Published in final edited form as:

Hear Res. 2018 September ; 367: 223–230. doi:10.1016/j.heares.2018.06.016.

Electric and acoustic harmonic integration predicts speech-in-noise performance in hybrid cochlear implant users

Damien Bonnard^{#a,2}, Adam Schwalje^{#a}, Bruce Gantz^a, and Inyong Choi^{a,b,*}

^aDepartment of Otolaryngology — Head and Neck Surgery, University of Iowa Hospitals and Clinics, Iowa City, IA, USA

^bDepartment of Communication Sciences and Disorders, University of Iowa, Iowa City, IA, USA

[#] These authors contributed equally to this work.

Abstract

Background: Pitch perception of complex tones relies on place or temporal fine structure-based mechanisms from resolved harmonics and the temporal envelope of unresolved harmonics. Combining this information is essential for speech-in-noise performance, as it allows segregation of a target speaker from background noise. In hybrid cochlear implant (H-CI) users, low frequency acoustic hearing should provide pitch from resolved harmonics while high frequency electric hearing should provide temporal envelope pitch from unresolved harmonics. How the acoustic and electric auditory inputs interact for H-CI users is largely unknown. Harmonicity and inharmonicity are emergent features of sound in which overtones are concordant or discordant with the fundamental frequency. We hypothesized that some H-CI users would be able to integrate acoustic and electric information for complex tone pitch perception, and that this ability would be correlated with speech-in-noise performance. In this study, we used perception of inharmonicity to demonstrate this integration.

Methods: Fifteen H-CI users with only acoustic hearing below 500 Hz, only electric hearing above 2 kHz, and more than 6 months CI experience, along with eighteen normal hearing (NH) controls, were presented with harmonic and inharmonic sounds. The stimulus was created with a low frequency component, corresponding with the H-CI user's acoustic hearing (fundamental frequency between 125 and 174 Hz), and a high frequency component, corresponding with electric hearing. Subjects were asked to identify the more inharmonic sound, which requires the perceptual integration of the low and high components. Speech-in-noise performance was tested in both groups using the California Consonant Test (CCT), and perception of Consonant-Nucleus-Consonant (CNC) words in quiet and AzBio sentences in noise were tested for the H-CI users.

Results: Eight of the H-CI subjects (53%), and all of the NH subjects, scored significantly above chance level for at least one subset of the inharmonicity detection task. Inharmonicity detection ability, but not age or pure tone average, predicted speech scores in a linear model. These results

*Corresponding author. Communication Sciences & Disorders, Wendell Johnson Speech and Hearing Center, Iowa City, IA 52242, USA., inyong-choi@uiowa.edu (I. Choi).

²Permanent addresses: Institut de Neurosciences Cognitives et Intégratives d'Aquitaine, CNRS (UMR 5287), Bordeaux, France; Department of Otorhinolaryngology and Skull Base Surgery, Pellegrin University Hospital F-33000, University of Bordeaux F-33000, Bordeaux, France.

were significantly correlated with speech scores in both quiet and noise for H-CI users, but not with speech in noise performance for NH listeners. Musical experience predicted inharmonicity detection ability, but did not predict speech performance.

Conclusions: We demonstrate integration of acoustic and electric information in H-CI users for complex pitch sensation. The correlation with speech scores in H-CI users might be associated with the ability to segregate a target speaker from background noise using the speaker's fundamental frequency.

Keywords

Hybrid cochlear implantation; Acoustic electric integration; Duplex pitch; Harmonicity; Speech perception

1. Introduction

Cochlear implants (CIs) provide electric stimulation to the auditory nerves of deaf and hard of hearing patients. There have been over 300,000 cochlear implantations worldwide as of 2013 (NIDCD, 2017), and they have been shown to improve speech outcomes, especially in quiet situations. However, there are limitations to this technology. The implant processors code the temporal envelope, but do not provide temporal fine structure information (Heng et al., 2011). In addition, place information, which plays some role in pitch processing, is degraded in electric hearing. Therefore, the fundamental frequency of a voice, which is an important cue especially for performance on speech tasks in noise, is not easily accessible to those who use electric hearing (Auinger et al., 2017).

To address this issue, hybrid cochlear implants (H-CIs) aim to preserve low frequency acoustic hearing while electrically stimulating the auditory nerve for high frequency sounds. They are designed for patients with severe to profound high frequency hearing loss, but a relatively good low frequency residual hearing. Generally, these patients have a limited functional benefit with conventional hearing aids, but they often don't meet the indication criteria for a traditional cochlear implant given their residual hearing in the low frequency region. The H-CI combines the principle of traditional CI and hearing aid; a shorter electrode array inserted in the basal cochlea with minimally traumatic surgery is complemented by acoustic amplification which stimulates the still-functional apical region of the cochlea. Outcomes for those with hybrid cochlear implantation are generally better than for those with electric hearing only, but there is wide outcome variability between patients (Gantz et al., 2016). Whether or how the acoustic and electric (A + E) signals are integrated in hybrid cochlear implantees is not known. We therefore aimed to investigate the integration of acoustic and electric information.¹

