Premiere Publications from The Triiological Society

Read all three of our prestigious publications, each offering high-quality content to keep you informed with the latest developments in the field.

**Laryngoscope**

FOUNDED IN 1896

Editor-in-Chief: Michael G. Stewart, MD, MPH

The leading source for information in head and neck disorders.

[Laryngoscope.com](http://Laryngoscope.com)

**Laryngoscope Investigative Otolaryngology**

Editor-in-Chief: D. Bradley Welling, MD, PhD, FACS

Rapid dissemination of the science and practice of otolaryngology-head and neck surgery.

[InvestigativeOto.com](http://InvestigativeOto.com)

**ENTtoday**

A publication of the Triological Society

Editor-in-Chief: Alexander Chiu, MD

Must-have timely information that Otolaryngologist-head and neck surgeons can use in daily practice.

[Enttoday.org](http://Enttoday.org)

WILEY
Control of Speech and Voice in Cochlear Implant Patients

Anirudh Gautam, BSc; James G. Naples, MD; Steven J. Eliades, MD, PhD

Objective: Hearing plays an important role in the learning and production of speech, but the benefits of cochlear implantation for such vocal control are unclear. Here, we present a perspective and review of recent work on the control of speech and voice following cochlear implantation. We further discuss insights provided on the mechanisms of normal vocal control and implications for future rehabilitative approaches.

Data Sources: Peer-reviewed articles on speech and voice production in cochlear implant patients were identified from PubMed. Relevant articles were supplemented with selected publications describing normal vocal control mechanisms and behaviors.

Review Methods: Publications that discussed speech and voice outcomes following cochlear implantation were chosen, with a focus on those presenting measurements of specific speech or voice parameters.

Results: Recent studies demonstrate that hearing restoration by cochlear implantation has significant effects on many aspects of voice and speech production. These include changes in vocal pitch and loudness, as well as improved control of both vowels and consonants. Despite these improvements, however, the speech of many implant recipients remains abnormal as compared to normal hearing individuals. Such differences likely result from the impoverished auditory feedback provided by the implant.

Conclusions: Cochlear implants provide valuable insights into the role of hearing in vocal production. Although implants improve vocal production for most patients, there remains considerable room for future study and therapeutic improvement.

Key Words: Cochlear implant, speech production, voice, hearing loss, vocal control.

Laryngoscope, 129:2158–2163, 2019

INTRODUCTION

Cochlear implantation has been a major advance in the rehabilitation of patients with hearing loss that has restored sensory perception in tens of thousands of recipients. Although cochlear implants (CIs) are intended to improve the global communication of recipients, assessment of implant outcomes has largely been focused on auditory perceptual abilities. Vocal communication, however, requires accurate speech production in addition to perception. Unfortunately, research into the speech productive abilities of CI recipients has lagged behind measurements of perceptual performance. Recently, there has been increasing interest in otolaryngology into the mechanisms of vocal production and control, and in this review we summarize recent work into the role of CIs. Understanding productive abilities and limitations of CI recipients will hopefully allow further improvements in communication abilities for patients with hearing loss.

Hearing, Deafness, and Vocal Production

It is well appreciated that hearing plays an important role in vocal production. Patients with congenital deafness have difficulty acquiring and maintaining normal speech.2 These children exhibit errors in voicing, vowel substitutions, increased speech duration, and difficulties in intonation and voice quality, to name a few.3 As a result, such speech is often less intelligible than that of normal hearing (NH) individuals. Even those with hearing loss later in life exhibit mild degradation in their speech over time, often producing speech with greater pitch variability and abnormal articulation, although overall intelligibility of their speech is not greatly deteriorated.5

Normal hearing individuals, on the other hand, have been shown to exhibit robust control of speech and voice. When faced with errors or alterations in auditory feedback, normal listeners will adjust their vocal production to compensate. This includes the classic Lombard effect, wherein speakers increase vocal loudness in the presence of background noise.6 Other evidence for feedback control of speech is found in the control of vocal pitch,7 vowel formants,8 and timing.9 These studies provide evidence that speakers can use hearing on a moment-to-moment basis to help control speech and voice, although the underlying mechanisms remain uncertain.10 Current models implicate the auditory system as part of a network of brain regions that evaluate auditory feedback while speaking and relay this information to speech
motor areas (Fig. 1). Despite the importance of vocal feedback control for NH listeners and its dysfunction in hearing loss, to what extent CIs can restore this feedback control and support normal speech is unclear. Restoring hearing through a CI should in theory provide auditory feedback of one’s own voice, thereby improving vocal control.

