Audio-vocal responses elicited in adult cochlear implant users

Torrey M. Loucks, Deepa Suneel, and Justin M. Aronoff

*Department of Speech and Hearing Science, University of Illinois, 901 South Sixth Street,
Champaign, Illinois 61820, USA*

tloucks@illinois.edu, dsuneel93@gmail.com, jaronoff@illinois.edu

Abstract: Auditory deprivation experienced prior to receiving a cochlear implant could compromise neural connections that allow for modulation of vocalization using auditory feedback. In this report, pitch-shift stimuli were presented to adult cochlear implant users to test whether compensatory motor changes in vocal F0 could be elicited. In five of six participants, rapid adjustments in vocal F0 were detected following the stimuli, which resemble the cortically mediated pitch-shift responses observed in typical hearing individuals. These findings suggest that cochlear implants can convey vocal F0 shifts to the auditory pathway that might benefit audio-vocal monitoring.

© 2015 Acoustical Society of America

[Q-JF]

Date Received: June 29, 2015 **Date Accepted:** October 1, 2015

1. Introduction

The voice is an intrinsic facet of our identity that marks our age, gender and even general health. Auditory feedback is used to maintain an intended vocal fundamental frequency (F0) in speech and singing,^{1,2} but also to adjust voicing rapidly as speaking conditions vary. Even though cochlear implants have transformed the potential for speech comprehension and production in children and adults, providing auditory feedback to perceive and produce the fine spectral changes associated with vocalization remains a challenge.³

The pitch-shift response paradigm (PSR) has confirmed that speakers with typical hearing adjust their vocal F0 rapidly during vowel production to compensate for unanticipated changes in perceived pitch.⁴ The PSR functions to help a speaker maintain an intended vocal F0 of a pitch target or particular intonation patterns. By briefly shifting vocal F0 up or down in frequency, non-volitional compensatory motor responses (opposite direction to shift) are elicited that begin approximately 100 ms after the stimulus and peak around 400 ms. Neuroimaging has demonstrated the PSR is a cortically mediated mechanism.⁵ Theoretically, the PSR reflects an internal model that operates to reduce the mismatch between an intended F0 and the perceived F0.⁴ Our goal was to test whether audio-vocal responses with characteristics of the PSR can be elicited in adult cochlear implant users. If audio-vocal responses are present, it may be possible to use production measures to gain insight into CI users' perception of vocal cues and the relative benefit of auditory feedback.

Post-lingually deafened individuals frequently shift into an inefficient high F0 range, employ overly loud speech, and do not vary voicing appropriately due to an absence of auditory feedback.² While some voice characteristics improve after implantation,⁶ likely due to partial restoration of auditory feedback, many CI users continue to show less than optimal voice perception and production. In any case, rapid audio-vocal manipulations, such as the PSR, may not occur in CI users for several reasons. First, the F0 range of the adult voice (~100–250 Hz) falls below the frequency range of cochlear implant processors. Second, most implant processors do not convey brief frequency changes to the auditory nerve faithfully, which could degrade the perception of spectral cues. Third, and most important, it is possible that auditory deprivation experienced by CI users prior to implantation degrades or compromises the neural connections that underlie rapid audio-vocal responses.

The technical limitations of conveying rapid spectral transitions of the voice might be mitigated, however. Alterations in F0 usually also alter a large range of frequencies, such as the natural harmonics of the voice, so a shift in vocal pitch that includes the harmonics could be encoded by processors. Additionally, limitations in transmitting spectral information can be reduced using strategies such as current steering⁷ that more faithfully transduce spectral changes. Without current steering, spectral changes within a frequency band do not alter stimulation while spectral changes that

cross frequency bands can result in a disproportionately large shift in stimulation. Current steering manipulates stimulation in two electrodes simultaneously to create an electrical peak between the electrodes and thus might transmit spectral changes more faithfully to the auditory pathway.⁷

This study tested whether rapid audio-vocal responses can be elicited in adult CI participants when listening with a map that uses current steering. We focused on adult participants who experienced a long period of hearing loss and deafness, which could have degraded auditory-to-vocal feedback. If audio-vocal responses are present, it suggests this feedback mechanism can be preserved despite a period of auditory deprivation.