When the overtones of a complex sound match with its fundamental frequency (f_0), that is, when the frequencies of its spectral components are small integer multiples of a common f_0 , they are in a harmonic relationship. Harmonicity is experienced as a single complex tone,

³Abbreviations: H-CI, hybrid cochlear implant; A + E, acoustic and electric; CNC, Consonant Nucleus Consonant; CCT, California Consonant Test; SNR, signal to noise ratio; NH, normal hearing; PTA, pure tone average.

from a single sound source and with a single pitch; overtones seem to perceptually fuse in a single auditory object, a phenomenon commonly known as *harmonic fusion*. When the overtones do not match with the f_0 , inharmonicity is experienced as a buzzing, an extra pitch, or beating (Bonnard et al., 2013; Plomp, 1967; Viemeister et al., 2001). Inharmonicity detection has been studied in normal hearing and hearing impaired individuals. For subjects with a cochlear hearing loss, inharmonicity detection is still possible but the threshold is generally worse than for normal hearing subjects (Bonnard et al., 2017).

In the current study, we used an inharmonicity detection task to investigate the integration of A + E information provided by a hybrid cochlear implant. We hypothesized that ability to integrate A + E signals, as demonstrated by inharmonicity detection performance, is related to speech in noise performance.

2. Materials and methods

All procedures were reviewed and approved by the local Institutional Review Board. Informed consent was obtained and all work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.1. Subjects

2.1.1. Hybrid CI users—Adult H-CI users, who had only acoustic hearing below 500 Hz, only electric hearing above 2 kHz, and more than 6 months CI experience, were selected from the research subject pool at a tertiary medical center. All testing was performed in an audiometric sound booth with sound stimuli presented by a single front and center speaker (JBL LOFT40). Distance to the sound source was fixed. Stimuli were presented with, and statistics were calculated with, Matlab software (The Mathworks, Inc., R2016b).

2.1.2. Normal hearing controls—Adult normal hearing (NH) controls were selected from the student population of a large university. Pure tone audiometry verified normal hearing status. Speech-in-noise performance was tested with the California Consonant Test as described for H-CI users below, except in multitalker babble at -3 and 3 dB SNR. The inharmonicity detection task proceeded as described below, including blocking of one ear with an earplug.

2.2. Experimental procedure

Within 24 h before their participation, subjects underwent audiometric testing and verification of CI and HA fittings by a licensed audiologist. A variant of the Iowa Musical Background Questionnaire (IMBQ) was used to quantify musical experience (Drennan et al., 2015). Musical experience score was calculated as number of cumulative years to date of formal music lessons added to the number of years of music ensemble experience. To test speech understanding in quiet for H-CI users, two lists of 50 words each were administered from the open-set Consonant Nucleus Consonant (CNC) word recognition test (Tillman and Carhart, 1966). The total number of correct words were scored as a percentage and the average of two lists was reported. This testing was performed in both a hybrid condition, in which electric and acoustic stimulation was provided to the implanted ear with an earplug

blocking sound transmission to the contralateral ear, and in a combined condition (with a different two lists), in which the contralateral ear had the benefit of acoustic amplification in addition to the ipsilateral acoustic plus electric stimulation. To test speech understanding in noise for H-CI users, we used the AzBio test, which requires the identification of words in sentences (Spahr et al., 2012). The sentences were presented in 10-talker babble at 5 dB SNR. The total number of words repeated correctly was scored for the combined condition. For further testing of speech in noise in both H-CI users and NH listeners, the California Consonant Test (CCT) was used (Owens and Schubert, 1977). The words of the CCT were presented in multitalker babble at 7 and 13 dB SNR in a four-alternative forced choice task. Scores for the combined condition were reported as percent correct.

Subjects were presented with one harmonic and one inharmonic sound in a two-interval, two-alternative, forced choice task in which they were asked to identify the more inharmonic sound. The stimulus was synthesized with a complex, presumably resolved, low frequency component, corresponding with acoustic hearing (f_0 roving between 125 and 174 Hz), and a complex unresolved high frequency component, corresponding with electric hearing. To keep the temporal envelope consistent between trials, phases of individual sine wave frequency components were not randomized. Interval duration was 0.5 s, with 0.02s cosine ramps and 0.4s interstimulus interval (ISI). The low frequency component consisted of f_0 and all harmonics up to either 400 Hz or 700 Hz, depending on the acoustic/electric crossover frequency for each subject, to ensure no bleeding of the low frequency component into the electric amplification. In addition, subject amplification maps were confirmed to not include overlapping electric amplification below the upper limit of the low frequency component. The high frequency component either matched the low component's f_0 (harmonic condition), or the upper f_0 was mistuned by $\pm 20\%$ or $\pm 40\%$ (inharmonic conditions). Only harmonics between 2300 and 3500 Hz were included in this component, leading to a temporal envelope oscillation whose frequency is at the missing fundamental frequency. Because of this constant bandpass filter, two to five electrodes were active for each subject, depending on their frequency maps, for the duration of the test. Depending on the f_0 and the mistuning values, the lower component harmonic ranks were from 1 to 5 (resolved for a NH listener), and the higher component harmonic ranks were from 11 to 38 (unresolved). All sounds were presented in threshold-equalizing noise (TE Noise), extending 0.5s before and after the stimulus, to avoid interactions between low frequency spectral components and any possible cochlear combination tones produced by the high frequency spectral components (Moore et al., 2000). The contralateral ear was blocked with an earplug rated for 39.5 ± 3.7 dB attenuation at 1 kHz, and confirmed to have attenuation of at least 25 dB at 1 kHz with our usage, to avoid acoustic input, especially in the high frequencies, from the contralateral ear. Identification of inharmonicity therefore required integration of the acoustic low frequency and electric high frequency sounds for hybrid cochlear implant subjects.

