically degraded speech signal, which improves with sufficient exposure or with sufficient context cues. This occurs regularly in the initial

activation period for CI patients, who often may not perceive speech during activation, yet show perceptual improvements with extended use. Audiologists counsel users on the importance of wearing the device regularly to improve how the brain learns the new sensory signals transmitted through the CI. This also occurs for NH listeners who may initially encounter a novel talker with an unfamiliar accent or dialect. Despite initial difficulty, listeners report a rapid improvement in their speech understanding with exposure. This phenomenon is referred to in the psychology literature as perceptual learning. Perceptual learning can be defined as adaptive changes in an organism's perceptual system that enhance the system for future interactions (Goldstone, 1998). In the speech domain, perceptual learning can result in long-lasting changes to the mapping process between the speech signal and phonetic categories (Goldstone, 1998; Norris, Cutler, & McQueen, 2003). A body of research has studied perceptual learning in lab-based settings in NH listeners to examine the factors that underlie improvements in speech understanding. Across many studies and different forms of acoustic degradations, research has demonstrated that NH listeners consistently rely on lexical information to promote improvements in comprehension of challenging speech signals, and that improvements can be leveraged through a brief training period.

4.1. A role for lexical information

One example of the influence of lexical information on speech learning occurs in the lexically guided perceptual learning paradigm. In this paradigm, listeners complete a brief training exposure phase followed by a test phase. During the training phase, participants listen to words and nonwords and complete a lexical decision task (Norris et al., 2003; Kraljic & Samuel, 2005; Drouin et al., 2016; Drouin et al., 2018). Critically, listeners are exposed to an atypical production during the training phase, such as a fricative that is spectrally ambiguous between /s/

and /ʃ/. For one group of listeners (e.g., /s/-bias group), the ambiguous replaces the medial /s/ in words that normally contain an /s/ (e.g., *dinosaur*). For a different group of listeners (e.g., /ʃ/-bias group), the same ambiguous sound replaced the medial /ʃ/ in words that normally containing an /ʃ/ (e.g., *publisher*). Accordingly, lexical context can be used to resolve the ambiguity in the atypical productions. In the test phase, listeners categorize items from a nonword continuum (e.g., /ɑʃi/-/asi/) to assess changes in the mapping between speech acoustics and the phonetic categories. The standard result in this paradigm is that performance differs between the two exposure groups at test in line with their exposure during training; specifically, listeners in the /s/-bias training group categorize more test items as /s/ compared to listeners in the /ʃ/-bias group. This finding demonstrates a persistent influence of lexical context on speech perception, even when disambiguating lexical context is removed. It has been confirmed that this effect is explicitly driven by lexical context as learning is absent when the ambiguous sound is placed in a nonword context during exposure (e.g., Norris et al., 2003).

Lexically guided perceptual learning is also observed for listeners learning to adapt to noise-vocoded speech. Noise-vocoding is a digital signal manipulation that is frequently used in the research domain to approximate the auditory experience of an individual using a CI. Noise-vocoding consists of dividing the natural speech signal into frequency bands, extracting the amplitude envelope, and replacing the fine structure in each band with noise (e.g., Faulkner, Rosen, & Smith, 2000; Loizou, Dorman, & Tu, 1999; Shannon et al., 1995; Davis, et al., 2005). This manipulation has been useful in research studies to determine the spectral resolution that is necessary for proficient speech comprehension. Standard noise-vocoded learning paradigms consist of a pre-test, or baseline measure of speech understanding, followed by a brief period of linguistic training with feedback. A post-test transcription measure is then compared to baseline

to gauge improvement as a function of training. Numerous training studies have found that while perception of noise-vocoded speech is poor at initial exposure, listeners show robust perceptual learning with only minutes of training exposure, which generalizes across sentences (e.g., Loebach, Bent, & Pisoni, 2008), talkers (Huyck, Smith, Hawkins, & Johnsrude, 2017), and stimulus types (e.g., Hervais-Adelman, Davis, Johnsrude, & Carlyon, 2008; Loebach, Pisoni, & Svirsky, 2009).

