THE PERCEPTION OF FOREIGN-ACCENTED SPEECH BY COCHLEAR IMPLANT USERS

Terrin N. Tamati and David B. Pisoni

University Medical Center Groningen, Indiana University
t.n.tamati@umcg.nl, pisoni@indiana.edu

ABSTRACT

The current study investigated the perception of foreign-accented speech by prelingually deaf, long-term cochlear implant (CI) users. CI and normal-hearing (NH) listeners made judgements about the intelligibility of short sentences produced by native and non-native speakers of American English. While both CI and NH listeners rated foreign-accented sentences as less intelligible than native sentences, the CI listeners perceived smaller differences in intelligibility between foreign-accented and native sentences. However, the CI listeners demonstrated substantial individual variability in their ratings. Additional analyses of these individual differences showed that CI listeners who were more sensitive to foreign accent differences also had better speech perception abilities. Taken together, these results suggest that CI listeners are sensitive to foreign accents, but less so than NH listeners. Further, their sensitivity to this source of variability in the speech signal may reflect the development and use of basic speech and language processing skills.

Keywords: speech perception, foreign accent, cochlear implants, individual differences

1. INTRODUCTION

In real-world listening environments, listeners encounter many different talkers with diverse developmental, linguistic, and social backgrounds. The speech signal conveys highly detailed indexical information about the talkers and their backgrounds, along with the linguistic content (i.e., sounds, words, etc.) of the utterance. To achieve robust speech communication, listeners must rapidly process both the linguistic and indexical information in the speech signal to recover the intended meaning of the utterance, while also extracting information about the talker who produced the utterance and how it was produced (e.g., [11]). As such, dealing with indexical variability from a variety of sources may play an important role in normal, everyday speech communication.

Hearing-impaired listeners who have received cochlear implants (CIs) as a medical treatment for profound deafness, like normal-hearing (NH) listeners, must also be able to deal with multiple sources of speech variability. However, CI children and adults have to rely on episodic contexts that have less detailed acoustic-phonetic information than is typically available to NH listeners, due to the limitations of the CI device. Despite advances in our understanding of CI listeners' basic speech perception abilities, little is currently known about how CI listeners perceive and use detailed indexical information relevant to real-world communication, or what factors contribute to individual differences in their basic speech perception abilities. Several recent studies have suggested that CI listeners may not be able to make use of detailed talker-specific acoustic-phonetic information to make judgements about the talker (e.g., [3, 4]) and the talker's linguistic background (e.g. [5]) to the same extent as NH listeners. Further, CI listeners may have difficulty recognizing highly variable speech produced by non-native talkers [7]. Thus, although the CI device improves hearing, different sources of indexical variability may pose substantial problems to CI speech perception and spoken word recognition.

Foreign-accented speech is a common source of real-world variability. In many communicative environments, NH and CI listeners encounter non-native talkers whose speech is highly accented and unintelligible (e.g., [10]). NH listeners are sensitive to foreign accent variability and can use information about the talkers' accent to make reliable judgements about the talkers' language background and speech intelligibility (e.g. [6]). Further, NH listeners are able to make rapid perceptual adjustments and learn systematic accented variation in order to improve the recognition and understanding of foreign-accented speech [1, 2]. CI listeners may be less able to detect subtle phonetic regularities in foreign-accented speech compared to native listeners, because they make use of limited acoustic-phonetic speech information. As such, they may not be able to perceive and use foreign accent information in speech to make judgements about talkers, or recognize foreign-accented sentences or words. The current study was carried out to evaluate and characterize the perception of foreign-accented

speech by prelingually deaf, early-implanted, long-term CI users. CI and NH listeners' ability to evaluate the intelligibility of unfamiliar native and non-native talkers was assessed using an intelligibility rating task with native and foreign-accented speech. Based on previous studies, we expected that the CI listeners would have more difficulty than the NH listeners at perceiving and using the foreign accent information to rate the intelligibility of foreign-accented and native speech.