Testing required two personalization steps, one training step, and the detection task itself:

- a. Pure tone detection in noise. A series of pure tones was played in noise in an adaptive staircase to determine the thresholds of hearing tones in noise at three

different frequencies in the subject's acoustic hearing. The SNR for the stimulus was set at 20 dB above the mean threshold SNR.

- b. Loudness matching. In a 2-interval, 3-alternative choice adaptive task, a low frequency component alone was followed by a high frequency component alone, and subjects were asked to identify the louder of the two, with a third answer option for if the stimuli were perceived as having approximately the same loudness. Fundamental frequencies were roved as in the main experiment (described above), and all sounds were presented in noise with SNR as set in the above task. Based on the subject's response, the loudness of each interval was changed to make them more equivalent. The task was continued until the intervals of three successive items were determined to have approximately the same loudness. The same ratio of RMS levels which provided this approximately equal perception of loudness was used to construct the harmonic and inharmonic sounds for the remainder of the experiment.
- c. Training. Subjects were trained on the concept of harmonicity using a slide presentation outlining the concepts of harmonicity and inharmonicity, several examples of harmonic and inharmonic sounds with free discussion about the stimuli, and ten sample test items with feedback given. During this step, overall loudness was set to a comfortable level, but remained under 70dBA for all subjects. Since we did not expect that subjects could perform this task, there was no cutoff for performance before proceeding to the main test.
- d. Inharmonicity detection. A block of 150 harmonic/inharmonic dyads were presented in random order. Subjects were asked to identify the position of the inharmonic sound using a keypad. The trial structure and two possible stimuli are shown in Fig. 1. Chance level was calculated using the chi-square test for difference from 0.5, based on the number of trials that were included in each subset of the test (e.g., for "negative inharmonicity detection" with upper f_0 mistuned by -20 or -40%) or the entire test ("all conditions combined," including all four mistuning values). Therefore, the number of trials and therefore the statistically significant chance level vary according to which subset is referenced.

2.3. Computer simulation of strategies

There were asymmetries necessary in the testing paradigm to maintain the harmonic number above eleven in the higher component of the sound, and to avoid a very low mistuned f_0 , for the different mistuning conditions. Though the fundamental frequency of the harmonic sounds roved between 125 and 174 Hz, the inharmonic sounds' fundamental frequencies were half that range for each condition: Positive mistuning fundamental frequency roved between 125 and 149 Hz, while negative mistuning was between 150 and 174 Hz. The difference in relative distribution of pitches between the harmonic and inharmonic conditions served as a type of "catch," which allowed us to test whether subjects were using pitch-based, as opposed to harmonic-integration-based, strategies for answering. These pitch-based strategies could lead to performance significantly above or below chance level in some conditions, without using harmonic fusion itself. However, given the random order of

the four experimental mistuning conditions in a block (−20%, +20%, −40%, +40%), a subject using one of these strategies would obtain a specific answer pattern depending on the condition. We therefore designed a computer simulation using Matlab (The Mathworks, Inc., R2016b) to implement four pitch-based strategies during 1,000,000 trials, and predict resulting answer patterns. The strategies were: a. listen only to the lower components of each of the two sounds and systematically choose the sound with the highest pitch in the lower component; b. listen only to the lower components and systematically choose the lowest pitch; c. listen only to the higher components and systematically choose the highest pitch; d. listen only to the higher components and systematically choose the lowest pitch.