Previous research suggests that perceptual learning of noise-vocoded speech, like other atypical input, is strongly influenced by stimulus lexicality and feedback. In a series of experiments, Davis et al. (2005) trained NH listeners on noise-vocoded sentences composed of either words or nonwords. During training, listeners were provided with interleaved feedback in which they heard the noise-vocoded sentence, then heard the clear version, and finally heard the same noise-vocoded sentence repeated again. At test, participants were asked to report as many words as possible and no feedback was provided. They found that listeners trained with vocoded sentences composed of words performed significantly better than those trained with nonword sentences. In fact, the participants trained with nonword sentences were indistinguishable from naïve listeners with no training, highlighting the role of lexical context in facilitating comprehension of degraded input.

In addition to stimulus lexicality, research has also examined the role of lexical feedback in perceptual learning of noise-vocoded speech. Listeners do not necessarily require explicit feedback to learn noise-vocoded speech (Davis et al., 2005); however, feedback may allow learning to occur more efficiently. Studies have examined the role of both the type of feedback and the time course of feedback on learning noise-vocoded speech. Top-down approaches to adaptation posit that learning occurs via a comparison process between a given input and its

target representation (e.g., Norris et al., 2003; Mirman, McClelland, & Holt, 2006). Providing explicit feedback as to what the target item is can allow for ambiguous representations to be adjusted to reflect the intended item (Hervais-Adelman et al., 2008). Research has shown that both written and auditory feedback are effective during training for impoverished input (Davis et al., 2005; Schwab, Nusbaum, Pisoni, 1985; Greenspan, Nusbaum, Pisoni, 1988) and that perceptual learning occurs more rapidly for listeners who receive feedback prior to hearing the stimulus, suggesting that knowing the identity of the target can allow learning to occur more rapidly (Davis et al., 2005; Hervais-Adelman et al., 2005).

4.2. Alternate routes to perceptual learning

The architecture of the TRACE model of spoken word recognition (McClelland & Elman, 1986) is consistent with a facilitative role for top-down auditory training on speech perception. In this framework, engaging the lexicon during training allows for an improved tuning in the mapping of the sensory signal to prelexical representations through a top-down feedback mechanism. Top-down approaches to adaptation posit that learning occurs via a comparison process between a given input and a target representation (e.g., Norris et al., 2003). Providing explicit lexical feedback as to what the target item is can allow for potentially ambiguous representations to be adjusted to reflect the intended item – leading to a refinement in the mapping process given experience (Hervais-Adelman et al., 2005). Support for this framework comes from the finding that perceptual learning of noise-vocoded speech is absent for listeners who are trained with nonword sentences (Davis et al., 2005), in line with other findings from the psycholinguistic domain (e.g., Norris et al., 2003). However, this finding was not replicated by listeners trained on single nonwords, suggesting that when the memory load is sufficiently reduced, listeners can utilize non-linguistic information for speech adaptation.

Nevertheless, lexical context may maximize perceptual learning for vocoded speech as the magnitude of learning was smaller for single word training (e.g., Hervais-Adelman et al., 2008), compared to training with sentences (e.g., Davis, Johnsrude, Hervais-Adelman, Taylor, & McGettigan, 2005). Research has shown that both written and auditory lexical feedback are effective for training of impoverished input (Davis et al., 2005; Schwab, Nusbaum, Pisoni, 1985; Greenspan, Nusbaum, Pisoni, 1988) and that perceptual learning occurs more rapidly for listeners who receive feedback prior to hearing the stimulus, which suggests that knowing the identity of the target can allow learning to occur more rapidly (Davis et al., 2005; Hervais-Adelman et al., 2005). This finding is in line with the TRACE supervised learning framework, which allows backpropagation between processing layers to integrate available contextual information into the percept (McClelland & Elman, 1986; Davis et al., 2005; Hervais-Adelman et al., 2005). Under this framework, lexically rich contexts during training promotes an optimal environment in which to learn degraded input.

An alternative account posits that learning may be mediated by the similarity of the tasks used during training and test. In a transfer-appropriate processing (TAP) framework (e.g., Franks et al., 2000), performance is maximized when the task used to assess learning at test is identical to the task used during training. Under this theoretical framework, learning occurs because of a match between training and test tasks. Support for this view comes from studies demonstrating perceptual learning on tasks that mirror that the test task (i.e., Davis et al., 2005). Thus, advocates for this view argue that linguistic focused training tasks are necessary to observe a linguistic benefit. This view has been challenged by findings that demonstrate a linguistic benefit using training tasks that direct attention towards non-linguistic characteristics of the signal. For example, recent work has demonstrated perceptual learning of noise-vocoded speech using non-