**Levels of Vocal Control: Voice Versus Speech**

It is important to make a distinction between related but often confused elements of vocal communication, including voice, speech, intelligibility, and language. Intelligibility is a subjective measure of the perceived comprehensibility of speech that is dependent on multiple acoustic features. Although speech intelligibility is decreased in deafness and improved by CIs, it does not yield much insight into underlying mechanisms. Language, broadly, is the information content of speech, rather than the signal itself, and has been extensively studied in CIs. Here we instead focus on control of voice and speech, which in a simplified fashion can be thought to consist of vocal production by laryngeal air flow (voicing) subsequently modified by pharyngeal and oral articulation. Past research into vocal control often divides measurements into two broad categories, reflecting distinct control mechanisms. The first set of parameters are so-called suprasegmental or postural measures that are continuous from one phoneme or word to the next, consisting of loudness, pitch, and duration, and best corresponding to parameters of voice. Such suprasegmental parameters are under tight feedback control in NH individuals, with short latency reflexes, as in control of pitch. In contrast, parameters that comprise the segmental measures are those that vary from utterance to utterance, including vowel formants and consonants, more closely related to speech and phonemes. These distinctions are not always simple because formant and fine structure information can also be used in the perception of voice pitch/prosody, particularly in CI patients. Control of such segmental parameters is less clear because they are transient speech events, and feedback processing is often insufficiently fast to correct errors. Many segmental parameters exhibit feedback control that is more gradual and learned, as in the control of vowel formants. The mechanisms that allow NH speakers

![Fig. 1. Interactions between the auditory pathway and vocal motor system to support vocal self-monitoring and feedback control of voice and speech. Green arrows illustrate ascending auditory information flow, blue arrows the descending vocal motor pathway, and orange the reciprocal connections between cortical areas thought to be involved in vocal feedback control.](https://www.laryngoscope.com)
to control speech are poorly understood, in part because of limitations in the degree of experimental manipulation that can be performed. Because of the ability to completely control feedback, CI studies have potential to yield unique insights into these vocal processes.

**Cochlear Implants and Voice**

Studies of voice parameters generally focus on measurements of vocal pitch/fundamental frequency (F0) and amplitude/loudness (sound pressure level, SPL), as well as vocal stability, termed jitter (pitch variability), and shimmer (loudness variability). Patients with hearing loss generally speak louder and with higher pitch than NH individuals, presumably attempting to improve self-perception of their voice.17,18 Multiple studies have examined effects of CI on such vocal control, both over short and long time-scales (Table I). For example, immediately following CI activation, patients generally exhibit decreased vocal amplitude in both adults17,19–25 and children.24 Such control operates on very short time-scales, with loudness changes being detectable even when simply turning the implant on and off.17,25–27 Further evidence for loudness control is found in CI patients’ increased vocal amplitude in the presence of background noise.28 the previously described Lombard effect.

Evidence for the short-term effects of implantation on vocal fundamental frequency has been mixed. Similar to loudness, some adult users exhibit elevated F0 when the implant is turned off and subsequent decrease in pitch when switched on.14,17,25–27 Other evidence for vocal pitch control in CI users is found in studies showing decreased F0 in the presence of delayed auditory feedback, a reflexive behavior seen in NH individuals.29,30 However, in contrast to studies showing changes in voice with acute changes in hearing, other studies have found more heterogeneous results. Some adults did not reduce F0 or SPL, whereas others actually increased these parameters.14,17,26,27,31 Such individual variability in pitch and loudness is primarily evident with sudden changes in auditory feedback (i.e., turning the implant on/off), and may suggest that, for some, vocal control was not dependent on moment-to-moment feedback, instead operating over longer time scales. Alternatively, it has been suggested that some individuals may actively suppress their voice to avoid overly loud speech.14,17,27

