The Perception of Regional Dialects and Foreign Accents by Cochlear Implant Users

Tamati, Terrin N.; Pisoni, David B.; Moberly, Aaron C.

Published in:
Journal of Speech Language and Hearing Research

DOI:
10.1044/2020_JSLHR-20-00496

JSLHR

Research Note

The Perception of Regional Dialects and Foreign Accents by Cochlear Implant Users

Terrin N. Tamati,[a,b] David B. Pisoni,[c] and Aaron C. Moberly[a]

Purpose: This preliminary research examined (a) the perception of two common sources of indexical variability in speech—regional dialects and foreign accents, and (b) the relation between indexical processing and sentence recognition among prelingually deaf, long-term cochlear implant (CI) users and normal-hearing (NH) peers.
Method: Forty-three prelingually deaf adolescent and adult CI users and 44 NH peers completed a regional dialect categorization task, which consisted of identifying the region of origin of an unfamiliar talker from six dialect regions of the United States. They also completed an intelligibility rating task, which consisted of rating the intelligibility of short sentences produced by native and nonnative (foreign-accented) speakers of American English on a scale from 1 (*not intelligible at all*) to 7 (*very intelligible*). Individual performance was compared to demographic factors and sentence recognition scores.
Results: Both CI and NH groups demonstrated difficulty with regional dialect categorization, but NH listeners significantly outperformed the CI users. In the intelligibility rating task, both CI and NH listeners rated foreign-accented sentences as less intelligible than native sentences; however, CI users perceived smaller differences in intelligibility between native and foreign-accented sentences. Sensitivity to accent differences was related to sentence recognition accuracy in CI users.
Conclusions: Prelingually deaf, long-term CI users are sensitive to accent variability in speech, but less so than NH peers. Additionally, individual differences in CI users' sensitivity to indexical variability was related to sentence recognition abilities, suggesting a common source of difficulty in the perception and encoding of fine acoustic–phonetic details in speech.

In real-world environments, listeners experience an immense amount of talker variability, hearing speech signals that originate from multiple talkers with different voices and diverse linguistic and developmental histories (Abercrombie, 1967; Pisoni, 1997). In normal-hearing (NH) listeners, real-world speech communication is generally successful in spite of this variability. Speech understanding is facilitated by NH listeners' ability to perceive, encode, and retain highly detailed talker information in memory (Nygaard, 2008; Pisoni, 1993). NH listeners are able to use indexical information encoded in the speech signal to make judgments about a talker's identity (e.g., Van Lancker, Kreiman, & Emmorey, 1985; Van Lancker, Kreiman, & Wickens, 1985), gender (e.g., Lass et al., 1976), and region of origin (e.g., Labov, 1972). Furthermore, NH listeners are able to learn talker- and group-specific indexical patterns to facilitate speech recognition (e.g., Nygaard & Pisoni, 1998; Nygaard et al., 1994).

Indexical variability may present a challenge for hearing-impaired children and adults who have received cochlear implants (CIs) as a medical treatment for profound deafness. CI users must rely on a signal that is highly degraded in spectro-temporal detail due to limitations of the electrode–nerve interface and relatively broad electrical stimulation of the auditory nerve (for a review, see Başkent et al., 2016). As a result, compared to NH listeners, CI users are less sensitive to subtle indexical cues; for example, they

[a]Department of Otolaryngology, Wexner Medical Center, The Ohio State University, Columbus
[b]Department of Otorhinolaryngology, The University Medical Center Groningen, University of Groningen, the Netherlands
[c]DeVault Otologic Research Laboratory, Department of Otolaryngology—Head and Neck Surgery, Indiana University School of Medicine, Indianapolis
Correspondence to Terrin N. Tamati: Terrin.Tamati@osumc.edu

Editor-in-Chief: Frederick (Erick) Gallun
Editor: Christian E. Stilp
Received August 21, 2020
Revision received October 9, 2020
Accepted October 9, 2020
https://doi.org/10.1044/2020_JSLHR-20-00496
Publisher Note: Part of the data in the current manuscript was presented at the 18th International Congress of Phonetic Sciences held in Glasgow, Scotland, and appears in the *Proceedings of 18th International Congress of Phonetic Sciences*.

__Disclosure:__ Aaron C. Moberly serves as a paid consultant for Cochlear Americas and Advanced Bionics. The other authors declared that no other competing interests existed at the time of publication.

often demonstrate poor same-gender talker discrimination (Cleary & Pisoni, 2002; Cleary et al., 2005; Massida et al., 2011; McDonald et al., 2003). Furthermore, while CI users are generally able to achieve some gender discrimination (e.g., Massida et al., 2011), where the indexical cues are more disparate, they do so abnormally by weighting voice cues differently than NH listeners (Fuller et al., 2014). Additionally, CI users may be less able to make use of detailed, talker-specific, acoustic–phonetic information in speech to make nonlinguistic judgments about a talker's developmental and linguistic backgrounds, such as the talker's region of origin (e.g., Clopper & Pisoni, 2004; Hay-McCutcheon et al., 2018; Tamati et al., 2014). Similarly, limiting spectral resolution via acoustic simulations of CI hearing detection has been shown to result in poorer discrimination of native and foreign-accented speech in NH listeners (Kapolowicz et al., 2016).

