frontiers in Neuroscience

ORIGINAL RESEARCH
published: 21 April 2020
doi: 10.3389/fnins.2020.00301

A New Approach for Noise Suppression in Cochlear Implants: A Single-Channel Noise Reduction Algorithm[1]

Huali Zhou[1], Ningyuan Wang[2], Nengheng Zheng[3], Guangzheng Yu[1] and Qinglin Meng[1]**

[1] Acoustics Lab, School of Physics and Optoelectronics, South China University of Technology, Guangzhou, China, [2] Nurotron Biotechnology Inc., Hangzhou, China, [3] The Guangdong Key Laboratory of Intelligent Information Processing, College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China

The cochlea "translates" the in-air vibrational acoustic "language" into the spikes of neural "language" that are then transmitted to the brain for auditory understanding and/or perception. During this intracochlear "translation" process, high resolution in time–frequency–intensity domains guarantees the high quality of the input neural information for the brain, which is vital for our outstanding hearing abilities. However, cochlear implants (CIs) have coarse artificial coding and interfaces, and CI users experience more challenges in common acoustic environments than their normal-hearing (NH) peers. Noise from sound sources that a listener has no interest in may be neglected by NH listeners, but they may distract a CI user. We discuss the CI noise-suppression techniques and introduce noise management for a new implant system. The monaural signal-to-noise ratio estimation-based noise suppression algorithm "eVoice," which is incorporated in the processors of Nurotron® Enduro™, was evaluated in two speech perception experiments. The results show that speech intelligibility in stationary speech-shaped noise can be significantly improved with eVoice. Similar results have been observed in other CI devices with single-channel noise reduction techniques. Specifically, the mean speech reception threshold decrease in the present study was 2.2 dB. The Nurotron society already has more than 10,000 users, and eVoice is a start for noise management in the new system. Future steps on non-stationary-noise suppression, spatial-source separation, bilateral hearing, microphone configuration, and environment specification are warranted. The existing evidence, including our research, suggests that noise-suppression techniques should be applied in CI systems. The artificial hearing of CI listeners requires more advanced signal processing techniques to reduce brain effort and increase intelligibility in noisy settings.

Keywords: cochlear implant, noise reduction, cocktail party problem, monaural, speech in noise, intelligibility, Nurotron, eVoice

OPEN ACCESS

Edited by:
Yuki Hayashida,
Osaka University, Japan

Reviewed by:
Sebastián Ausili,
University of Miami, United States
Li Xu,
Ohio University, United States

****Correspondence:***
Guangzheng Yu
scgzyu@scut.edu.cn
Qinglin Meng
mengqinglin@scut.edu.cn

Specialty section:
This article was submitted to Neural Technology, a section of the journal Frontiers in Neuroscience

Received: *30 November 2019*
Accepted: *16 March 2020*
Published: *21 April 2020*

Citation:
Zhou H, Wang N, Zheng N, Yu G and Meng Q (2020) A New Approach for Noise Suppression in Cochlear Implants: A Single-Channel Noise Reduction Algorithm. Front. Neurosci. 14:301. doi: 10.3389/fnins.2020.00301

[1]Portions of this work were presented in "Implementation and evaluation of a single-channel noise reduction method in cochlear implants" at the 2017 Conference on Implantable Auditory Prostheses, Lake Tahoe, CA, United States, July 2017; "Neural Interface: Frontiers and Applications" in Advances in Experimental Medicine and Biology, Volume 1101; and "Speech intelligibility test of "eVoice", a new noise-reduction algorithm in Nurotron Enduro systems" at the 2019 Asia Pacific Symposium on Cochlear Implants and Related Sciences, Tokyo, Japan, November 2019.

INTRODUCTION

The cochlear implant (CI) is one of the most successful prostheses ever developed and aims to rehabilitate hearing by transmitting acoustic information into the brains of people with severe to profound hearing impairment by electrically stimulating auditory nerve fibers (Shannon, 2014). The artificial electric hearing provided by current CIs is useful for speech communication but is still far from satisfactory compared with normal hearing (NH), especially in the aspect of speech-in-noise recognition.