Although evidence for control of voice over short-time scales is somewhat mixed, there is more robust evidence for control of vocal pitch and loudness after longer-term CI use, spanning months to years. These studies show consistent reductions in both F0 and SPL following implantation of adults with postlingual hearing loss.17,20,22,25,27 Surprisingly, improvements have also been seen in patients with prelingual deafness who did not receive a CI until adulthood,33 patients who generally have poorer CI outcomes. Similar results have been seen in pediatric recipients, regardless of whether hearing loss was pre- or postlingual.24 Changes in vocal pitch and loudness can begin as early as 2 weeks after implant activation and continue to improve for at least 6 months afterward.19 Voice improvements were not as prominent in early studies of single-channel CIs,32 consistent with multichannel implants providing greater auditory information more effectively utilized to control voice.

In addition to overall improvements in vocal pitch and loudness, CIs also affect the ability to maintain a stable voice. Hearing impaired individuals exhibit less stability in their vocal production compared to NH, presumably due to loss of feedback that allows detection and correction of errors. This reduced control is most obvious in vocal instability between successive words, with increases in variability of pitch (vF0) and loudness (vAM). Following

---

**TABLE I.**

**Summary of Speech and Voice Changes in Deafness and the Effects of CI.**

<table>
<thead>
<tr>
<th></th>
<th>Effects of Deafness</th>
<th>After CI</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Voice: pitch (F0)</strong></td>
<td>Increased</td>
<td>Decreased*</td>
<td>Adult: 17,22–23,25–27, Pediatric: 24</td>
</tr>
<tr>
<td><strong>Loudness (SPL)</strong></td>
<td>Increased</td>
<td>Decreased*</td>
<td>Adult: 14,17–23,25–27,32, Pediatric: 24</td>
</tr>
<tr>
<td><strong>Jitter (and F0 variation)</strong></td>
<td>Increased</td>
<td>Variable decreases*</td>
<td>Adult: 19,34, Pediatric: 34–36</td>
</tr>
<tr>
<td><strong>Shimmer (and SPL variation)</strong></td>
<td>Increased</td>
<td>Variable decreases*</td>
<td>Adult: 23,44–47, Pediatric: 48–51</td>
</tr>
<tr>
<td><strong>Vowels: formant separation</strong></td>
<td>Reduced</td>
<td>Slightly increased separation*</td>
<td>Adult: 17–18,23, Pediatric: 36,48,51</td>
</tr>
<tr>
<td><strong>Individual formants</strong></td>
<td>Decreased F1, Variable F2</td>
<td>Increased contrasts*</td>
<td>Adult: 56–58</td>
</tr>
<tr>
<td><strong>Consonants: contrasts</strong></td>
<td>Poor /s/ vs. /ʃ/</td>
<td>Variable increases*</td>
<td>Adult: 18,26,47,52,59</td>
</tr>
<tr>
<td><strong>Voice onset time</strong></td>
<td>Poor /r/ vs. /ɹ/</td>
<td>Reduced</td>
<td>Variable increases*</td>
</tr>
<tr>
<td><strong>Speech timing: duration</strong></td>
<td>Increased</td>
<td>Decreased*</td>
<td>Adult: 19, Pediatric: 24</td>
</tr>
<tr>
<td><strong>Rate</strong></td>
<td>Slower</td>
<td>Variable changes*</td>
<td>Adult: 23,44–47, Pediatric: 48–51</td>
</tr>
</tbody>
</table>

*Change relative pre-CI or CI turned off.
†Results closer to NH listeners in nonlongitudinal studies.
CI = cochlear implant; F0 = fundamental frequency; NH = normal hearing; SPL = sound pressure level.
implantation, adult vF0 has been found to decrease. On the other hand, CIs have only been found to improve vAM in children. Common short-term measurements of vocal stability—jitter and shimmer—which measure variation of pitch and loudness, respectively, within single utterances, have similarly been found to increase in hearing impaired individuals, which is consistent with decreased vocal stability. Following CI, both jitter and shimmer decrease in some prelingually deaf children, approaching normal ranges. Other studies have suggested that both jitter and shimmer in deaf children do not significantly differ from NH either before or after implantation. However, studies have been limited to a small number of children, and the result variability between them may reflect this. Nonetheless, overall results support CIs providing sufficient feedback to improve both short- and long-term voice stability.