Limitations in the perception of talker-specific indexical details may be associated with difficulties in adapting to and recognizing the linguistic content of speech in CI users (Cleary et al., 2005). Differences in the perception and encoding of fine acoustic–phonetic details of speech may influence listeners' ability to take advantage of episodic context information to make talker judgments as well as facilitate speech recognition (Nygaard & Pisoni, 1998; Nygaard et al., 1994). Poorer discrimination of talker details has previously been associated with less accurate recognition of highly variable speech in NH listeners (Tamati & Pisoni, 2014; Tamati et al., 2013) and CI users (Cleary et al., 2005; Rodman et al., 2020). CI users have also been found to have more difficulty recognizing highly variable speech produced by multiple talkers from different dialect regions (e.g., Sommers et al., 1997; Tamati et al., 2020) and foreign-accented talkers (Ji et al., 2014; Kapolowicz et al., 2020), compared to idealized lab speech, which typically involves carefully articulated speech produced by a single talker with no discernable accent. Similarly, studies using simulations of CI hearing have also reported less accurate recognition of highly variable speech (Faulkner et al., 2015) and foreign-accented speech (Kapolowicz et al., 2016, 2018). Thus, although CIs improve hearing, different sources of indexical variability may pose substantial problems to speech recognition.

The main goal of the current study was to examine the perception of regional dialects and foreign accents—two common sources of indexical variability in speech—in prelingually deafened adolescents and adults with long-term CI use. The prelingually deafened CI users were congenitally deafened or had experienced hearing loss very early in childhood. As such, they had experienced little or no access to sound during early development and had learned speech and language with their CIs. While previous studies suggest that reduced spectral resolution in CI users results in a relative deficit in the perception of detailed talker information, little is known about how early sensory deprivation and exposure to talker variability primarily via the degraded signals of the CI impacts the perception of talker variability, with long-term CI use. A secondary goal was to explore the relation between indexical processing and speech recognition skills in CI users, who routinely display enormous variability in speech recognition outcomes that largely remains unexplained (e.g., Ruffin et al., 2013; van Wieringen & Wouters, 2015). Therefore, CI users and a group of NH peers completed two indexical processing tasks, including a regional dialect categorization task and a foreign-accent intelligibility rating task, and sentence recognition tasks. Based on previous studies (e.g., Clopper & Pisoni, 2004; Hay-McCutcheon et al., 2018; Tamati et al., 2014), we hypothesized that CI users would have more difficulty than the NH listeners at categorizing unfamiliar talkers by region of origin based only on acoustic–phonetic dialectal features, as well as in perceiving and using accent information to rate the intelligibility of foreign-accented and native speech. We further hypothesized that better indexical processing skills would be related to more accurate sentence recognition.

Method

Participants

A group of prelingually deaf, long-term CI users and a group of NH peers participated in the current study. The CI group consisted of 43 prelingually deaf, long-term CI users who had received their CIs early in childhood. At the time of testing, the mean age of the CI users was 17.2 years (SD = 5.4), with a range of 9.3–30.0 years. The average age of CI implantation was 3.2 years (SD = 1.7), with a range of 0.7–6.3 years, and the average length of CI use was 14.0 years (SD = 4.5), with a range of 7.3–24.5 years. Twenty-four were unilateral CI users, 17 were bilateral CI users, and only two were bimodal users (CI + hearing aid in the contralateral ear). The CI users used their everyday CI programming settings during testing, but the two bimodal CI users did not use their hearing aids during testing to reduce the possible effects of residual hearing on performance. The nonimplanted ear was not plugged for bimodal or unilateral CI users. However, preoperative, better-ear unaided pure-tone average at 500 Hz, 1000 Hz, and 2000 Hz was 85 dB HL or worse for all CI users, suggesting that residual hearing was likely too poor to have significantly impacted performance. The NH group consisted of 44 adolescents and adults with NH, with a mean age of 17.8 years (SD = 5.3), and a range of 10.0–29.3 years. The CI and NH listener groups did not differ significantly in age, $t(85) = 0.53$, $p = .598$. All listeners were recruited from central Indiana as part of a larger study of long-term CI outcomes and benefits (Kronenberger et al., 2013). Local institutional review board approval was obtained, and all participants provided written informed consent prior to participation.

Procedures and Materials

Participants completed two indexical processing tasks: a regional dialect categorization task and a foreign accent intelligibility rating task. Due to testing and timing constraints, a small number of participants were unable to complete both tasks. A total of 39 CI users and 44 NH listeners completed the regional dialect categorization task, and a total of 41 CI users and 43 NH listeners completed the foreign accent intelligibility rating task. For comparison, sentence recognition scores were obtained for CI users from the larger study (Kronenberger et al., 2013). Participants were tested individually in sound-attenuated rooms under the supervision of speech-language pathologists. They were seated in front of a computer touchscreen monitor and a high-quality external speaker (Advent AV570, Audiovox Electronics), located approximately 1 m from them. All stimulus items were presented at 65 dB SPL via the speaker.