3. Results

3.1. Subject characteristics and demographics

Of 34 H-CI users who participated in research at a single tertiary medical center over the seven-month period starting 3/1/2017, fifteen H-CI users (44%) met inclusion criteria. These subjects had only acoustic hearing below 500 Hz, only electric hearing above 2 kHz, and more than 6 months CI experience. All eligible subjects participated (ages 60–85 years, mean 70.37 years, median 70.08 years, standard deviation 7.60 years, 60% female). Demographic information is summarized in Table 1. Mean ipsilateral unaided (residual) hearing pure tone average (PTA) at 0.5 kHz, 1 kHz, and 2 kHz was 82.4 dB HL (median 81.7 dB HL, standard deviation 10.4 dB HL), while on the contralateral side the unaided mean PTA was 73.11 dB HL (median 75 dB HL, standard deviation 12.62 dB HL). In no case did the maximum overall sound presentation level exceed a single frequency (at or over 2 kHz) threshold of hearing loss plus earplug attenuation. Audiograms are shown in Fig. 2.

3.2. Some H-CI subjects can detect inharmonicity across acoustic and electric domains

Five of fifteen H-CI subjects and 16/18 NH subjects were able to detect inharmonicity (scored significantly above the calculated chance level) in all conditions combined. However, negative mistuning was easier than positive mistuning, and 40% mistuning was easier than 20% mistuning, for all subjects. These results are shown in Fig. 3

3.3. Computer simulations of pitch-based answering strategies match three H-CI subjects' answer patterns

Simulation of possible answering strategies revealed specific answer patterns depending on the condition. Results of the simulation for each mistuning condition are shown in Table 2. We found three H-CI users whose answer patterns matched one of these simulated patterns. Of these, one (CI-15) had performance significantly above chance level in all conditions combined, and two (CI-12 and CI-14) had performance significantly above chance level in the negative mistuning condition only. In all of these cases, subjects performed significantly above chance level in the negative mistuning condition and significantly below chance level in the positive mistuning condition, as predicted by the simulation using strategies a and d. Consequently, given the ambiguity in interpretation of their results, we elected to exclude these three subjects from the correlation analyses described below.

3.4. Inharmonicity detection correlates with speech scores for H-CI users but not for NH listeners

We investigated the relationship between medians for groups created by the categorical variable above-chance/below-chance performance on harmonicity (negative mistuning only, chance level at 66%), using the Wilcoxon rank sum test, with the scores on CCT words in babble (combined condition, $p = 0.0455$), CNC words in quiet (combined condition, $p = 0.0051$) and AzBio sentences in babble (combined condition, $p = 0.1136$). Our statistical “chance” level for this analysis was set at 66% for percent correct score on the negative mistuning condition, compared to the actual chance level of 50%, in order to be as conservative as possible in the interpretation of who could and could not perform the task.

This conservative interpretation of “statistical chance level” does not preclude us from using the continuous variable of percent correct for inharmonicity detection ability in regression analyses. Therefore, using Pearson correlations, we further investigated the relationship between inharmonicity detection and speech-in-noise performance. Percent correct for negatively mistuned inharmonicity detection was strongly correlated with percent correct for CNC words in quiet (combined condition, $r = 0.769$, $p = 0.0034$), CNC words in quiet (hybrid condition, $r = 0.768$, $p = 0.0036$), AzBio sentences in babble (combined condition, $r = 0.630$, $p = 0.028$), and CCT words in babble (combined condition, $r = 0.639$, $p = 0.025$), for H-CI users. These results are shown in Fig. 4.

Inharmonicity detection in the negative mistuning condition was not correlated with identification of CCT words in babble for NH listeners ($r = -0.088$, $p = n.s.$). There was no statistically significant correlation between length of pre-implantation hearing loss and performance on inharmonicity detection ($r = 0.16$, $p = n.s.$).

3.5. Inharmonicity detection ability, but not pure tone average or current age, predicts speech performance for H-CI users

To evaluate predictors of speech performance, we used a linear regression model and three-way ANOVA with independent variables harmonicity performance (inharmonicity detection percentage in the negative mistuning condition), low frequency ipsilateral unaided PTA (0.25 kHz, 0.5 kHz, and 1 kHz) and age at testing. Of these, only harmonicity performance contributed significantly to performance on CNC words in quiet (hybrid condition, $F_{1,8} = 7.476$, $p = 0.0257$), CCT words in noise (combined condition, $F_{1,8} = 10.237$, $p = 0.0126$), and AzBio sentences in noise (combined condition, $F_{1,8} = 9.629$, $p = 0.0146$). The speech scores predicted by the linear model are compared to the actual speech scores in Fig. 5.

3.6. Musical experience, but not pure tone average or current age, predicts inharmonicity detection ability

To evaluate predictors of harmonicity performance, we used a linear regression model and three-way ANOVA with the following independent variables: low frequency ipsilateral PTA, age at testing, and musical experience. Of these, musical experience was the only predictor of performance on the negative inharmonicity detection task ($F_{1,8} = 5.961$, $p = 0.0405$). But, musical experience did not predict performance on most of the speech tasks except for AzBio sentences in noise ($F_{1,8} = 6.239$, $p = 0.0371$).