In addition to their benefit in improving vocal stability during speech and vowel production, recent studies also show that implants may improve the accuracy in producing specific vocal pitches. Prelingually deaf children demonstrate improved pitch matching while singing but still fail to match the performance of their NH counterparts. In contrast, children still exhibit diminished abilities to produce pitch contours associated with semantic content, such as increased pitch at the end of a question. In tonal languages such as Mandarin, most children with CIs also exhibit difficulty in the production of appropriate speech pitch contours. Overall, CIs appear to improve vocal production by reducing vocal pitch and loudness and by reducing variability. These results suggest that CIs provide adequate auditory feedback during vocal production to allow stabilization and correction of voice, although patients often still do not match the performance of NH individuals. Such results suggesting improved pitch control are surprising, given that most implant electrodes do not extend into vocal pitch frequencies and that implant patients exhibit poorer performance in pitch detection and musical appreciation. On the other hand, pitch perception is not purely dependent on F0 and can be extracted from timing cues or higher harmonic frequencies within the CI range. It is possible that subjects use these harmonics instead of F0 alone. It is also unclear to what extent vocal control uses moment-to-moment feedback because sudden changes in the implant (on/off) have mixed outcomes. Alternatively, subjects might gradually learn to use averaged feedback over longer periods. Future investigations may be better able to disambiguate the two and may have implications for speech rehabilitation.

**Cochlear Implants and Vowel Formants**

A second major component of speech is the production of vowels by modification of voiced sounds using the upper airway to produce distinct formant frequency peaks. In NH speakers, each vowel has a characteristic set of formants, measured by the position of the first and second (F1 and F2) formant frequencies. However, in deaf speakers, there typically is a reduction in the difference between vowel formants, resulting in sounds that are acoustically and perceptually less distinct from one another. Such reductions result in a contracted vowel space in which multiple vowels are closely spaced and overlapping in frequency rather than spread out and distinct (Table I).

Several studies have examined the effects of CI on vowel formants. Following implantation in adults, there is a gradual expansion of the formant space that results in improved contrasts between vowels. Similar results are present in implanted prelingually deaf children. These changes not only lead to more distinct vowel contrasts but also result in vowels that are more similar to those expected from age-matched peers. Interestingly, vowel changes appear to more specifically involve F1 decreases, with F2 changes being more variable in adults and children. This is surprising given that CIs provide better access to high-frequency sound components. Improvements in vowels can be seen as early as 2 weeks, although other studies have not found effects until 2 months, in contrast to the more immediate changes noted above for vocal pitch.

Maintenance of these vowel changes appears to require continuous CI use. When an implant is turned off, there is a re-contraction of the vowel space back toward less distinct formants than during implant use. This effect is strongest early after implantation, with evidence for contraction of the vowel space when implants are briefly turned off at 1 month postactivation. However, there is increased speech stability with increasing experience, with persistent vocal improvements even during brief hearing interruptions, suggesting that there may be some learned or plastic changes in speech as a result of CI use. Implant failures also result in similar formant changes.

Despite these strong trends, not all CI recipients exhibit improvements. A few reports have documented adults whose vowel spaces actually contracted after implantation. Other studies failed to find improvements in postlingually deaf adults with relatively normal speech prior to implantation, possibly due to a higher performing baseline. Even for children who exhibit improvements, they are often still unable to match NH individuals. The origin of this considerable variability remains uncertain but may reflect variability in preimplant hearing experience and speech quality, whether speech developed prior to implantation, or variability in implant electrode placement and function.

Overall, CIs appear to result in substantial improvements in vowel production for most recipients, with expanded contrasts between formants resulting in greater acoustic distinctions. The effects of CI on formants appear to be more robust and consistent than that observed for voicing parameters, perhaps due to the higher formant frequencies that are more likely to overlap frequency programming of the CI speech processor, or possibly the role of slower and more stable learning and plasticity in vowel formants, as has been observed in NH individuals.