linguistic training tasks that differ from the test task (Loebach, Bent, & Pisoni, 2008). In this study, three groups of NH listeners heard noise-vocoded sentences produced by multiple talkers during a training period. During training, one group was asked to transcribe the sentence, one group was asked to identify the talker, and another group was asked to identify the gender of the talker. All listeners received feedback in line with the assigned task. At test, all participants, regardless of training group, were asked to complete a transcription task to measure language comprehension. The results showed that perceptual learning was equivalent for listeners trained with either the linguistic transcription task or the non-linguistic talker identification task. This finding suggests that a linguistic benefit can be obtained regardless of whether the training phase mirrors the task used during test, and is in line with other forms of perceptual learning for adaptation to ambiguous speech sounds (e.g., Drouin et al., 2018).

The findings from Loebach et al. (2008) are considered with respect to a depth of processing framework (e.g., Craik & Lockhart, 1972). In this framework, perceptual learning occurs if the task used during training is sufficiently challenging so that the listener needs to attend to fine-grained acoustic information in order to accurately perform the task. Under this view, perceptual learning is directly mediated by the attentional requirements of the task and thus can be achieved using a variety of top-down or bottom-up focused training strategies, so long as the listener is sufficiently challenged to engage with the stimuli. Support from this view comes from the Loebach et al. (2008) finding showing that perceptual learning, as measured through a linguistic transcription task, was equally maximized for two groups of listeners who completed a challenging task during training (i.e., lexical transcription and talker identification), but was less robust for listeners who completed an easier gender identification task during training. They argue that this finding reflects the attentional requirements of the training task. Specifically, the

talker identification and transcription tasks were more cognitively challenging tasks relative to the gender identification training group, which resulted in a disparity in how listeners encoded and learned the novel input. Together, these findings suggest that similarity between training and test asks in addition to the attentional requirements of the task are both important factors to consider when designing adaptation studies where the outcome goal is comprehension.

5. Translating principles of learning to auditory training for CI users

Predicting variability in speech perception outcomes following implantation remains a challenge in the CI literature. Here we suggest that active auditory training may offer a means to maximize adaptation for both low and high performers. In this review, we have focused on findings from the psycholinguistics literature demonstrating that robust perceptual learning can occur even in the face of acoustic degradation. Human speech perception is a highly plastic and adaptive process that allows for the compensation of a variety of listening conditions with relative ease. The conditions that promote ease of understanding for atypical speech input has been the primary focus of this review. We have examined the influential effects of lexical knowledge on processing impoverished input for NH listeners, with a focus on how active top-down feedback can propagate to improve the mapping between a sensory signal and prelexical representations. This is an important consideration with respect to the rehabilitation literature because it suggests that passive listening through the implant, without engagement in active training, may not sufficiently maximize opportunities to improve perception. Translational research characterizing how CI users benefit from lexical context and how lexically driven training approaches promote improvements in perception is needed to begin to formalize concrete rehabilitation recommendations. While many outstanding questions still exist, the literature reviewed here for NH listeners lays the foundation for establishing a benefit of lexical

context and feedback for degraded speech signals – raising the possibility that modeling training paradigms in this way for CI users may also promote similar benefits. If lexical context is the putative factor for driving perceptual learning of degraded speech input transmitted through a CI, then listeners trained using lexically oriented paradigms may show the greatest long-term benefit from their device.

We have also reviewed alternate routes to perceptual learning under different theoretical frameworks and consider how task-based listening strategies may mediate learning outcomes. This is particularly important when considering how to design optimal training paradigms. If training using cognitively easier tasks (i.e., talker identification instead of lexical transcription) promotes equivalent learning outcomes on linguistic tasks, then it opens the opportunity to better customize training to patient needs. An understudied area in the current research domain with respect to bottom-up or top-down focused listening strategies is the degree to which gains are maintained over long-term time periods. Only through a longitudinal design may we better document potential differences among training groups. Namely, would two groups of listeners trained using different tasks show equivalent performance in the short-term, but differ with respect to long-term benefit of gains? Currently, we do not have the evidence base to definitively answer this question, which contributes to hesitations about recommending a specific training protocol.