4. Discussion

4.1. Integration of acoustic and electric stimulation can occur in hybrid cochlear implantees

In order to detect inharmonicity in our task, H-CI subjects must integrate acoustic and electric information. Alternative strategies have been ruled out: First, H-CI subjects could not integrate low frequency acoustic information from the implanted ear with high frequency acoustic information from the non-implanted ear given the stimulus level and residual hearing attenuated by the earplug. Second, we excluded three subjects whose answer patterns matched computer simulations of pitch-based, as opposed to harmonic-integration-based, strategies. Four subjects performed significantly better than chance level, while using a harmonic-integration-based strategy, in all four mistuning conditions combined. An additional subject performed significantly better than chance level, but only in the negative mistuning condition, without using a pitch-based strategy. Because contralateral masking was not used, it is possible that low frequency acoustic information from the non-implanted ear was integrated with the high frequency electric information from the implanted ear. More acoustic information was available to the implanted ear, since that ear was not plugged, and so we believe that ipsilateral harmonic integration is the more likely driver of our results. In any case, one third of H-CI users tested could integrate electric and acoustic information from two frequency registers.

Electric/acoustic integration has been suggested in the past using pitch perception and vowel perception across ears in bimodal long electrode (Guerit et al., 2014; Reiss et al., 2014a) and single sided deaf (Guerit et al., 2014) cochlear implant users. Difficulties with integration of acoustic and electric information during dichotic listening were suggested for those populations. However, ours is the first experiment we are aware of which focuses on ipsilateral integration of acoustic and electric information. Direct comparisons between these studies is challenging due to differences in study populations and specifics of the perceptual tasks, but our finding of A + E integration ability in some H-CI users, contrasted with the difficulties in dichotic listening shown previously, suggests that within-ear integration of A + E signals may be easier. Further study could help answer this question.

During training for the task, some H-CI subjects reported that the inharmonic sound was more “buzzing” or more “harsh” than the harmonic sound. We don’t know, however, whether harmonic sounds were more or less fused, or more or less pleasing or concordant, for the H-CI subjects as for the normal hearing individuals. Indeed, at least one H-CI subject reported that “both sound bad, but [the inharmonic] one is worse.” While we cannot be sure that the percepts are the same for H-CI and NH subjects, the simple ability to detect a difference between the two within the constraints of our experimental design argues for integration of electric and acoustic information in those H-CI subjects who can successfully complete the task.

4.1.1. Impact on theories of duplex (place and temporal) pitch—According to the duplex theory of pitch, normal hearing individuals can judge the harmonicity of complex tones based on place of cochlear stimulation and/or temporal information, the latter including temporal envelope and temporal fine structure information. Because processing

strategies discard fine temporal structure, our H-CI subjects only had access to the temporal envelope of the high component of our sounds. We know that cochlear place coding of the high component was useless for H-CI users in performance of our task because 1) we used the same frequency band for the high component of the sound in all stimuli so the same electrodes were always activated in all stimuli, and 2) there is a significant mismatch between the natural tonotopic organization of the cochlea and the actual electrode positions in the cochlea (Reiss et al., 2014b). For our H-CI subjects, the temporal envelope information of the unresolved (high frequency) harmonics was presented to the wrong place in the cochlea, i.e., with a tonotopic mismatch (Lee et al., 2010), even if plasticity of central auditory pathways may sometimes reduce this mismatch (Reiss et al., 2014b). However, the subjects were still able to perform our task. These data are consistent with a previous study on normal hearing listeners, showing that subjects can simultaneously integrate f_0 information from both resolved and unresolved groups of harmonics (Carlyon, 1992). Our study demonstrates that this simultaneous integration is still possible in H-CI subjects, even when unresolved components are presented with a tonotopic mismatch and without temporal fine structure information. This result is consistent with models taking into account resolved and unresolved components for complex tone pitch extraction (e.g., Meddis and Omand, 1997).

4.1.2. Asymmetry between positive and negative mistuning—Inharmonicity detection is easier with negative than positive mistuning. Pitch-based answer strategies could account for this asymmetry in three H-CI subjects, and none of the NH-controls, while the asymmetry is found in the majority of subjects in both groups. Given the differences in our f_0 ranges, since negative mistunings were presented for slightly higher fundamental frequencies (150–174 Hz for negative mistuning vs 125–149 Hz for positive mistuning), we cannot exclude the possibility that this asymmetry could reflect easier detection of inharmonicity when f_0 is higher. Nevertheless, it seems unlikely that such a large asymmetry in difficulty detecting positive and negative mistuning would be the result of such a small, less than half-octave, difference between the two frequency registers.

Previous studies found an asymmetry between “stretched” and “compressed” frequency ratios in inharmonicity detection tasks on NH-listeners, with better performance for smaller ratios (Demany et al., 1991). Although their experimental paradigms were different and based only on resolved components, the similarity of the results suggests that they might be linked to a common – but still unexplained – origin. Interestingly, the asymmetry was present in our study for NH and H-CI subjects, suggesting that inharmonicity detection relies on the same mechanism in both groups.