2. Methods

Six adults with bilateral cochlear implants were recruited for this study (three female). As per Table 1, five out of six participants were post-lingually deaf with varying durations of deafness and implant use. The other participant experienced hearing loss in a peri-lingual stage. All participants used spoken English as their primary language and regularly used their cochlear implants throughout the day and were able to perceive open set speech in quiet conditions. PSR data were also collected from a comparison group of six typical hearing English speaking adults in the same age range. Informed consent was obtained from each participant. The study was approved by the Institutional Review Board at the University of Illinois.

Only participants with bilateral Advanced Bionics implants were selected because of the availability of current steering. All CI participants were tested unilaterally with a map with a HiRes 120 strategy, a pulse rate of approximately 967 pps, a phase duration of approximately 32 μ s, and a frequency range of 306 to 8054 Hz. M and T levels were set individually for each participant, with T levels set at 10% of M levels.

The pitch shifts were delivered online through a TC-Helicon multi-effects processor as participants vocalized the vowel /a/. Vocalizations were recorded using a headworn microphone, routed through the multi-effects processor, and delivered to the CI participant unilaterally via the processors' auxiliary input port to the participant's preferred ear. The selection of the preferred ear was based on which ear the patient felt was their stronger ear. During testing the patient only wore a processor on the preferred ear. This greatly reduced the signal available to the non-preferred ear, restricting it to only what could be heard with their minimal residual hearing. Participants produced 60 vocalizations while three pitch shift stimuli were delivered per vocalization trial. The amplitude of the pitch-shift stimulus was 200 cents and lasted 200 ms. The direction of the pitch-shift stimulus varied randomly.⁸ The setup for typical hearing participants was identical except the feedback signal was delivered bilaterally through circumaural headphones. For all participants, perceived loudness was adjusted to a loud but comfortable level. Figure 1(a) shows spectra from 100 ms samples before and during an upward shift to illustrate how the entire voice spectrum is shifted.

The vocal acoustic signals were sampled at 4 kHz using an anti-aliasing filter. Potential pitch shift responses were extracted beginning 200 ms before the shift stimulus until 800 ms post-stimulus. In order for a trial to be included, the change in pitch had to begin at least 80 ms after stimulus onset, exceed 15 cents in amplitude and peak before 700 ms. Compensatory responses showed a primary amplitude response in the

Table 1. Characteristics of cochlear implant participants.

Subject	Age	Gender	Hearing loss onset	Onset of profound hearing loss	Cause	Implant experience
I02	60	Female	2 years old	3 years old	Meningitis	2 years (L) ^a 5 years (R)
I05	69	Male	5 years old	14 years old	Unknown (Injury or Genetic)	12 years (L) ^a 12 years (R)
I06	56	Female	36 years old	47 years old	Genetic - Maternal	1 year (L) 3 years (R) ^a
I07	53	Male	30 years old	51 years old	Familial	6 months (L) 6 months (R) ^a
I10	49	Female	29 years old	46 years old	Autoimmune	1 year (L) ^a 2 years (R)
I11	68	Male	10 years old (L) 58 years old (R)	10 years old (L) 58 years old (R)	Sudden	2 years (L) 10 years (R) ^a

^aEar used for experiment.

Fig. 1. (a) Spectra of a woman's voice before and during the pitch-shift stimulus. (b) A representative waveform showing the averaged upward compensatory vocal response (cents) of a CI participant to a negative pitch shift stimulus (−200 cents stimulus, mean peak amplitude: 38 cents; mean peak time: 342 ms).

opposite direction of the shift while following responses occurred in the same direction as the shift. Compensatory and following responses were averaged separately according to the direction of the response (absolute values determined for all conditions). The average peak amplitude, average peak time and proportion of responses relative to the total number of stimuli are reported for each category (Up Compensatory, Down Compensatory, Up Follow, and Down Follow).

3. Results

We classified pitch-shift responses according to the criteria described above. Figure 1(b) displays an example of an averaged upward compensatory PSR from a CI participant that closely resembles published PSR waveforms (stimulus: −200 cents). We confirmed through questioning that the CI participants could detect the pitch shift stimulus. Later in post-testing, we also confirmed that CI participants could detect the audibility of pitch changes by shifting the pitch up or down for longer periods of time. When the PSR analysis criteria were applied to the data of the CI participant I11, who did not hear the shifts, we found that most trials did not meet our criteria for valid responses because of highly damped responses and ambiguous waveforms relative to the other participants. As a result, his data were not included in the averaged results. Figure 2(a) shows means of the PSR amplitude, PSR peak time and proportion of responses from