Finally, there is a significant need to better understand how individual patient factors interact with training outcomes and to characterize differences in plasticity across patient profiles. Even within the NH population there remains significant variation in how individuals adapt to acoustically poor input. While most training studies have focused exclusively on outcomes at the group level, characterizing individual differences in learning is emerging as an

active area of research. As reviewed, both bottom-up and top-down processes drive perceptual adaptation to acoustically poor input, therefore differences may emerge from variation in how listeners code the low-level sensory signal or how listeners integrate higher-order contextual cues in the mapping process. Could customized training programs be used to specifically target weaknesses in either domain to close the gap between high and low performers? We propose that the next phase of auditory training research should be aimed at examining the use of novel outcome measures, long-term assessments, training variables (e.g., task, duration, time of intervention), and individual difference metrics. By establishing these parameters, scientists and clinicians will be able to facilitate recommendations for the conditions that promote general speech learning as well as how training paradigms might be customized to the needs of individual CI users.

6. References

- Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America*, *113*(1), 544-552.
- Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*(4), 419-439.
- Başkent, D., Eiler, C. L., & Edwards, B. (2010). Phonemic restoration by hearing-impaired listeners with mild to moderate sensorineural hearing loss. *Hearing Research*, *260*(1), 54-62.
- Battmer, R. D., Gupta, S. P., Allum-Mecklenburg, D. J., & Lenarz, T. H. (1995). Factors influencing cochlear implant perceptual performance in 132 adults. *The Annals of Otolaryngology, Rhinology & Laryngology. Supplement*, *166*, 185-187.

- Busby, P. A., Roberts, S. A., Tong, Y. C., & Clark, G. M. (1991). Results of speech perception and speech production training for three prelingually deaf patients using a multiple-electrode cochlear implant. *British Journal of Audiology*, 25(5), 291-302.
- 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.
- Cole, R. A., & Cooper, W. E. (1975). Perception of voicing in English affricates and fricatives. *The Journal of the Acoustical Society of America*, 58(6), 1280-1287.
- Conway, C. M., Deocampo, J. A., Walk, A. M., Anaya, E. M., & Pisoni, D. B. (2014). Deaf children with cochlear implants do not appear to use sentence context to help recognize spoken words. *Journal of Speech, Language, and Hearing Research*, 57(6), 2174-2190.
- Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671-684.
- Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General*, 134(2), 222-241.
- Dawson, P. W., Blamey, P. J., Rowland, L. C., Dettman, S. J., Clark, G. M., Busby, P. A., ... & Rickards, F. W. (1992). Cochlear implants in children, adolescents, and prelinguistically deafened adults: Speech perception. *Journal of Speech, Language, and Hearing Research*, 35(2), 401-417.
- Dawson, P. W., & Clark, G. M. (1997). Changes in synthetic and natural vowel perception after specific training for congenitally deafened patients using a multichannel cochlear implant. *Ear and Hearing*, 18(6), 488-501.

- DeFilippo, C.L. & Scott, B.L. (1978). A method for training and evaluation of the reception of on-going speech. *Journal of the Acoustical Society of America*, 63, 1186-1192.
- Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *The Journal of the Acoustical Society of America*, 27(4), 769-773.
- Dorman, M. F., Loizou, P. C., & Rainey, D. (1997). Simulating the effect of cochlear-implant electrode insertion depth on speech understanding. *The Journal of the Acoustical Society of America*, 102(5), 2993-2996.
- Drouin, J. R., & Theodore, R. M. (2018). Lexically guided perceptual learning is robust to task-based changes in listening strategy. *The Journal of the Acoustical Society of America*, 144(2), 1089-1099.
- Drouin, J. R., Theodore, R. M., & Myers, E. B. (2016). Lexically guided perceptual tuning of internal phonetic category structure. *The Journal of the Acoustical Society of America*, 140(4), EL307-EL313.
- Eisner, F., McGettigan, C., Faulkner, A., Rosen, S., & Scott, S. K. (2010). Inferior frontal gyrus activation predicts individual differences in perceptual learning of cochlear-implant simulations. *Journal of Neuroscience*, 30(21), 7179-7186.
- Farris-Trimble, A., McMurray, B., Cigrand, N., & Tomblin, J. B. (2014). The process of spoken word recognition in the face of signal degradation. *Journal of Experimental Psychology: Human Perception and Performance*, 40(1), 308-327.
- Faulkner, A., Rosen, S., & Smith, C. (2000). Effects of the salience of pitch and periodicity information on the intelligibility of four-channel vocoded speech: Implications for cochlear implants. *The Journal of the Acoustical Society of America*, 108(4), 1877-1887.
- Finley, C. C., & Skinner, M. W. (2008). Role of electrode placement as a contributor to