4.2. Inharmonicity detection and speech-in-noise performance

While inharmonicity detection is not linked with better speech-in-noise (SiN) performance in normal hearing listeners, ability to integrate the temporal envelope of the electric stimulation with the place and temporal fine structure information provided by acoustic hearing is strongly correlated with speech perception, both in quiet and in noise, for hybrid CI users. Even when using a categorical variable for inharmonicity detection ability, this relationship is maintained at a statistically significant level for all but the AzBio test. These

results are consistent with previous studies demonstrating the importance of the sensitivity to harmonicity and perceptual grouping between components in presence of simultaneous sound sources, e.g. for segregation of harmonic complex tones (Roberts and Brunstrom, 1998) or concurrent vowel identification (de Cheveigné et al., 1997).

If the NH and H-CI subjects were age matched, we might have seen a different correlation between negative inharmonicity detection and speech scores for the NH group. However, this potential difficulty is outweighed by the need to have a truly normal hearing control group; with an age-matched group we would not expect completely normal hearing, even in the presence of normal audiograms (Plack et al., 2014).

4.2.1. A central mechanism is likely to be responsible for differences in inharmonicity detection performance in H-CI users—

The lack of correlation between inharmonicity detection ability and crossover frequency, and the lack of contribution of the low frequency PTA in a model predicting inharmonicity detection performance, argues against peripheral differences as the driving factor behind this variance in ability. With the caveat that the audiogram is a coarse measure of peripheral transmission, and pure tone audiometry does not exclude the possibility that poorer-performing subjects may have hidden hearing loss affecting the fidelity of their acoustic pitch representation, these results suggest a more central mechanism behind the integration of acoustic and electric information. The neurophysiological bases of harmonic fusion are still largely unknown, but data indicate that some neurons in the primary auditory cortex of marmosets show a much stronger response when two harmonically related tones are simultaneously presented (i.e., one at the neuron's characteristic frequency and the other at double the characteristic frequency) than when tones are individually presented (Feng and Wang, 2017; Wang, 2013). These kinds of neurons could support the mechanism of harmonic integration, and we would not necessarily expect that their function would be dramatically affected in H-CI users despite the degradation of peripheral information. Analysis of cortical pathways using neuroimaging techniques like EEG will help guide our understanding of the mechanisms underlying the integration of acoustic and electric information and associated perceptual improvement.

4.2.2. Inharmonicity detection ability may drive a fundamental frequency benefit for auditory scene segregation—

The correlation between inharmonicity detection and speech scores in H-CI users might be associated with the ability to segregate a target speaker from background noise using the speaker's fundamental frequency. Segregation of audio streams from different sources requires separation of simultaneous harmonic sounds with different fundamental frequencies, which could benefit from inharmonicity detection ability (Carlyon and Gockel, 2008; de Cheveigné et al., 1997).

At our center, one strategy for rehabilitation of H-CI users following implantation is to maximize use of the implant. If we speculate that the relationship between inharmonicity detection is causal and not simply correlative, more focused training strategies might be warranted; for example, it may be possible to train subjects to integrate their acoustic and electric hearing using the concept of inharmonicity.

5. Conclusions

H-CI users' ability to integrate acoustic and electric information may impact their speech-in-noise performance. Variation in this ability may be caused by central factors, but further study is necessary to characterize peripheral contributions to this ability. Our findings suggest that further characterization of integration of acoustic and electric inputs in H-CI users might provide clues for improvements in the device design or in user rehabilitation strategies.

Acknowledgements

This work was supported by NIDCD P50 DC000242 31A1 (PI: Gantz), T32 DC000040–24, Hearing Health Foundation Emerging Research Grant (PI: Choi), and AFON 2017 grant (Association Française d'Otologie et d'Otoneurologie). We thank Subong Kim and Dr. Jihwan Woo for helping with data collection.