2. METHODS

2.1. Listeners

Two listener groups were recruited for this study. The CI group consisted of 44 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 listeners was 17.2 years, with a range of 9.3 to 30.0 years. The average age of CI implantation was 3.2 years, with a range of 0.7 to 6.3 years, and the average length of CI use was 14.0 years, with a range of 7.3 to 24.5 years. The CI listeners used their everyday CI map settings during testing. The NH group consisted of 47 children and adults with normal hearing. The mean age of the NH listeners was 17.8 years, with a range of 10.0 to 29.3 years. All listeners were part of an on-going study of long-term, cochlear implant outcomes and benefits [9].

2.2. Stimulus materials

Seventeen talkers (8 females and 9 males) were selected from the Multitalker Corpus of Foreign-Accented English (MCFAE) [12] for the current study. Nine talkers (4 females and 5 males) were non-native speakers of English with different native languages, including Japanese (N = 1), Kannada (N = 1), Korean (N = 1), Mandarin (N = 1), Portuguese (Brazil) (N = 1), Portuguese (Portugal) (N = 1), Spanish (Columbia) (N = 1), Taiwanese (N = 1), and Turkish (N = 1). The other 8 talkers (4 females and 4 males) 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 SPIN sentences [8] were used in the task, which consisted of 2 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 talkers, both non-native speakers of English, producing a unique high probability sentence. One of the selected male non-native talkers was only used for the practice trials. On the test trials, the participants listened to all the other 16 talkers (8 non-native and 8 native) producing a unique high or low probability sentence.

2.3. Procedure

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, located approximately 1m from them. All stimulus items were presented at 65 dB via the speaker. 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. The participants responded by touching a dialog box labelled with one of the response alternatives, i.e., the numbers 1-7, on the touchscreen monitor. The participants were not given any feedback, but they were allowed to ask questions about the task. Once they felt ready to continue, they began the test trials. For test trials, on each trial, participants were again presented with a sentence produced by a single talker, but they could only listen to that sentence one time. After listening to the single presentation of each utterance, participants could take as long as they wanted to respond, again by touching a labelled box with the response alternatives on the touchscreen monitor. Once they responded, the next trial began. Again, no feedback was given on the test trials.

3. RESULTS

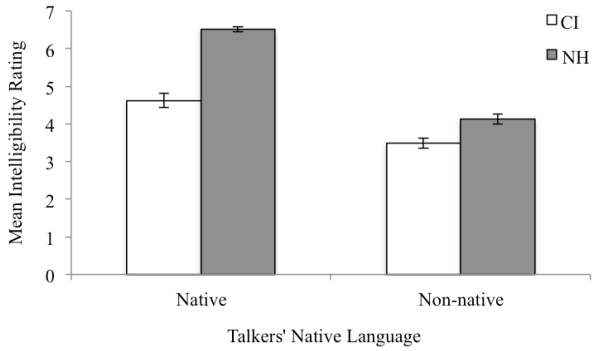
3.1. Group differences

Overall, both CI and NH listener groups indicated that the native talkers were on average more intelligible than the non-native talkers. Figure 1 displays the mean intelligibility ratings separately for the CI and NH groups. CI listeners gave an average rating of 4.6 (SD = 1.3) for native talkers and an average rating of 3.5 (SD = 0.9) for non-native talkers, while the NH listeners gave an average rating of 6.5 (SD = 0.5) for native talkers and an average rating of 4.1 (SD = 1.0) for non-native talkers.

A repeated-measures ANOVA on the intelligibility ratings with the talker's native language (native or non-native) as the within-subjects factor and listener group (CI or NH) as the between-subjects factor revealed two significant main effects of the talker's native language ($F(1,89)$ = 194.9, p < .001) and listener group ($F(1,89)$ =