- variability in cochlear implant outcomes. *Otology & Neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 29(7), 920-928.
- Franks, J. J., Bilbrey, C. W., Lien, K. G., & McNamara, T. P. (2000). Transfer-appropriate processing (TAP). *Memory & Cognition*, 28(7), 1140-1151.
- Friedland, D. R., Venick, H. S., & Niparko, J. K. (2003). Choice of ear for cochlear implantation: The effect of history and residual hearing on predicted postoperative performance. *Otology & Neurotology*, 24(4), 582-589.
- Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America*, 110(2), 1150-1163.
- Fu, Q. J., & Galvin III, J. J. (2007). Perceptual learning and auditory training in cochlear implant recipients. *Trends in Amplification*, 11(3), 193-205.
- Fu, Q. J., & Galvin, J. J. (2008). Maximizing cochlear implant patients' performance with advanced speech training procedures. *Hearing Research*, 242(1), 198-208.
- Fu, Q. J., Nogaki, G., & Galvin, J. J. (2005). Auditory training with spectrally shifted speech: implications for cochlear implant patient auditory rehabilitation. *Journal of the Association for Research in Otolaryngology*, 6(2), 180-189.
- Fu, Q. J., Shannon, R. V., & Galvin III, J. J. (2002). Perceptual learning following changes in the frequency-to-electrode assignment with the Nucleus-22 cochlear implant. *The Journal of the Acoustical Society of America*, 112(4), 1664-1674.
- Fu, Q. J., Galvin, J., Wang, X., & Nogaki, G. (2005). Moderate auditory training can improve

- speech performance of adult cochlear implant patients. *Acoustics Research Letters Online*, 6(3), 106-111.
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, 6(1), 110-125.
- Gantz, B.J., Tyler, R.S., Abbas, P., Tye-Murray, N., Knutson, J.F., McCare, B.F.... & Kuk, F. (1988). Evaluation of five different cochlear implant designs: Audiologic assessment and predictors of performance. *Laryngoscope*, 98, 1100-1106.
- Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, 49(1), 585-612.
- Greenspan, S. L., Nusbaum, H. C., & Pisoni, D. B. (1988). Perceptual learning of synthetic speech produced by rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 421-433.
- Hedrick, M. S., & Carney, A. E. (1997). Effect of relative amplitude and formant transitions on perception of place of articulation by adult listeners with cochlear implants. *Journal of Speech, Language, and Hearing Research*, 40(6), 1445-1457.
- Helms, J., Weichbold, V., Baumann, U., von Specht, H., Schön, F., Müller, J., ... & D'Haese, P. (2004). Analysis of ceiling effects occurring with speech recognition tests in adult cochlear-implanted patients. *ORL*, 66(3), 130-135.
- Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S., & Carlyon, R. P. (2008). Perceptual learning of noise vocoded words: Effects of feedback and lexicality. *Journal of Experimental Psychology: Human Perception and Performance*, 34(2), 460-474.
- Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, 97(5), 3099-3111.

- Hu, Y., & Loizou, P. C. (2008). A new sound coding strategy for suppressing noise in cochlear implants. *The Journal of the Acoustical Society of America*, *124*(1), 498-509.
- Huyck, J. J., Smith, R. H., Hawkins, S., & Johnsrude, I. S. (2017). Generalization of perceptual learning of degraded speech across talkers. *Journal of Speech, Language, and Hearing Research*, *60*(11), 3334-3341.
- Iverson, P., Smith, C. A., & Evans, B. G. (2006). Vowel recognition via cochlear implants and noise vocoders: Effects of formant movement and duration. *The Journal of the Acoustical Society of America*, *120*(6), 3998-4006.
- Jusczyk, P. W. (1997). *The Discovery of Spoken Language*. Cambridge: MIT Press.
- Kral, A., & Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends in Neurosciences*, *35*(2), 111-122.
- Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal?. *Cognitive Psychology*, *51*(2), 141-178.
- Kessler, D. K., Loeb, G. E., & Barker, M. J. (1995). Distribution of speech recognition results with the Clarion cochlear prosthesis. *Annals of Otology, Rhinology, & Laryngology*, *104*(Suppl 166), 283-285.
- Levitt, H., Waltzman, S. B., Shapiro, W. H., & Cohen, N. L. (1986). Evaluation of a cochlear prosthesis using connected discourse tracking. *Journal of Rehabilitation Research and Development*, *23*(1), 147-154.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431-461.
- Licklider, J. C. R., & Pollack, I. (1948). Effects of differentiation, integration, and infinite peak