Regional dialect categorization task. A regional dialect categorization task was used to assess the listeners' ability to categorize unfamiliar talkers by their region of origin, requiring sensitivity to dialect-specific, acoustic–phonetic cues as well as use of stored knowledge of regional dialect variation in the United States, based on previous studies (Clopper & Pisoni, 2004). Twelve talkers (six women and six men) from the Texas Instruments/Massachusetts Institute of Technology acoustic–phonetic speech corpus (Garofolo et al., 1993) were selected for this task. Two talkers (one woman and one man) were from each of the following six dialect regions: New England, North Midland, South Midland, North, South, and West. Throughout the task, participants were presented with a single talker producing a sentence and were asked to select the region where they thought the talker was from using a closed set of six response alternatives, choosing from the six dialect regions represented on a graphical map of the United States displayed on a computer monitor. Participants entered their responses by touching a labeled box located within each dialect region on the computer touchscreen.

The task consisted of six practice trials and 12 test trials. During the practice trials, the participants listened to one talker from each dialect region producing the sentence, "She had your suit in greasy wash water all year." This sentence is one of the baseline calibration sentences collected from all talkers in the Texas Instruments/Massachusetts Institute of Technology database and was designed to obtain dialect-specific, acoustic–phonetic differences. For the practice trials, participants heard a single talker but could listen to that talker as many times as they wanted before responding when they felt ready. They did not receive feedback, and the next trial began following their response.

After the practice trials, they began the test trials when ready. For the test trials, the participants listened to the entire set of 12 talkers producing one unique sentence, which was selected to contain representative phonetic features of each dialect region. Sentences did not contain dialect-specific lexical or grammatical features. On each trial, participants heard a single talker but could only listen to that talker once. Participants could take as long as they wanted to respond, and once they responded, the next trial began. Again, no feedback was given. Responses were collected and coded for the dialect region selected and scored for accuracy.

Foreign accent intelligibility rating task. The foreign accent intelligibility rating task was used to assess a listener's ability to discriminate the intelligibility of unfamiliar native and nonnative talkers whose speech varied by accent strength. Seventeen talkers (eight women and nine men) were selected for this task from the Multi-talker Corpus of Foreign-Accented English (Tamati et al., 2011). Nine talkers (four women and five men) were nonnative speakers of English with nine different native languages (Japanese, Kannada, Korean, Mandarin, Portuguese [Brazil], Portuguese [Portugal], Spanish [Columbia], Taiwanese, and Turkish). Talkers with multiple native languages were selected in order to include talkers with a wide range of baseline accentedness and intelligibility. Furthermore, using multiple native languages allowed us to obtain a broad measure of foreign accent perception that was not language dependent, and that would not be greatly influenced in a case where a listener has familiarity with any one foreign accent or native language. The other eight talkers (four women and four men) were monolingual native speakers of American English from a General American dialect region (Midland, West, parts of New England). Eighteen unique high probability or low probability Speech Perception in Noise (SPIN) sentences (Kalikow et al., 1977) were used in the task, which consisted of two practice trials and 16 test trials. The practice trials were given so that the participants could become familiar with the task methodology. On the practice trials, the participants heard two trials; each practice trial included a nonnative speaker of English producing a unique high probability sentence. One of the male nonnative talkers was randomly selected for the practice trials. On the test trials, the participants listened to all the other 16 talkers (eight nonnative and eight native) producing a unique high or low probability sentence.

For the two practice trials, participants listened to a sentence produced by a single talker and were asked to indicate how intelligible they thought the talker was using a scale from 1 (*not intelligible at all*) to 7 (*very intelligible*). The participants could replay the sentence as many times as they wanted before responding. Participants responded by touching a dialog box labeled with one of the response alternatives (i.e., the numbers 1–7, on the touchscreen monitor). No feedback was given. Participants began the test trials when ready. On each test trial, participants were again presented with a sentence produced by a single talker, but they could only listen to that sentence 1 time. Participants could take as long as they wanted to respond, but again responded by touching a labeled box with the response alternatives on the touchscreen monitor. Once they responded, the next trial began. Again, no feedback was given.

Sentence recognition measures. Sentence recognition abilities were assessed with simple sentences produced by a single talker (Hearing in Noise Test for Children [HINT-C]; Nilsson et al., 1994), sentences produced by multiple talkers from different U.S. dialect regions (Perceptually Robust English Sentence Test: Open-set [PRESTO]; Gilbert et al., 2013), and foreign-accented sentences produced by multiple nonnative talkers (Perceptually Robust English Sentence Test

Open-Set—Foreign Accented English [PRESTO-FAE]). The sentence recognition tasks involved one 10-sentence list of HINT-C; two PRESTO lists consisting of 18 sentences, where each sentence is spoken by a different talker across U.S. dialect regions; and one list of PRESTO-FAE sentences, consisting of 26 SPIN sentences produced by 28 (14 men and 14 women) nonnative speakers of American English with different native languages. Test order was randomized for each participant. Note that HINT-C scores were not collected from the NH listeners because of concerns with possible ceiling effects. All sentences were presented in quiet, and participants repeated the words they heard. Sentence recognition scores represent percent correct keywords recognized.

Data Analyses

Groups were examined for differences in dialect categorization using independent-samples *t* tests. Foreign accent intelligibility ratings for the groups were compared using a two-way mixed design analysis of variance with the talker's native language (native or nonnative) as the within-subject factor and listener group (CI or NH) as the between-subjects factor. Individual differences in performance on the two indexical processing tasks were explored using correlational analyses with demographic variables and performance on additional speech recognition tasks from the larger study (Kronenberger et al., 2013). For all measures, an alpha of .05 was set. When $p > .05$, correlations are reported as *not significant*. Bivariate correlations are shown, with a Holm–Bonferroni correction applied to correct for multiple comparisons. To normalize sentence recognition scores for statistical analyses, proportion correct responses on sentence recognition tasks were transformed into rationalized arcsine units (Studebaker, 1985); however, scores in tables and figures are reported as proportion correct.