References

- Auinger AB, Riss D, Liepins R, Rader T, Keck T, Keintzel T, Kaider A, Baumgartner WD, Gstöttner W, Arnoldner C, 2017 Masking release with changing fundamental frequency: electric acoustic stimulation resembles normal hearing subjects. *Hear. Res.* 350, 226–234. [PubMed: 28527538]
- Bonnard D, Dauman R, Semal C, Demany L, 2017 The effect of cochlear damage on the sensitivity to harmonicity. *Ear Hear.* 38, 85–93. [PubMed: 27992390]
- Bonnard D, Michey C, Semal C, Dauman R, Demany L, 2013 Auditory discrimination of frequency ratios: the octave singularity. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 788–801. [PubMed: 23088507]
- Carlyon RP, 1992 The psychophysics of concurrent sound segregation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 336, 347–355. [PubMed: 1354374]
- Carlyon RP, Gockel HE, 2008 Effects of harmonicity and regularity on the perception of sound sources In: Yost WA, Popper AN, Fay RR (Eds.), *Auditory Perception of Sound Sources*. Springer Handbok of Auditory Research, vol. 29 Springer, Boston, MA, pp. 191–213.
- de Cheveigné A, McAdams S, Marin CMH, 1997 Concurrent vowel identification. II. Effects of phase, harmonicity, and task. *J. Acoust. Soc. Am.* 101, 2848–2856.
- Demany L, Semal C, Carlyon RP, 1991 On the perceptual limits of octave harmony and their origin. *J. Acoust. Soc. Am.* 90, 3019–3027.
- Drennan WR, Oleson JJ, Gfeller K, Crosson J, Driscoll VD, Won JH, Anderson ES, Rubinstein JT, 2015 Clinical evaluation of music perception, appraisal and experience in cochlear implant users. *Int. J. Audiol.* 54, 114–123. [PubMed: 25177899]
- Feng L, Wang X, 2017 Harmonic template neurons in primate auditory cortex underlying complex sound processing. *Proc. Natl. Acad. Sci. U. S. A.* 114, E840–E848. [PubMed: 28096341]
- Gantz BJ, Dunn C, Oleson J, Hansen M, Parkinson A, Turner C, 2016 Multicenter clinical trial of the Nucleus Hybrid S8 cochlear implant: final outcomes. *Laryngoscope* 126, 962–973. [PubMed: 26756395]
- Guerit F, Santurette S, Chalupper J, Dau T, 2014 Investigating interaural frequency-place mismatches via bimodal vowel integration. *Trends Hear* 18.
- Heng J, Cantarero G, Elhilali M, Limb CJ, 2011 Impaired perception of temporal fine structure and musical timbre in cochlear implant users. *Hear. Res.* 280, 192–200. [PubMed: 21664263]
- Lee J, Nadol JB, Jr., Eddington DK, 2010 Depth of electrode insertion and post-operative performance in humans with cochlear implants: a histopathologic study. *Audiol. Neuro. Otol.* 15, 323–331.
- Meddis R, O'Mard L, 1997 A unitary model of pitch perception. *J. Acoust. Soc. Am.* 102, 1811–1820. [PubMed: 9301058]
- Moore BC, Huss M, Vickers DA, Glasberg BR, Alcantara JJ, 2000 A test for the diagnosis of dead regions in the cochlea. *Br. J. Audiol.* 34, 205–224. [PubMed: 10997450]

- NIDCD, 2017 Cochlear Implants [Online]. Available by NIH. <https://www.nidcd.nih.gov/health/cochlear-implants> (posted 3/6/2017; verified 12/16/2017).
- Owens E, Schubert ED, 1977 Development of the California consonant test. *J. Speech Lang. Hear. Res.* 20, 463–474.
- Plack CJ, Barker D, Prendergast G, 2014 Perceptual consequences of “hidden” hearing loss. *Trends Hear* 18.
- Plomp R, 1967 Beats of mistuned consonances. *J. Acoust. Soc. Am.* 42, 462–474. [PubMed: 6075940]
- Reiss LA, Ito RA, Eggleston JL, Wozny DR, 2014a Abnormal binaural spectral integration in cochlear implant users. *J Assoc Res Otolaryngol* 15, 235–248. [PubMed: 24464088]
- Reiss LA, Turner CW, Karsten SA, Gantz BJ, 2014b Plasticity in human pitch perception induced by tonotopically mismatched electro-acoustic stimulation. *Neuroscience* 256, 43–52. [PubMed: 24157931]
- Roberts B, Brunstrom JM, 1998 Perceptual segregation and pitch shifts of mistuned components in harmonic complexes and in regular inharmonic complexes. *J. Acoust. Soc. Am.* 104, 2326–2338. [PubMed: 10491697]
- Spahr AJ, Dorman MF, Litvak LM, Van Wie S, Gifford RH, Loizou PC, Loiselle LM, Oakes T, Cook S, 2012 Development and validation of the AzBio sentence lists. *Ear Hear.* 33, 112–117. [PubMed: 21829134]
- Tillman TW, Carhart R, 1966 An Expanded Test for Speech Discrimination Utilizing CNC Monosyllabic Words. SAM-TR USAF School of Aerospace Medicine. SAM-TR-66-55 [Technical Report].
- Viemeister NF, Rickert M, Stellmack M, Breebar DJ, Houtsma AJM, Kohlrausch A, Prijs VF, Schoonhoven R, 2001 Beats of Mistuned Consonances, Physiological and Psychophysical Bases of Auditory Function, pp. 113–120. Shaker, Maastricht, The Netherlands.
- Wang X, 2013 The harmonic organization of auditory cortex. *Front. Syst. Neurosci* 7, 114. [PubMed: 24381544]

Fig. 1.

Trial structure and two possible stimuli S1 and S2. S1 is harmonic, S2 is inharmonic with a mistuning value of +20%. Subjects were asked to identify the more inharmonic sound (S2 in this example).

Fig. 2.

a H-CI subject composite audiogram, unaided implanted ear. Error bars reflect standard deviations. b H-CI subject composite audiogram, unaided unimplanted ear. Error bars reflect standard deviations. c NH subject composite audiogram.