The noise issue is a common complaint of CI users (e.g., Ren et al., 2018). Because of variability associated with implant surgery time, hearing history, rehabilitation and training, surgical conditions, devices and signal processing, and so on, large differences in hearing abilities have always been reported within any group of CI users. These reasons behind the CI-NH gap and intersubject CI variance may be classified into "top-down" and "bottom-up" types (Moberly and Reed, 2019; Tamati et al., 2019).

From a practical standpoint, knowledge about "top-down" memory and cognition is useful for rehabilitation and making surgical decisions (Kral et al., 2019), whereas the relationship between speech performance and the "bottom-up" signal processing functions—especially those on the electrode interface—determines the engineering approaches used in current CI systems (Wilson et al., 1991; Loizou, 1999, 2006; Rubinstein, 2004; Zeng, 2004; Zeng et al., 2008; Wouters et al., 2015; Nogueira et al., 2018). Although the "top-down" approach has been suggested to be incorporated into CI systems to form an adaptive closed-loop neural prothesis (Mc Laughlin et al., 2012), we only introduce "bottom-up"–related techniques that might be useful for CI users to tackle the problem of noise masking, as discussed below.

How to send more useful information upward? Sound pressure waveforms are decomposed by healthy cochleae into fine temporal-spectral "auditory images". CIs attempt to capture and deliver the same images but, unfortunately, in a coarse way. Theories in grouping, scene analysis, unmasking, and attention have demonstrated the significance of precise coding of acoustic cues including pitch or resolved harmonics, common onset, and spatial cues. For most CI systems, only temporal envelopes from a limited number of channels can be transferred to the nerve, and current interactions between channels are a key limitation of the multichannel CI framework.

Several research directions have been explored to improve the CI recognition performance of speech in noise by updating the technology of contemporary multichannel devices: (1) stimulating auditory nerves in novel physical ways such as optical stimulation (Jeschke and Moser, 2015) and penetrating nerve stimulation (Middlebrooks and Snyder, 2007); (2) developing intracochlear electrode arrays with different lengths, electrode shapes, and mechanical characteristics (Dhanasingh and Jolly, 2017; Rebscher et al., 2018; Xu et al., 2018); (3) steering and focusing the current spread by simultaneously activating multiple electrodes (Berenstein et al., 2008; Bonham and Litvak, 2008); (4) refining the strategies in the temporal domain by introducing harmonics (Li et al., 2012), timing of zero crossings (Zierhofer, 2003) or peaks (Van Hoesel, 2007), and slowly varying temporal fine structures (Nie et al., 2005; Meng et al., 2016); and (5) enhancing speech or suppressing noise before or within the core signal processing strategies. The first and second directions are developed from the perspective of neurophysiology; the third is mainly based on psychophysical tests; the fourth uses a combination of signal processing and psychophysics, and the fifth mainly concentrates on signal processing. All of these aspects are worth further investigation.

In the last two decades, the fifth approach of enhancing speech or suppressing noise before or within the core signal processing strategies has become a hot topic in academic and industrial research. Noise reduction and speech enhancement are two sides of the same coin, and the goal is to improve intelligibility or quality of speech in noise, in most cases with a signal-to-noise ratio (SNR) enhancement signal processing system. Some noise reduction techniques in telecommunications and hearing aids have been used to process noisy speech signals, and then the processed signals are presented through loudspeakers to CI users (e.g., classic single-channel spectral subtraction) (Yang and Fu, 2005) for feasibility verification. Now there are more sophisticated single-channel noise-reduction algorithms (NRAs) (Chen et al., 2015), directional microphone, or multimicrophone-based beamformers of hearing aids (Chung et al., 2004; Buechner et al., 2014), and more recently deep neural network–based algorithms (Lai et al., 2018; Goehring et al., 2019) that have been tried with CI listeners. Another line of research is to specifically optimize algorithm parameters with a consideration of the differences between CI and NH listeners. The parameters are generally related to the noise estimation or gain function for noise reduction (Hu et al., 2007; Kasturi and Loizou, 2007; Mauger et al., 2012a,b; Wang and Hansen, 2018). All these studies demonstrated significant improvements, which can be explained by the higher SNR yielded by the techniques before or within the CI core strategies.