Results

Group Differences

A summary of performance across all tasks is provided in Table 1. The regional dialect categorization task was difficult for both the CI and NH groups. Overall accuracy was 18.6% (SD = 12.5) for the CI group and 26.5% (SD = 11.8) for the NH group, as displayed in Figure 1. The performance of the CI group was not significantly different than chance, $t(38) = 0.96$, $p = .341$, which is approximately 16.7% for the six-alternative, forced-choice categorization task. An independent-samples *t* test demonstrated that the NH group was significantly more accurate than the CI group, $t(81) = 2.97$, $p = .004$, and performed significantly above chance level, $t(43) = 5.53$, $p < .001$.

On the foreign accent intelligibility task, both CI and NH groups rated the native talkers as being on average more intelligible than the nonnative talkers, as displayed in Figure 2. CI users gave an average intelligibility rating of 4.6 (SD = 1.3) for native talkers and an average rating of 3.7 (SD = 0.8) for nonnative talkers, while the NH listeners gave an average rating of 6.5 (SD = 0.5) for native talkers and an average rating of 4.5 (SD = 0.8) for nonnative talkers. A repeated-measures analysis of variance on the intelligibility ratings with talker accent (native or nonnative) as the within-subject factor and listener group (CI or NH) as the between-subjects factor revealed significant main effects of talker accent, $F(1, 82) = 236.1$, $p < .001$, and listener group, $F(1, 82) = 63.0$, $p < .001$, and a significant talker Accent × Listener Group interaction, $F(1, 82) = 27.6$, $p < .001$. To explore the interaction, post hoc paired

Table 1. Performance scores on indexical processing and sentence recognition tasks for cochlear implant (CI) users and normal-hearing (NH) listeners.

	CI users *M* (*SD*)	NH listeners *M* (*SD*)
Indexical processing tasks		
Regional dialect categorization accuracy	18.6 (12.5)	26.5 (11.8)
Intelligibility ratings		
-Native	4.6 (1.3)	6.5 (0.5)
-Nonnative	3.7 (0.8)	4.5 (0.8)
-Difference scores (native – nonnative)	1.0 (1.0)	2.0 (0.8)
Sentence recognition accuracy		
HINT-C	79.8 (30.0)	N/A
PRESTO	48.5 (27.1)	96.1 (4.2)
PRESTO-FAE	32.8 (19.2)	81.9 (6.9)

Note. N/A = not applicable; HINT-C = Hearing in Noise Test for Children; PRESTO-FAE = Perceptually Robust English Sentence Test Open-Set—Foreign Accented English.

Figure 1. Box plot showing the mean regional dialect categorization accuracy (% Correct) by cochlear implant (CI) users (open box) and normal-hearing (NH) listeners (filled grey box). The boxes extend from the lower to the upper quartile (the interquartile range, IQ), the solid midline indicates the median, and the star indicates the mean. The whiskers indicate the highest and lowest values no greater than 1.5 times the IQ, and the plus signs indicate outliers, which are defined as data points larger than 1.5 times the IQ. The grey dashed line represents chance performance on the categorization task (16.7%). Individual participant scores are shown to the right of their corresponding box plot.

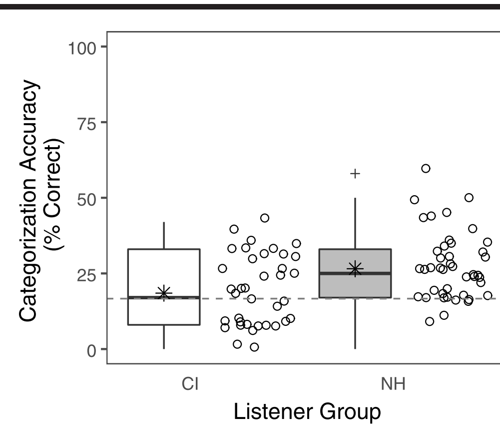
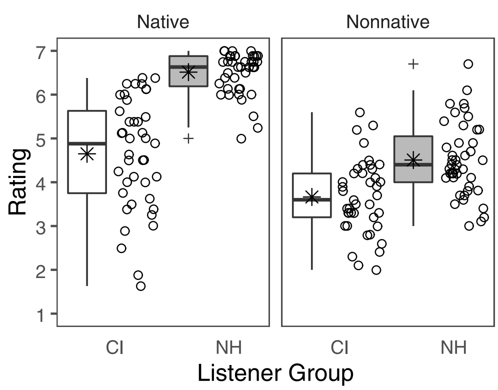


Figure 2. Box plot showing the mean intelligibility ratings for native talkers (left panel) and nonnative talkers (right panel) by cochlear implant (CI) users (open boxes) and normal-hearing (NH) listeners (filled grey boxes). The boxes extend from the lower to the upper quartile (the interquartile range, IQ), the solid midline indicates the median, and the star indicates the mean. The whiskers indicate the highest and lowest values no greater than 1.5 times the IQ, and the plus signs indicate outliers, which are defined as data points larger than 1.5 times the IQ. Individual participant scores are shown to the right of their corresponding box plot.