- clipping upon the intelligibility of speech. *The Journal of the Acoustical Society of America*, 20(1), 42-51.
- Loebach, J. L., Bent, T., & Pisoni, D. B. (2008). Multiple routes to the perceptual learning of speech. *The Journal of the Acoustical Society of America*, 124(1), 552-561.
- Loebach, J. L., Pisoni, D. B., & Svirsky, M. A. (2009). Transfer of auditory perceptual learning with spectrally reduced speech to speech and nonspeech tasks: Implications for cochlear implants. *Ear and Hearing*, 30(6), 662-674.
- Loizou, P. C. (1998). Mimicking the human ear. *IEEE signal processing magazine*, 15(5), 101-130.
- Loizou, P. C. (2006). Speech processing in vocoder-centric cochlear implants. In *Cochlear and Brainstem Implants* (Vol. 64, pp. 109-143). Karger Publishers.
- Loizou, P. C., Dorman, M., & Tu, Z. (1999). On the number of channels needed to understand speech. *The Journal of the Acoustical Society of America*, 106(4), 2097-2103.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1-86.
- McMurray, B., Farris-Trimble, A., & Rigler, H. (2017). Waiting for lexical access: Cochlear implants or severely degraded input lead listeners to process speech less incrementally. *Cognition*, 169, 147-164.
- Miller, J. L. (1994). On the internal structure of phonetic categories: A progress report. *Cognition*, 50(1), 271-285.
- McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception? *Trends in Cognitive Sciences*, 10(8), 363-369.
- Moberly, A. C., Bates, C., Harris, M. S., & Pisoni, D. B. (2016). The enigma of poor

- performance by adults with cochlear implants. *Otology & Neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 37(10), 1522-1528.
- Moberly, A. C., Lowenstein, J. H., Tarr, E., Caldwell-Tarr, A., Welling, D. B., Shahin, A. J., & Nittrouer, S. (2014). Do adults with cochlear implants rely on different acoustic cues for phoneme perception than adults with normal hearing? *Journal of Speech, Language, and Hearing Research*, 57(2), 566-582.
- Most, T., Shrem, H., & Duvdevani, I. (2010). Cochlear implantation in late-implanted adults with prelingual deafness. *American Journal of Otolaryngology*, 31(6), 418-423.
- Neel, A. T. (2008). Vowel space characteristics and vowel identification accuracy. *Journal of Speech, Language, and Hearing Research*, 51(3), 574-585.
- Nelson, P. B., Jin, S. H., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *The Journal of the Acoustical Society of America*, 113(2), 961-968.
- Newman, R. S., Clouse, S. A., & Burnham, J. L. (2001). The perceptual consequences of within-talker variability in fricative production. *The Journal of the Acoustical Society of America*, 109(3), 1181-1196.
- Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N. Y., Quittner, A. L., ... & CDaCI Investigative Team. (2010). Spoken language development in children following cochlear implantation. *Journal of the American Medical Association*, 303(15), 1498-1506.
- Nittrouer, S., & Caldwell-Tarr, A. (2016). Language and literacy skills in children with cochlear

- implants: Past and present findings. In *Pediatric Cochlear Implantation* (pp. 177-197). Springer New York.
- Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, 47(2), 204-238.
- Parkin, J.L., Stewart, B.E., Dankowski, K., & Haas, L.J. (1990). Prognosticating speech performance in multi-channel cochlear implant patients. *Journal of Otolaryngology – Head & Neck Surgery*, 101, 314-319.
- Pelizzone, M., Cosendai, G., & Tinembart, J. (1999). Within-patient longitudinal speech reception measures with continuous interleaved sampling processors for ineraid implanted subjects. *Ear and Hearing*, 20(3), 228-237.
- Peterson, G.E., & Barney, H.L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 24, 175-184.
- Pisoni, D. B., Cleary, M., Geers, A. E., & Tobey, E. A. (1999). Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: New process measures of performance. *The Volta Review*, 101(3), 111-164.
- Pisoni, D. B., Kronenberger, W. G., Harris, M. S., & Moberly, A. C. (2017). Three challenges for future research on cochlear implants. *World Journal of Otorhinolaryngology – Head and Neck Surgery*, 3(4), 240-254.
- Prendergast, S. G., & Kelley, L. A. (2002). Aural rehab services: Survey reports who offers which ones and how often, and by whom. *The Hearing Journal*, 55(9), 30-34.
- Rayes, H., Al-Malky, G., & Vickers, D. (2019). Systematic review of auditory training in pediatric cochlear implant recipients. *Journal of Speech, Language, and Hearing Research*, 62(5), 1574-1593.

- Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science*, *212*(4497), 947-949.
- Schramm, D., Fitzpatrick, E., & Séguin, C. (2002). Cochlear implantation for adolescents and adults with prelinguistic deafness. *Otology & Neurotology*, *23*(5), 698-703.
- Schwab, E. C., Nusbaum, H. C., & Pisoni, D. B. (1985). Some effects of training on the perception of synthetic speech. *Human factors*, *27*(4), 395-408.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303-304.
- Sharma, A., Gilley, P. M., Dorman, M. F., & Baldwin, R. (2007). Deprivation-induced cortical reorganization in children with cochlear implants. *International Journal of Audiology*, *46*(9), 494-499.
- Shea III, J. J., Domico, E. H., & Orchik, D. J. (1990). Speech recognition ability as a function of duration of deafness in multichannel cochlear implant patients. *The Laryngoscope*, *100*(3), 223-226.
- Skinner, M. W., Holden, L. K., Whitford, L. A., Plant, K. L., Psarros, C., & Holden, T. A. (2002). Speech recognition with the nucleus 24 SPEAK, ACE, and CIS speech coding strategies in newly implanted adults. *Ear and Hearing*, *23*(3), 207-223.
- Spivak, L. G., & Waltzman, S. B. (1990). Performance of cochlear implant patients as a function of time. *Journal of Speech, Language, and Hearing Research*, *33*(3), 511-519.
- Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *The Journal of the Acoustical Society of America*, *64*(5), 1358-1368.
- Stevens, K. N., & Klatt, D. H. (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *The Journal of the Acoustical Society of America*, *55*(3), 653-659.

- Sweetow, R., & Palmer, C. V. (2005). Efficacy of individual auditory training in adults: A systematic review of the evidence. *Journal of the American Academy of Audiology, 16*(7), 494-504.
- Sweetow, R. W., & Sabes, J. H. (2007). Technologic advances in aural rehabilitation: Applications and innovative methods of service delivery. *Trends in Amplification, 11*(2), 101-111.
- Theodore, R. M., Miller, J. L., & DeSteno, D. (2009). Individual talker differences in voice-onset-time: Contextual influences. *The Journal of the Acoustical Society of America, 125*(6), 3974-3982.
- Tong, Y. C., Busby, P. A., & Clark, G. M. (1988). Perceptual studies on cochlear implant patients with early onset of profound hearing impairment prior to normal development of auditory, speech, and language skills. *The Journal of the Acoustical Society of America, 84*(3), 951-962.
- Son, R. J. V., & Pols, L. C. (1997). The correlation between consonant identification and the amount of acoustic consonant reduction. In *Fifth European Conference on Speech Communication and Technology*.
- Tye-Murray, N., & Tyler, R. S. (1988). A critique of continuous discourse tracking as a test procedure. *Journal of Speech and Hearing Disorders, 53*(3), 226-231.
- van Dijk, J. E., van Olphen, A. F., Langereis, M. C., Mens, L. H., Brokx, J. P., & Smoorenburg, G. F. (1999). Predictors of cochlear implant performance. *Audiology, 38*(2), 109-116.
- Walley, A. C., & Carrell, T. D. (1983). Onset spectra and formant transitions in the adult's and child's perception of place of articulation in stop consonants. *The Journal of the Acoustical Society of America, 73*(3), 1011-1022.

Waltzman, S. B., Cohen, N. L., & Shapiro, W. H. (1986). Long- Term effects of multichannel cochlear implant usage. *The Laryngoscope*, *96*(10), 1083-1087.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, *167*(3917), 392-393.

Wu, J. L., Yang, H. M., Lin, Y. H., & Fu, Q. J. (2007). Effects of computer-assisted speech training on Mandarin-speaking hearing-impaired children. *Audiology and Neurotology*, *12*(5), 307-312.