Fig. 2. Average amplitude, average peak time and proportion of responses in the CI group and the reference sample of typical hearing participants: Up Compensatory, Down Compensatory, Up Follow, and Down Follow. (a) Cochlear implant participants. (b) Typical hearing participants. (Error bars indicate 1 standard deviation.)

the five participants. The data from six typical hearing individuals that were analyzed in the same manner as the CI participants are shown in Fig. 2(b). It is clear from a descriptive comparison that the magnitude and timing of the compensatory responses show similar patterns in CI and typical hearing participants, although the amplitude of the CI participants' responses is larger. A repeated measures analysis of variance indicated that the amplitudes of responses in CI participants were significantly greater [$F(1,9)=7.75$, $p=0.021$] and particularly the following responses [Group by Response Type— $F(1,9)=14.50$, $p=0.004$]. Across both groups, compensatory responses were greater than following responses [$F(1,9)=15.82$, $p=0.003$]. The increased amplitude in the CI group denote individual differences that are mostly attributable to three participants, who had a higher range of amplitude values than the typical hearing group. The typical hearing listeners showed individual differences as well but their range of responses was reduced relative to the CI group and consistent with our previous studies.⁸

An equivalent comparison of peak time showed that the only difference was an interaction [Group by Response Type— $F(1,9)=5.31$, $p=0.047$] in which the following responses of the reference sample were relatively earlier. Typical hearing participants clearly tended towards having a majority of compensatory responses, whereas the CI participants showed more variation, although with a mean preference for the compensatory direction. Correlations between the dependent variables and the duration of hearing loss, duration of deafness, onset age of hearing loss and length of CI use were small and nonsignificant.

An unexpected finding suggests some cochlear implant users rapidly adjust their vocal F0 based on auditory input in the absence of the experimental stimulus. Two CI participants displayed rapid, large increases in F0 immediately following phonation onset but before the first pitch shift stimulus was presented. One of these participants (IO2) heard the pitch shifts while the other was I11 who did not perceive the stimuli. As per Fig. 3, the F0 from participant I11 rose rapidly at onset and peaked between 500 to 700 ms after phonation onset whereas F0 remained stable after that point or declined gradually. The amplitude of these initial adjustments across both participants had a combined average of a 20 Hz increase following the onset of vocalization (standard deviation = 10 Hz).

4. Discussion

In this study, we tested whether cochlear implant users are able to produce rapid audio-vocal responses to unanticipated pitch shifts despite the alterations of the spectral signal by the processor and implant and periods of partial or complete auditory deprivation. We presented unilateral pitch-shifts to six CI participants and found that five of six participants perceived and responded to the stimuli in a manner that resembles the rapid audio-vocal responses in typical hearing participants. It appears that CI participants produced greater response amplitudes and less contrast between compensatory and following responses but had similar timing as the typical hearing reference

Fig. 3. Representative traces of F0 during vocalization in a CI participant who showed rapid and large F0 increases starting at the onset of vocalization.

sample. Variation in PSR parameters and proportion of following responses in typical hearing have been noted previously,^{4,8} so a range of differences in these variables is not unexpected.

The CI participants resembled the typical hearing group in showing a general preference for more compensatory responses than following responses. Also, the close similarity in peak time across the groups suggest the time course of audio-vocal responses in CI users can be as rapid as typical hearing individuals. Based on clear changes in pitch that generally tended to be compensatory and occurred over the same time window as typical hearing participants, it is reasonable to infer that rapid audio-vocal responses akin to the PSR can be elicited in adult CI users.

As each of these adult participants experienced some typical acoustic hearing during the critical language development period, they had some previous familiarity with auditory feedback for vocal control. It appears that short-latency audio-vocal control remained functional in most of these CI users despite a period of auditory deprivation, although some form of neuroimaging or evoked potential data are eventually needed to support this inference. It was not clear *a priori* what CI listener's responses would be to pitch shifts, given the alterations of the signal by the processors and the atypical nature of their hearing experience, but their responses suggest sensitivity to pitch variations. It is also unclear whether audio-vocal responses would be found in CI users who experienced pre-lingual deafness, but it is a relevant question for future research.