Scale: 1 = *not intelligible at all* to 7 = *very intelligible*

comparison *t* tests were carried out on ratings for each listener group. Native talker ratings were significantly higher than nonnative talker ratings for both the CI, $t(40) = 6.3$, $p < .001$, and NH, $t(42) = 17.0$, $p < .001$, groups.

To further explore the talker Accent × Listener Group interaction and to obtain a measure of the sensitivity to the difference between native and foreign-accented speech, difference scores were calculated by subtracting the average intelligibility rating for the nonnative talkers from the average intelligibility rating for the native talkers (native − nonnative). The NH listeners showed significantly larger differences scores ($M = 2.0$, $SD = 0.8$) than CI users ($M = 1.0$, $SD = 1.0$), $t(82) = 5.3$, $p < .001$.

Individual Differences

To investigate individual differences, we first examined the relation between individual performance on the two indexical processing tasks. For CI users, regional dialect categorization accuracy was not related to intelligibility ratings for both native ($r = -.03$, $p = .84$) and nonnative talkers ($r = .07$, $p = .69$), or to difference scores ($r = -.1$, $p = .56$), likely reflecting restricted range due to floor effects in this listener population. For NH listeners, regional dialect categorization accuracy was significantly related to intelligibility ratings for both native ($r = .37$, $p = .014$) and nonnative talkers ($r = .35$, $p = .02$), but not difference scores ($r = -.1$, $p = .67$).

Second, a series of correlational analyses was carried out between performance on the indexical processing tasks with accuracy on the additional set of speech recognition scores and demographic measures (i.e., age, age at CI implantation, length of CI use). Results are displayed in Table 2. For the CI users, regional dialect categorization accuracy scores were not related to sentence recognition accuracy on any task after Holm–Bonferroni correction, again possibly reflecting floor effects. Intelligibility ratings for native talkers were related to sentence recognition accuracy on all sentence recognition tasks after Holm–Bonferroni correction ($r = .59$–$.63$, all $p < .001$), but not to any demographic factor. Intelligibility ratings for nonnative talkers were not significantly related to any sentence recognition or demographic factor. Difference scores were related to sentence recognition accuracy on all sentence recognition tasks after Holm–Bonferroni correction ($r = .54$–$.65$, all $p < .001$), but not significantly related to any demographic factor. For the NH listeners, none of the indexical processing measures were significantly related to any sentence recognition tasks, or to age, likely reflecting restricted range due to ceiling effects.

Discussion

The main goal of this study was to assess the perception of two common sources of indexical variability—regional dialects and foreign accents—by prelingually deaf, long-term CI users and NH peers. Based on previous findings, we predicted that the CI users would have more difficulty than NH peers at reliably perceiving and using the regional dialect and foreign accent information. Consistent with our initial hypothesis, CI users showed less accurate regional dialect categorization and perceived less difference in intelligibility between native and nonnative listeners, compared to NH listeners. These results suggest that CI users were less sensitive to the dialect and accent information and/or less able to reliably use this information to make categorization and discrimination judgments.

Both CI and NH listeners demonstrated relatively poor performance on the regional dialect categorization task, compared to previous research with NH adults and adult CI users (e.g., Clopper & Pisoni, 2004; Tamati et al., 2014). In the current study, CI users were around chance performance of 16.7% correct ($M = 18.6\%$, $SD = 12.5$), and NH listeners were slightly, but significantly, above chance ($M = 26.5\%$, $SD = 11.8$). In previous studies, NH listeners have achieved around 28%–34% categorization accuracy with similar materials (e.g., Clopper & Pisoni, 2004; Tamati et al., 2014), and two individual postlingually deafened, adult CI users have also been able to reach 27.1% and 32.3% accuracy (Clopper & Pisoni, 2004; Tamati et al., 2014). Although the participants in the current study were younger than the participants in those previous studies, performance was not significantly correlated with age. The relatively poor performance observed in the current study may have been at least partly related to task difficulty. Given that the task involved two talkers from each dialect region, it is possible that these talkers were not strongly representative of the typical speech patterns of their region of origin. Nevertheless, CI users were significantly less accurate than NH peers, with categorization around chance level,

Table 2. Pearson correlations between scores on indexical processing tasks and demographic and other speech recognition measures for cochlear implant (CI) users (left) and normal-hearing (NH) listeners (right).

		CI users				NH listeners			
			Intelligibility ratings				Intelligibility ratings		
		Regional dialect categorization	Native	Nonnative	Difference Score (native – nonnative)	Regional dialect categorization	Native	Nonnative	Difference score (native – nonnative)
Demographics									
Age at testing	*r*	–.08	–.27	–.19	–.19	.14	.21	.16	–.05
	p	.629	.086	.248	.236	.361	.182	.292	.753
Age at CI implantation	*r*	–.23	–.18	–.16	–.10	N/A	N/A	N/A	N/A
	p	.153	.258	.334	.536				
Length of CI use	*r*	.0	–.26	–.16	–.19	N/A	N/A	N/A	N/A
	p	1.0	.104	.305	.236				
Sentence recognition									
HINT-C	*r*	–.23	**.59**	.12	**.65**	N/A	N/A	N/A	N/A
	p	.166	**< .001**	.325	**< .001**				
PRESTO	*r*	–.06	**.60**	.26	**.54**	–.02	–.01	.09	–.11
	p	.744	**< .001**	.107	**< .001**	.890	.970	.552	.500
PRESTO-FAE	*r*	–.10	**.63**	.32	**.54**	–.15	–.11	.17	–.26
	p	.559	**< .001**	.043	**< .001**	.329	.483	.269	.096