In the newest versions of CI processors from current commercial companies such as Cochlear® (Hersbach et al., 2012), Advanced Bionics® (Buechner et al., 2010), and MED-EL® (Hagen et al., 2019), one or multiple algorithms of SNR-based monaural noise reduction and spatial cue-based directional microphone or multimicrophone beamformers have been implemented and evaluated. Multimicrophone beamformers significantly improve speech intelligibility for CI recipients in noise. However, it is based on the assumption that target speech and noise sources are spatially separated. Thus, single-microphone NRAs in CI systems are still worthy of attention to improve speech perception in noise, especially in scenarios when the target speech and noise sources are not spatially separated.

Some single-microphone NRAs that are already implemented in commercial CI products have been reported in the literature. ClearVoice is a monaural NRA implemented with the HiRes 120 speech processing strategy (Buechner et al., 2010; Holden et al., 2013). It first estimates noise by assuming that speech energy amplitude changes frequently and background noise energy is less modulated. Then, gain is reduced for channels identified as having mainly noise energy. The noise estimation works at

a time window of 1.3 s, which is the activation time of this algorithm. Experiments showed that ClearVoice can improve speech intelligibility in stationary noise (Buechner et al., 2010; Kam et al., 2012). Another monaural NRA is implemented with the ACE (advanced combination encoder) strategy in Nucleus devices. It uses a minimum statistics algorithm with an optimal smoothing method for noise estimation (Martin, 2001) and an *a priori* SNR estimate (McAulay and Malpass, 1980) in conjunction with a modified Wiener gain function (Loizou, 2007). It was reported to significantly improve hearing in stationary noise (Dawson et al., 2011).

We introduce a recently developed single-channel estimated-SNR–based NRA, termed "eVoice," which has been implemented in the second-generation research processor Enduro^{TM} of Nurotron. Nurotron, a young company based in Irvine, CA, United States, and Hangzhou, Zhejiang, China, currently has more than 10,000 patients implanted. The Nurotron system has 24 electrode channels, and its users' speech performance in quiet and postsurgery development status are comparable with previous data from other brands (Zeng et al., 2015; Gao et al., 2016). The noise estimation in eVoice is processed on a frame-by-frame basis, which is using a relatively short time window. It is based on classical signal processing algorithms and is not the first CI device to use this kind of approach. The aims of this study include reporting the intelligibility experiment results for eVoice and rethinking noise management of a new CI system, which in this case is the Nurotron system.

EVOICE OF NUROTRON: A SINGLE-CHANNEL NRA

The default core strategy of Nurotron is the advanced peak selection (APS) strategy, which is similar to an "*n*-of-*m*" strategy (Zeng et al., 2015). The APS strategy is based on a short-time Fourier transform (STFT) and typically selects eight maxima (an automatic process defined in the coding strategy) for stimulation in each frame (Ping et al., 2017). A block diagram of the APS strategy and eVoice is shown in **Figure 1**. In APS, acoustic input signal is first preamplified, followed by bandpass filtering (the band number m typically equals the active electrode number, i.e., m = 24 in Nurotron devices) and envelope calculation. Then, in peak selection, n bands with the largest amplitude are selected for further non-linear compression and electrical stimulation (typically, n = 8 in Nurotron devices). The eVoice is an envelope-based noise reduction method implemented between envelope calculation and peak selection. It consists of two steps: noise estimation and gain calculation (Wang et al., 2017).

Noise Estimation

The noise estimation algorithm is based on an improved minima-controlled recursive averaging (MCRA-2) algorithm (Rangachari and Loizou, 2006). Noise power in each channel is estimated on a frame-by-frame basis instead of a time window that includes several frames to reduce response time.