The variation across individuals in response proportion suggests a bias in directionality of responses. It is not clear if this is due to individual variation in how the stimulus is perceived or an atypical mode of vocal control. Varying the stimulus parameters and vocalization instructions along with controlling participant characteristics could help in understanding how different response patterns emerged. The unilateral or monaural testing of the CI users may have contributed to some differences in their responses but time limitations did not permit a comparison of binaural vs monaural responses, yet this is a goal for our continuing study of bilateral cochlear implant users. We focused on Advanced Bionics processors with current steering activated as a likely candidate for capturing spectral changes in this initial study. Given this evidence for rapid audio-vocal responses in CI users, further research with and without current steering, different processing strategies and perceptual discrimination of spectral shifts is warranted.

The unexpected but prominent pitch increases seen at vocalization onset in two CI participants are interesting because it suggests some form of phonatory feedback was associated with these pitch changes. These changes were observed at the beginning of the experiment so they did not arise from experience with the pitch-shift stimulus. Albeit speculative, there are several possibilities. The prolonged vocalization may have allowed the participant to detect a mismatch between their intended vocalization frequency and the modified form delivered to their auditory system from the implant. This mismatch drove a change in vocal frequency resembling the general internal model interpretation of typical pitch shift responses. Alternately, the individual may habitually seek additional sensory feedback that required a higher phonation frequency. This might involve a combination of laryngeal somatosensation and auditory sensation that was only detectable when pushing the voice outside a typical F0 range. This is a less than optimal strategy that places stress on the vocal folds over the long term. A better understanding of this phenomenon is clearly required, but it generally offers further evidence of rapid audio-vocal responses in CI users, but this type of response would not be expected in typical hearing listeners.

There are numerous directions for further research based on our findings that the cortically mediated PSR might be preserved despite auditory deprivation. It contributes to our understanding of whether a CI is helpful in restoring functional feedback for vocal control and how adult CI users process pitch alterations. The PSR findings may contribute to the development of production based assessments of auditory feedback. As one example, an audio-vocal feedback test could improve CI fitting cases where it is difficult to obtain reliable subjective feedback, such as with pediatric populations. It could also provide a means to create maps that reduce the mismatch between the production and perception of a CI user's own voice. Finally, there is considerable interest in processing strategies that can convey the vocal F0 changes of lexical tones in order to promote tone contrasts in CI users of tonal languages.⁹ If pitch-shift responses can be elicited in pediatric CI populations, it provides a new means for assessing processing strategies designed to deliver F0 contours that could promote production of these linguistic contrasts.

Acknowledgments

We thank our participants for their time and effort. We also thank Advanced Bionics for providing equipment for this study. This work was supported by NIH Grant No. R03-DC013380.

References and links

- ¹M. S. Smotherman, "Sensory feedback control of mammalian vocalizations," *Behav. Brain. Res.* **182**(2), 315–326 (2007).
- ²R. Cowie, E. Douglas-Cowie, and A. G. Kerr, "A study of speech deterioration in post-lingually deafened adults," *J. Laryngol. Otol.* **96**(2), 101–112 (1982).
- ³S. C. Peng, J. B. Tomblin, and C. W. Turner, "Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing," *Ear Hear.* **29**(3), 336–351 (2008).
- ⁴T. A. Burnett, M. B. Freedland, C. R. Larson, and T. C. Hain, "Voice F0 responses to manipulations in pitch feedback," *J. Acoust. Soc. Am.* **103**(6), 3153–3161 (1998).
- ⁵N. S. Kort, S. S. Nagarajan, and J. F. Houde, "A bilateral cortical network responds to pitch perturbations in speech feedback," *Neuroimage* **86**, 525–535 (2014).
- ⁶M. T. Ubrig, M. V. Goffi-Gomez, R. Weber, M. H. Menezes, N. K. Nemr, D. H. Tsuji, and R. K. Tsuji, "Voice analysis of postlingually deaf adults pre- and postcochlear implantation," *J. Voice* **25**(6), 692–699 (2011).
- ⁷X. Luo, D. M. Landsberger, M. Padilla, and A. G. Srinivasan, "Encoding pitch contours using current steering," *J. Acoust. Soc. Am.* **128**(3), 1215–1223 (2010).
- ⁸B. A. Sturgeon, R. J. Hubbard, S. A. Schmidt, and T. M. Loucks, "High F0 and musicianship make a difference: Pitch-shift responses across the vocal range," *J. Phon.* **51**, 70–81 (2015).
- ⁹S. C. Peng, J. B. Tomblin, H. Cheung, Y. S. Lin, and L. S. Wang, "Perception and production of Mandarin tones in prelingually deaf children with cochlear implants," *Ear Hear.* **25**(3), 251–264 (2004).