Note. Bolded values represent comparisons that are significant after Holm–Bonferroni correction. N/A = not applicable; HINT-C = Hearing in Noise Test for Children; PRESTO-FAE = Perceptually Robust English Sentence Test Open-Set—Foreign Accented English.

demonstrating that many were unable to reliably detect and use dialect-specific features to categorize talkers by region of origin. Moreover, compared with postlingually deafened adult CI users, prelingually deafened CI users likely do not benefit from prior linguistic experience to the same extent. For the CI users in the current study, dialect-specific information may only be weakly encoded and stored in memory, resulting in further difficulties in using stored knowledge to identify the region of origin of an unfamiliar talker, particularly when no supporting context or category labels are provided (e.g., Clopper, 2008). Thus, the CI users' poor performance on the regional dialect categorization task in the current study is likely related both to their developmental history and prior linguistic experience with compromised auditory input, as well as task difficulty.

In the intelligibility rating task, both CI and NH listeners rated nonnative talkers as less intelligible than native talkers, suggesting that both were sensitive to intelligibility differences between native and nonnative speech. Interestingly, CI users rated native and foreign-accented speech as less intelligible compared to the NH listeners, which suggests that both groups are sensitive to their recognition abilities and is consistent with previous findings demonstrating poorer spoken word recognition scores than NH listeners even after long-term CI use (e.g., Ruffin et al., 2013). Further comparing the listener groups, CI users did not perceive the foreign-accented speech to be drastically less intelligible than the native speech, as suggested by the smaller difference scores compared to the NH listeners. Thus, consistent with previous findings with NH listeners under CI simulation (Kapolowicz et al., 2016), the CI users in the current study were overall less sensitive to accent information, at least with respect to intelligibility judgments. Accent information, especially cues conveyed by fine spectral detail, may be poorly encoded via the CI, resulting in smaller perceived differences between native and foreign-accented speech compared to NH listeners. Interestingly, however, smaller difference scores in CI users appeared to be mainly related to lower intelligibility ratings for native rather than nonnative speech (see Figure 2). The relatively low intelligibility ratings for native speech may reflect underlying processing differences in CI users, in addition to differences in just hearing or audibility, as mentioned above. Prelingually deafened CI users may rely upon degraded and partially specified phonological representations for spoken words (e.g., Nittrouer et al., 2014; Pisoni et al., 1999), reflecting poor encoding of fine acoustic–phonetic episodic details as well as the indexical properties of the vocal source. As a result, speech recognition is challenging even when listening to native signals, potentially with relatively little difference in perceived degree of intelligibility between native and nonnative signals.

A second goal of the current study was to examine individual differences in performance on the two indexical processing tasks, and the relation between indexical processing skills and speech recognition abilities. For CI users, regional dialect categorization scores were not related to sentence recognition or demographic factors. The lack of association between regional dialect categorization and sentence recognition may again reflect the task difficulty, as many CI users were unable to complete the categorization task at an above-chance level (see Figure 1). Stronger associations were observed between intelligibility ratings and sentence recognition scores. CI users who rated native speech as being more intelligible showed more accurate sentence recognition, across the four sets of sentence materials. Similarly, those who had larger difference scores (likely, due to having rated native listeners as more intelligible) also showed more accurate sentence recognition. The NH listeners' indexical processing scores were not related to sentence recognition accuracy, likely because the NH listeners generally showed near-ceiling accuracy on the sentence recognition tasks. The findings with the CI users suggest that individual differences on these tasks reflect fundamental differences in encoding fine acoustic–phonetic episodic details as well as the indexical properties of the vocal source, and are consistent with previous findings demonstrating a close link between the perception of linguistic information and the perception of indexical information in speech (e.g., Cleary & Pisoni, 2002; Cleary et al., 2005; Tamati et al., 2013).

Several weaknesses of the current study should be noted. Because the current study used an intelligibility rating task that only indirectly examines the perception of foreign-accented speech, future studies should use a task that allows for more direct assessment of the perception of foreign-accented speech. Furthermore, to more fully understand potential sources of individual differences in the perception of indexical variability in speech, language experience and other demographic factors, as well as the settings and characteristics of listeners' CI devices, should be taken into account in future studies (e.g., van Wieringen & Wouters, 2015). Finally, other sources of indexical variability, such as within-talker variability in speaking style and emotion, and their relation to speech recognition outcomes in this population should also be considered.

The current study was a first step in understanding how early sensory deprivation and long-term exposure to speech primarily through a CI influences the perception of indexical variability in prelingually deaf, long-term CI users. The results of this study suggest that CI users are able to discriminate native and nonnative speech, but are overall less sensitive to accent variability than NH listeners. Furthermore, CI users' sensitivity to indexical variability in speech, at least for the intelligibility rating task, was related to sentence recognition skills. These findings further establish that linguistic and indexical channels of speech are closely coupled and linked together in speech perception. CI users have a selective weakness in the encoding and processing of fine acoustic–phonetic details of the signal, displayed in both the indexical processing and sentence recognition tasks. Future studies will be required to further elucidate the relations of linguistic and indexical processing of speech, along with the implications of these relations to real-world functioning in CI users.