**Cochlear Implants and Consonants**

Although there has been significant research looking at the effects of CI on vowels, there has been relatively less research examining effects on consonant production.
Cochlear implant use has been found to improve production and contrasts between sibilants /s/ and /ʃ/, particularly in adults with poor preimplantation speech.56 Although these changes can take 6 months or more to develop, other studies have found decreased variability in the production of consonant /r/57 and improved contrasts between /l/ and /ɾ/58, although such improvements can, again, take a year or more.

Another common measure of consonant production is voice-onset time (VOT), the duration between the start of a consonant and subsequent voiced vowel (i.e., /da/). In hearing impaired listeners, VOT is generally shorter than for NH individuals.59 Following CI, adults show a correction of the VOT toward normal values.26,59 This correction appears rapidly, occurring as early as the first day of activation. Similarly, turning the implant off leads to rapid changes, with shortened VOTs within 24 hours that correct immediately after the implant is turned back on. Other studies, however, have found more variable VOT results for some subjects.47,52 Overall, these results suggest rapid changes in VOT with hearing restoration.

### Cochlear Implants and Speech Rate or Duration

In contrast to somewhat mixed results in other voice and speech parameters, CIs appear to provide better control in the temporal patterns of speech. Individuals with hearing loss exhibit speech with longer syllable and word durations60 and slower speaking rates19 than NH individuals. Following adult CIs, there is a reduction in the duration of individual words or vowels.17,18,20,32,46 These effects can be seen as early as 1 day after implant activation22 but have not been seen in measurements immediately upon activation.27 For experienced users, the effects of feedback on duration appear to be rapid; turning an implant off results in longer word durations within one to two utterances, and turning it back on results in improvements over similar timescales.16,25,31 Children with CIs exhibit similar vocal durations to NH peers during choral singing.29 These results suggest that use of CI feedback to control speech duration is a rapid and robust phenomenon. Strong changes in timing are not surprising given the relative accuracy of temporal cues provided by CIs compared with much coarser frequency information.

Studies looking at the effects of CIs on speaking rate have not found any consistent improvements following implantation in adults19 or children.24 This absent effect has been hypothesized to result from increased efforts to articulate more precisely with an implant, resulting in slower speech. Despite the absence of long-term speaking rate effects, there is some evidence that auditory feedback can affect speaking rate over shorter periods. Delays in auditory feedback, known to cause slower speech and speech errors in NH listeners, also result in slower speaking rates in CI recipients.29

### Predictors of Speech Outcomes

Although it is clear that CIs improve the production of speech and voice in many recipients, there remains considerable variability in performance, and many factors may influence speech control. For example, in children the age of implantation plays a significant role both in perceptual performance61 and accuracy of speech production.51,53 The greatest improvements appear for those implanted before 4 years,62,63 with those implanted before 24 to 30 months producing speech similar to age-matched peers.64–66 Other factors that affect speech quality in children also arise from the learning environment, with those immersed in total auditory/oral communication showing improved voice and speech compared with those who continue to use signficant sign language.19,67 Likely reflecting increased daily practice with oral communication. Perhaps the best predictor of vocal production skills, however, is auditory perceptual abilities. Not surprisingly, pediatric production accuracy correlates with perceptual performance for consonants68; vowel formants69; and vocal pitch, jitter, and shimmer.70 Such results suggest that the ability to accurately control speech and voice after CI is dependent on phonologic awareness of these aspects of speech. Other possible contributions, including programming strategies, are intriguing possibilities but not consistently documented.

### CONCLUSION

Auditory feedback has an important role in vocal communication, with many deaf and hearing loss patients exhibiting abnormal patterns of speech and voice. Unfortunately, although we have good therapeutic options to improve hearing loss using CIs, the effects of restoration on vocal production have not been well studied. Recent work has shown that, although CIs provide sufficient auditory input to support the acquisition and control of speech, much of this speech remains atypical. To what extent speech and voice patterns reflect impoverished auditory information or abnormally learned speech motor patterns remains an open question. Many aspects of speech and voice require constant, on-line feedback, and likely reflect the quality and precision of the information provided by the implant, whereas others appear to exhibit more robust stability over time. It is our hope that future work will better distinguish how different aspects of speech are specifically affected by cochlear implantation, and that this can lead to better outcomes through new processing or rehabilitation strategies.

### BIBLIOGRAPHY