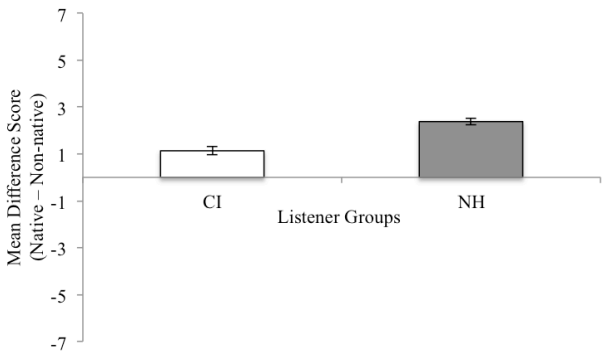
10.7, p = .002), and a significant talker language x listener group interaction ($F(1,89) = 22.0$, $p < .001$). Posthoc paired comparison t-tests on ratings for each listener group revealed that native talker ratings were higher than non-native talker ratings for both the CI ($t(44) = 6.5$, $p < .001$) and NH ($t(47) = 13.4$, $p < .001$) listener groups.

Figure 1: Mean intelligibility ratings for native and non-native talkers for CI (open bars) and NH (filled bars) listener groups. Error bars represent +/- 1 SE.



To further assess the listeners' perception of foreign-accented speech 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 non-native talkers from the average intelligibility rating for the native talkers (native – non-native). The NH listeners showed larger differences scores (M = 2.4, SD = 0.9) than the CI listeners (M = 1.1, SD = 1.1). A t-test confirmed that the difference scores were significantly different for the two listener groups ($t(89) = 4.7$, $p < .001$). Figure 2 displays the average difference score for the CI and NH listeners.

Figure 2: Mean intelligibility rating difference scores (native – non-native) for CI (open bar) and NH (filled bar) listener groups. Error bars represent +/- 1 SE.



3.2. Individual differences

Further examining the performance on the task, both CI and NH listeners displayed a great deal of variability in their sensitivity to the accent difference, as measured by the intelligibility rating difference scores. Difference scores for the CI listeners ranged from -1.4 to 2.9, while the difference scores for the NH listeners ranged from -0.1 to 4. To explore the sources of these individual differences, correlational analyses were carried out to compare the difference scores obtained from the intelligibility rating task to a small set of demographic variables and performance measures obtained from other speech perception tasks, which were analysed in the larger study [9]. Speech perception measures included sentence recognition scores using simple sentences produced by a single talker in quiet and in noise, anomalous sentences in quiet, sentences produced by multiple talkers in quiet, and foreign-accented sentences produced by multiple talkers in quiet. Scores from a nonword repetition task were also included. (Note that NH listeners did not complete all sentence recognition tasks.) The results of the correlational analyses between the intelligibility rating difference scores and other demographic and speech perception measures for CI and NH listeners are given in Table 1.

Table 1: Correlations between intelligibility rating difference scores and other demographic and speech perception measures for CI listeners (left) and NH listeners (right). *p < .05, **p < .01

Difference Scores	CI Listeners	NH Listeners
Age at testing	r = -.26	r = .04
Age at CI implantation	r = -.13	N/A
Length of CI use	r = -.21	N/A
Single Talker – Quiet	r = .65**	N/A
Single Talker – Noise	r = .69**	N/A
Anomalous – Quiet	r = .58**	r = .03
Multiple Talkers – Quiet	r = .62**	r = -.07
Foreign-Accented – Quiet	r = .58**	r = -.29
Nonword Repetition	r = .45**	r = .44**

For the CI group, none of the conventional demographic variables were related to the difference scores on the intelligibility rating task. However, difference scores were significantly related to all speech perception measures, which included tests of sentence recognition abilities across a variety of sentence types and presentation conditions as well as

nonword repetition abilities. For the NH group, only nonword repetition abilities were related to the intelligibility rating difference scores.

4. DISCUSSION

The goal of this study was to assess and characterize the perception of foreign-accented speech by a large group of prelingually deaf, early-implanted, long-term CI users. Overall, both CI and NH control listeners rated non-native talkers as less intelligible than native talkers, suggesting that they were sensitive to subtle differences in accents. However, the CI listeners rated both foreign-accented and native speech as less intelligible compared to the NH listeners. Given that CI listeners show poorer spoken word recognition scores even after long-term CI use (e.g., [9]), the present findings suggest that the CI listeners in the current study were aware of their own difficulties in recognizing spoken words, regardless of whether they were produced by a native or non-native speaker.