Acknowledgments

Aaron C. Moberly received grant funding support from Cochlear Americas for an unrelated investigator-initiated research study. Preparation of this research note was primarily supported by National Institute on Deafness and Other Communication Disorders Training Grant T32-DC00012 and National Institute on Deafness and Other Communication Disorders Research Grant R01-DC00111 awarded to David B. Pisoni, and in part by VENI Grant No. 275-89-035 from the Netherlands Organization for Scientific Research and funding from the President's Postdoctoral Scholars Program at The Ohio State University awarded to Terrin N. Tamati.

References

Abercrombie, D. (1967). *Elements of general phonetics*. Aldine.

Başkent, D., Gaudrain, E., Tamati, T. N., & Wagner, A. (2016). Perception and psychoacoustics of speech in cochlear implant users. In A. T. Cacace, E. de Kleine, A. G. Holt, & P. van Dijk (Eds.), *Scientific foundations of audiology: Perspectives from physics, biology, modeling, and medicine* (pp. 285–319). Plural.

Cleary, M., & Pisoni, D. B. (2002). Talker discrimination by prelingually deaf children with cochlear implants: Preliminary results. *Annals of Otology, Rhinology & Laryngology, 111*(Suppl. 5), 113–118. https://doi.org/10.1177/00034894021110S523

Cleary, M., Pisoni, D. B., & Kirk, K. I. (2005). Influence of voice similarity on talker discrimination in children with normal hearing and children with cochlear implants. *Journal of Speech, Language, and Hearing Research, 48*(1), 204–223. https://doi.org/10.1044/1092-4388(2005/015)

Clopper, C. G. (2008). Auditory free classification: Methods and analysis. *Behavior Research Methods, 40,* 575–581. https://doi.org/10.3758/BRM.40.2.575

Clopper, C. G., & Pisoni, D. B. (2004). Perceptual dialect categorization by an adult cochlear implant user: A case study. *International Congress Series, 1273,* 235–238. https://doi.org/10.1016/j.ics.2004.08.012

Faulkner, K. F., Tamati, T. N., Gilbert, J. L., & Pisoni, D. B. (2015). List equivalency for the clinical evaluation of speech recognition with PRESTO. *Journal of the American Academy of Audiology, 26*(6), 582–594. https://doi.org/10.3766/jaaa.14082

Fuller, C. D., Gaudrain, E., Clarke, J. N., Galvin, J. J., Fu, Q. J., Free, R. H., & Başkent, D. (2014). Gender categorization is abnormal in cochlear implant users. *Journal of the Association of Research in Otolaryngology, 15*(6), 1037–1048. https://doi.org/10.1007/s10162-014-0483-7

Garofolo, J. S., Lamel, L. F., Fisher, W. M., Fiscus, J. G., Pallett, D. S., & Dahlgren, N. L. (1993). The DARPA TIMIT acoustic–phonetic continuous speech corpus. *Linguistic Data Consortium.* https://doi.org/10.6028/NIST.IR.4930

Gilbert, J. L., Tamati, T. N., & Pisoni, D. B. (2013). Development, reliability, and validity of PRESTO: A new high-variability sentence recognition test. *Journal of the American Academy of Audiology, 24*(1), 26–36. https://doi.org/10.3766/jaaa.24.1.4

Hay-McCutcheon, M. J., Peterson, N. R., Pisoni, D. B., Kirk, K. I., Yang, X., & Parton, J. (2018). Performance variability on perceptual discrimination tasks in profoundly deaf adults with cochlear implants. *Journal of Communication Disorders, 72,* 122–135. https://doi.org/10.1016/j.jcomdis.2018.01.005

Ji, C., Galvin, J. J., Chang, Y., Xu, A., & Fu, Q.-J. (2014). Perception of speech produced by native and nonnative talkers by listeners with normal hearing and listeners with cochlear implants. *Journal of Speech, Language, and Hearing Research, 57*(2), 532–554. https://doi.org/10.1044/2014_JSLHR-H-12-0404

Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America, 61*(5), 1337–1351. https://doi.org/10.1121/1.381436

Kapolowicz, M. R., Montazeri, V., & Assmann, P. F. (2016). The role of spectral resolution in foreign-accented speech perception. In N. Morgan (Ed.), *Proceedings of the 17th Annual Conference of the International Speech Communication Association, Interspeech*

2016 (pp. 3289–3293). International Speech Communication Association. https://doi.org/10.21437/Interspeech.2016-1585

Kapolowicz, M. R., Montazeri, V., & Assman, P. F. (2018). Perceiving foreign-accented speech with decreased spectral resolution in single- and multiple-talker conditions. *The Journal of the Acoustical Society of America, 143*(2), EL99. https://doi.org/10.1121/1.5023594

Kapolowicz, M. R., Montazeri, V., Baese-Berk, M. M., & Assman, P. F. (2020). Rapid adaptation to non-native speech is impaired in cochlear implant users. *The Journal of the Acoustical Society of America, 148*(3), EL267. https://doi.org/10.1121/10.0001941

Kronenberger, W. G., Pisoni, D. B., Harris, M. S., Hoen, H. M., Xu, H., & Miyamoto, R. T. (2013). Profiles of verbal working memory growth predict speech and language development in children with cochlear implants. *Journal of Speech, Language, and Hearing Research, 56*(3), 805–825. https://doi.org/10.1044/1092-4388(2012/11-0356)

Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.