Fig. 3.

a-b H-CI vs NH results for inharmonicity detection in the negative mistuning condition.

Unshaded region indicates significantly above chance level performance. Some hybrid CI users can detect inharmonicity in our paradigm, indicating that A + E integration can occur.

c Negative versus positive mistuning. Red diamonds indicate significantly above chance level performance. Negative mistuning is easier than positive mistuning for all subjects. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 4.

a-c. Correlations between inharmonicity detection and speech scores, H-CI users. Red diamonds indicate significantly above chance level performance on inharmonicity detection. The CNC words were presented in the hybrid condition, with the contralateral ear blocked, while the CCT words and AzBio sentences were presented in a combined condition, making use of both the implanted and non-implanted ears. d Correlation between musical experience and inharmonicity detection, H-CI users. Red diamonds indicate significantly above chance level performance on inharmonicity detection. Musical experience score was calculated as number of cumulative years to date of formal music lessons added to the number of years of music ensemble experience. The linear regression curve (dotted line) was fit after excluding the outlier, while with the outlier included we used the exponential fitting function of Matlab (The Mathworks, Inc., R2016b) to create the dashed best-fit curve. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 5.
a-c Correlation between prediction of linear model and actual scores on CNC, AzBio, and CCT tests.

Subject demographic information – Hybrid cochlear implantees. Length of preoperative deafness was calculated starting from self-report of first need for a hearing aid. The crossover frequency represents the lowest frequency assigned to electrical stimulation. Musical experience score was calculated as number of cumulative years to date of formal music lessons added to the number of years of music ensemble experience.

Table 1a

Subject number	Gender	Age at Stimulation (y)	Length of Preoperative Deafness (y)	Device Type	Device Ear	Post-Activation Interval (m)	Crossover Frequency (Hz)	Musical Experience (y)
<i>CI-1</i>	M	65	9	Nucleus Hybrid L24	R	6	563	4
<i>CI-2</i>	F	75	11	Nucleus Hybrid L24	R	6	563	9
<i>CI-3</i>	M	75	8	Nucleus Hybrid L24	R	12	1188	7
<i>CI-4</i>	M	70	14	Nucleus Hybrid L24	R	12	813	0
<i>CI-5</i>	F	62	24	Nucleus Hybrid S12	L	96	813	5
<i>CI-6</i>	M	83	3	CI24RE	L	6	688	4
<i>CI-7</i>	F	64	22	Nucleus Hybrid L24	L	6	563	74.5
<i>CI-8</i>	F	55	12	S8	R	132	630	7
<i>CI-9</i>	M	84	19	Nucleus Hybrid L24	L	12	938	0
<i>CI-10</i>	F	76	16	Nucleus Hybrid L24	R	12	813	0
<i>CI-11</i>	M	48	10	Nucleus EAS2	R	169	563	20
<i>CI-12</i>	M	65	25	Nucleus Hybrid L24	R	12	688	0
<i>CI-13</i>	F	71	35	Nucleus CI422	L	25	563	10
<i>CI-14</i>	M	57	7	Nucleus Hybrid L24	L	43	813	0
<i>CI-15</i>	F	54	20	Nucleus Hybrid S12	L	72	1188	10

Table 1b

Subject demographic information – Normal hearing controls. Musical experience score was calculated as number of cumulative years to date of formal music lessons added to the number of years of music ensemble experience.

Subject number	Gender	Age (y)	Musical Experience (y)
<i>NH-1</i>	F	21	12
<i>NH-2</i>	F	22	1
<i>NH-3</i>	F	20	7
<i>NH-4</i>	F	20	12
<i>NH-5</i>	F	20	7
<i>NH-6</i>	M	19	12
<i>NH-7</i>	F	26	15
<i>NH-8</i>	M	22	3
<i>NH-9</i>	F	21	8
<i>NH-10</i>	F	20	7
<i>NH-11</i>	F	23	12
<i>NH-12</i>	F	19	4
<i>NH-13</i>	M	23	*
<i>NH-14</i>	M	23	*
<i>NH-15</i>	F	20	12
<i>NH-16</i>	M	22	13
<i>NH-17</i>	M	19	3
<i>NH-18</i>	M	17	8

* Data not available.

Table 2

Results of the simulation of pitch-based strategies (as percentage of correct answers) for each mistuning condition. Strategies are as follows: a. listen only to the lower components of each of the two sounds and systematically choose the sound with the highest pitch in the lower component; b. listen only to the lower components and systematically choose the lowest pitch; c. listen only to the higher components and systematically choose the highest pitch; d. listen only to the higher components and systematically choose the lowest pitch.

Strategy	Mistuning condition			
	-20%	+20%	-40%	+40%
<i>a</i>	74	23.9	74	23.9
<i>b</i>	23.9	74	23.9	74
<i>c</i>	10.5	78.3	0	100
<i>d</i>	87.9	20	100	0