Suppose that the noise is additive, then in the time domain, the input signal $y(n)$ can be denoted as

$$y(n) = x(n) + d(n) \tag{1}$$

where $x(n)$ is the clean speech signal and $d(n)$ is the additive noise signal. We use $Y(\lambda, k)$, i.e., the STFT of $y(n)$, to represent the summation magnitude of channel k in frame λ in the frequency domain. The power spectrum of the noisy signal can be smoothed and updated on a frame-by-frame basis using the recursion below:

$$P(\lambda, k) = \eta P(\lambda - 1, k) + (1 - \eta)\left|Y(\lambda, k)\right|^2 \tag{2}$$

where η is a smoothing factor. Then, the local minimum of the power spectrum in each channel can be tracked as follows:

$$P_{min}(\lambda, k) = \begin{cases} P(\lambda, k) & , \ P_{min}(\lambda - 1, k) \geq P(\lambda, k) \\ \gamma P_{min}(\lambda - 1, k) + \frac{1-\gamma}{1-\beta}(P(\lambda, k) & - \beta P(\lambda - 1, k)), \\ & P_{min}(\lambda - 1, k) < P(\lambda, k) \end{cases} \tag{3}$$

where $P_{min}(\lambda, k)$ is the local minimum of the noisy speech power spectrum, and β and γ are constant parameters. The ratio of noisy speech power spectrum to its local minimum can be calculated as follows:

$$S_r(\lambda, k) = \frac{P(\lambda, k)}{P_{min}(\lambda, k)} \tag{4}$$

This ratio is compared against a threshold $T(\lambda, k)$ to determine the speech-presence probability $I(\lambda, k)$ using the criterion below:

$$I(\lambda, k) = \begin{cases} 1, & S_r(\lambda, k) \geq T(\lambda, k) \\ 0, & S_r(\lambda, k) < T(\lambda, k) \end{cases} \tag{5}$$

where $T(\lambda, k)$ is the threshold that is dynamically updated according to the estimated SNR of the previous frame. It is worth mentioning that this threshold is set at a constant level in the literature, and we found from our pilot data that dynamic thresholds performed better than constants during our assessment, so we decided to use dynamic thresholds.

This speech-presence probability $I(\lambda, k)$ can be smoothed as follows:

$$K(\lambda, k) = \alpha K(\lambda, k) + (1 - \alpha) I(\lambda, k) \tag{6}$$

where $K(\lambda, k)$ is the smoothed speech-presence probability, and α is a smoothing constant. The smoothing factor to be used for noise estimation can be updated using the above calculated speech-presence probability:

$$\alpha_s(\lambda, k) = \alpha_d + (1 - \alpha_d) K(\lambda, k) \tag{7}$$

where α_s is the smoothing factor to be used for noise estimation, and α_d is a constant. Finally, the noise power of each channel is estimated as follows:

$$D(\lambda, k) = \alpha_s(\lambda, k) D(\lambda - 1, k) + (1 - \alpha_s(\lambda, k))\left|Y(\lambda, k)\right|^2 \tag{8}$$

Gain Function for Noise Reduction

Using the estimated noise power, the SNR can be estimated according to

$$SNR(\lambda, k) = \delta SNR(\lambda - 1, k) + (1 - \delta)\left(\frac{P(\lambda, k)}{D(\lambda, k)} - 1\right) \tag{9}$$

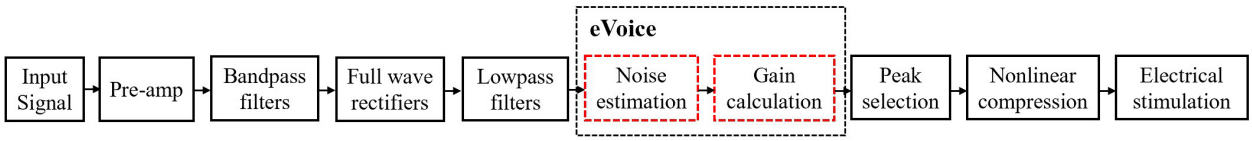

FIGURE 1 | Block diagram of the APS strategy (black) and eVoice (red).