Lass, N. J., Hughes, K. R., Bowyer, M. D., Waters, L. T., & Bourne, V. T. (1976). Speaker sex identification from voiced, whispered, and filtered isolated vowels. *The Journal of the Acoustical Society of America, 59*(3), 675–678. https://doi.org/10.1121/1.380917

Massida, Z., Belin, P., James, C., Rouger, J., Fraysse, B., Barone, P., & Deguine, O. (2011). Voice discrimination in cochlear-implanted deaf children. *Hearing Research, 275*(1–2), 120–129. https://doi.org/10.1016/j.heares.2010.12.010

McDonald, C., Kirk, K. I., Krueger, T., & Houston, D. (2003). *Talker discrimination and spoken word recognition by adults with cochlear implants*. Poster presented at the 26th Midwinter Meeting of the Association for Research in Otolaryngology, St. Petersburg, FL, United States.

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America, 95*(2), 1085–1099. https://doi.org/10.1121/1.408469

Nittrouer, S., Sansom, E., Low, K., Rice, C., & Caldwell-Tarr, A. (2014). Language structures used by kindergartners with cochlear implants: Relationship to phonological awareness, lexical knowledge and hearing loss. *Ear and Hearing, 35*(5), 506–518. https://doi.org/10.1097/AUD.0000000000000051

Nygaard, L. C. (2008). Perceptual integration of linguistic and nonlinguistic properties of speech. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 390–413). Blackwell. https://doi.org/10.1002/9780470757024.ch16

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific perceptual learning in speech perception. *Perception & Psychophysics, 60,* 355–376. https://doi.org/10.3758/BF03206860

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science, 5*(1), 42–46. https://doi.org/10.1111/j.1467-9280.1994.tb00612.x

Pisoni, D. B. (1993). Long-term memory in speech perception: Some new findings on talker variability, speaking rate and perceptual learning. *Speech Communication, 13*(1–2), 109–125. https://doi.org/10.1016/0167-6393(93)90063-Q

Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9–32). Academic Press.

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. *Volta Review, 101*(3), 111–164.

Rodman, C., Moberly, A. C., Janse, E., Başkent, D., & Tamati, T. N. (2020). The impact of speaking style on speech recognition in quiet and multi-talker babble in adult cochlear implant users. *The Journal of the Acoustical Society of America, 147*(1), 101–107. https://doi.org/10.1121/1.5141370

Ruffin, C. V., Kronenberger, W. G., Colson, B. G., Henning, S. C., & Pisoni, D. B. (2013). Long-term speech and language outcomes in prelingually deaf children, adolescents and young adults who received cochlear implants in childhood. *Audiology and Neurotology, 18*(5), 289–296. https://doi.org/10.1159/000353405

Sommers, M. S., Kirk, K. I., & Pisoni, D. B. (1997). Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal-hearing, and cochlear implant listeners. I: The effects of response format. *Ear and Hearing, 18*(2), 89–99. https://doi.org/10.1097/00003446-199704000-00001

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research, 28*(3), 455–462. https://doi.org/10.1044/jshr.2803.455

Tamati, T. N., Gilbert, J. L., & Pisoni, D. B. (2013). Some factors underlying individual differences in speech recognition on PRESTO: A first report. *Journal of the American Academy of Audiology, 24*(7), 616–634. https://doi.org/10.3766/jaaa.24.7.10

Tamati, T. N., Gilbert, J. L., & Pisoni, D. B. (2014). Influence of early linguistic experience on dialect categorization by an adult cochlear implant user: A case study. *Ear and Hearing, 35*(3), 383–386. https://doi.org/10.1097/AUD.0000000000000016

Tamati, T. N., Park, H.-Y., & Pisoni, D. P. (2011). *The development of a new corpus of foreign-accented English*. Poster presented at the 2011 VLSP New Tools and Methods for Very Large-Scale Phonetics Research Workshop. Philadelphia, PA, United States.

Tamati, T. N., & Pisoni, D. B. (2014). Non-native listeners' recognition of high-variability speech using PRESTO. *Journal of the American Academy of Audiology, 25*(9), 869–892. https://doi.org/10.3766/jaaa.25.9.9

Tamati, T. N., Ray, C., Vasil, K. J., Pisoni, D. B., & Moberly, A. C. (2020). High- and low-performing adult cochlear implant users on high-variability sentence recognition: Differences in auditory spectral resolution and neurocognitive functioning. *Journal of the American Academy of Audiology, 31*(5), 324–335. https://doi.org/10.3766/jaaa.18106

Van Lancker, D., Kreiman, J., & Emmorey, K. (1985). Familiar voice recognition: Patterns and parameters: Part I. Recognition of backward voices. *Journal of Phonetics, 13*(1), 19–38. https://doi.org/10.1016/S0095-4470(19)30723-5

Van Lancker, D., Kreiman, J., & Wickens, T. (1985). Familiar voice recognitions: Patterns and parameters: Part II. Recognition of rate-altered voices. *Journal of Phonetics, 13*(1), 39–52. https://doi.org/10.1016/S0095-4470(19)30724-7

van Wieringen, A., & Wouters, J. (2015). What can we expect of normally-developing children implanted at a young age with respect to their auditory, linguistic and cognitive skills? *Hearing Research, 332,* 171–179. https://doi.org/10.1016/j.heares.2014.09.002