Examining the listeners' sensitivity to foreign-accented speech, the difference in the perceived intelligibility of the foreign-accented and native speech was smaller for the CI listeners than the NH listeners. CI listeners did not perceive the foreign-accented speech to be much less intelligible than the native speech, compared to the NH listeners. Thus, CI listeners were less sensitive to the accent information in the speech signal and were less able to use this information to make reliable intelligibility rating judgements. This suggests that foreign accent information, especially cues conveyed by fine spectral detail, may be poorly encoded and foreign-accented speech may be less robustly represented in long-term memory for the CI listeners.

However, not all CI listeners were equally sensitive to the accent information, and there was a great deal of listener variability in performance on the task. To explore individual differences, the CI and NH listeners' difference scores were compared to three demographic variables and several additional speech perception measures. The CI listeners' difference scores were related to their ability to accurately recognize spoken words in sentences, as measured by sentence recognition tests using a range of different sentence materials, and their ability to accurately perceive and repeat nonwords. The NH listeners' difference scores, however, were only related to their nonword repetition ability, likely because the NH listeners show very accurate sentence recognition scores in quiet. The results of these correlational analyses suggest that CI listeners' ability to perceive and use indexical variability in speech may rely heavily on basic speech perception abilities. The present results support the findings of previous studies demonstrating a close link between the perception of linguistic information and the perception of indexical information in speech (e.g., [3, 4]). Thus, it would be expected that the CI listeners in the current study who showed better speech perception abilities would also show greater sensitivity to other sources of indexical variability, e.g., talker gender or regional dialect, in similar or different types of speech perception tasks, e.g., categorization or discrimination. More research should be carried out to explore the relation between the perception of linguistic and indexical information in speech in this clinical population.

The current study reports some preliminary data on the perception of indexical variability in speech by CI users. Because the current study used a task that only indirectly examines the perception of foreign-accented speech, future studies should use tasks that more directly assess how the indexical information is perceived, encoded, and stored in long-term memory. In addition, other sources of indexical variability relevant to speech communication in real-world environments could be used, and compared. These studies may also include a more diverse sample of CI listeners, including CI users with different socioeconomic, educational, or language backgrounds. Further, the settings and characteristics of their CI devices should be taken into consideration when assessing individual listener performance on speech perception tasks (e.g. [13]).

In addition to addressing basic research questions, findings from the current study can contribute to the development of new clinical tools. Because the intelligibility rating task was related to the basic speech perception abilities of the individual CI listeners, it may be useful as a quick and easy tool for assessing CI speech perception and spoken word recognition, in combination with other, more standard tests. Additional research on CI perception of indexical variability may help to identify sources of strengths and weaknesses of individual CI listeners, and motivate the development of novel training and intervention programs to improve speech perception and recognition abilities in real-world listening conditions.

5. REFERENCES

[1] Bradlow, A.R, Bent, T. 2008. Perceptual adaptation to non-native speech. *Cognition*, 106, 707-729.

[2] Clarke, C. M., Garrett, M. R. 2004. Rapid adaptation to foreign-accented English. *Journal of the American Academy of Audiology*, 116,

3647-3658.

[3] Cleary, M. Pisoni, D. B. 2002. Talker discrimination by prelingually deaf children with cochlear implants: Preliminary results. *Annals of Otology, Rhinology & Laryngology*, 111, 113-118.

[4] Cleary, M., Pisoni, D. P., 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, 204-223.

[5] 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.

[6] Flege, J. E., Munroe, M. J., MacKay, I. R. A. 1995. Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97, 3125-3134.

[7] 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, 532-554.

[8] 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. *Journal of the Acoustical Society of America*, 61, 1337–1351.

[9] 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, 805–825.

[10] Munro, M. J., Derwing, T. M. 1995. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45, 73-97.

[11] Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In: Johnson, K., Mullennix, J. W. (eds.), *Talker Variability in Speech Processing.* San Diego, CA: Academic Press, 9-32.

[12] 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.

[13] van Wieringen, A., Wouters, J. 2014. What can we expect of normally-developing children implanted at a young age with respect to their auditory, linguistic and cognitive skills? *Hearing Research*.