Then, we use a gain function like:

$$G(\lambda, k) = \frac{SNR(\lambda, k)}{SNR(\lambda, k) + 1} \tag{10}$$

To suppress the noise to the maximum extent, the gain can be further adjusted:

$$G_0(\lambda, k) = \begin{cases} g & , \ G(\lambda, k) < T_g \\ G(\lambda, k) & , \ G(\lambda, k) \geq T_g \end{cases} \tag{11}$$

where g is a minor constant value, and T_g is a dynamic threshold determined by SNR. T_g is also one of the key factors that determine algorithm sensitivity.

Finally, the signal power after noise reduction is as follows:

$$S(\lambda, k) = G_0(\lambda, k) \ P(\lambda, k) \tag{12}$$

Example

An example of eVoice working in a speech-shaped noise (SSN) at +5 dB SNR is shown in **Figure 2**. eVoice was implemented with the APS coding strategy with a channel selection of 8-of-24 at a sampling rate of 16,000 Hz. **Figure 2** shows the power comparison in the eighth channel, including the signals for clean speech, noisy speech, processed speech, and estimated noise plotted in different colors.

EXPERIMENT 1: SUBJECTIVE PREFERENCE AND SPEECH RECOGNITION IN NOISE

This experiment was designed to evaluate speech intelligibility with eVoice (denoted by "NR1") compared with another NRA (denoted by "NR2") that used a binary masking for noise reduction, as well as the APS strategy with no NRA (denoted by "APS"). NR2 uses the same noise estimation method with NR1 as described in *Noise Estimation*. After noise estimation, NR2 calculates an SNR that is used to set the gain. That is, if the SNR is higher than a threshold, set the gain to 1 (speech dominant), or a small constant if lower (noise dominant). NR2 was selected for comparison because it is as computationally effective as eVoice and the method of ideal binary masking had been studied in other CI systems (Mauger et al., 2012b). Speech intelligibility was measured with a speech-in-noise recognition test and a subjective rating questionnaire.

Methods

Participants

This experiment involved 11 experienced CI users (six females and five males), aged from 20 to 59 years (mean age = 41.2 years). All were postlingually deafened adults unilaterally implanted with a CS-10A implant and using a Venus^{TM} sound processor (i.e., first generation) programed with the APS strategy. The Enduro^{TM} sound processor was fitted instead of the Venus^{TM} in this experiment. There is an option in a remote control to select whether to use an NRA (one with NR1-eVoice and the other one with NR2-Binary Masking). Demographics for individual participants are presented in **Table 1**. All participants' native language was Mandarin Chinese, and participants were paid for their time and traveling expenses. Written informed consent was obtained before the experiment, and all procedures were approved by the local institution's ethical review board.

Procedures and Materials

In this experiment, NR1 and NR2 performances were assessed first in a subjective evaluation, followed by a speech-in-noise recognition test.

The subjective evaluation lasted for 2 weeks. At the beginning of week 1, participants were fitted with an Enduro^{TM} processor that was incorporated with the NR1 and were asked to have a take-home trial for 1 week. During that week, participants were free to turn the NR1 on and off and use it in various everyday listening scenarios. At the end of week 1, subjective ratings were collected using the questionnaire shown in **Table 2**. Similar procedures were followed for the NR2 in week 2. The questionnaire consists of eight questions that cover various everyday listening scenarios. A 5-point rating scale was used to collect participants' subjective ratings of the NR1 or NR2 in each listening scenario after each 1-week take-home use: 2, strongly agree; 1, agree; 0, neutral; -1, disagree; -2, strongly disagree.

In the test of speech recognition in noise, we used two noise types (an SSN and a babble noise) at three SNRs (5, 10, and 15 dB) to compare the three algorithms (APS, NR1, and NR2). This yielded a total of 21 test blocks (two noise types × three SNRs × three algorithms + baselines of the three algorithms in quiet). The three baseline blocks (three algorithms in quiet) were conducted first in a random order, followed by the remaining 18 blocks in a random order. We used sentence materials from two published Mandarin speech databases: the PLA General Hospital sentence recognition test (Xi et al., 2012) corpus and the House Research Institute sentence recognition test (Fu et al., 2011) corpus. The PLA General

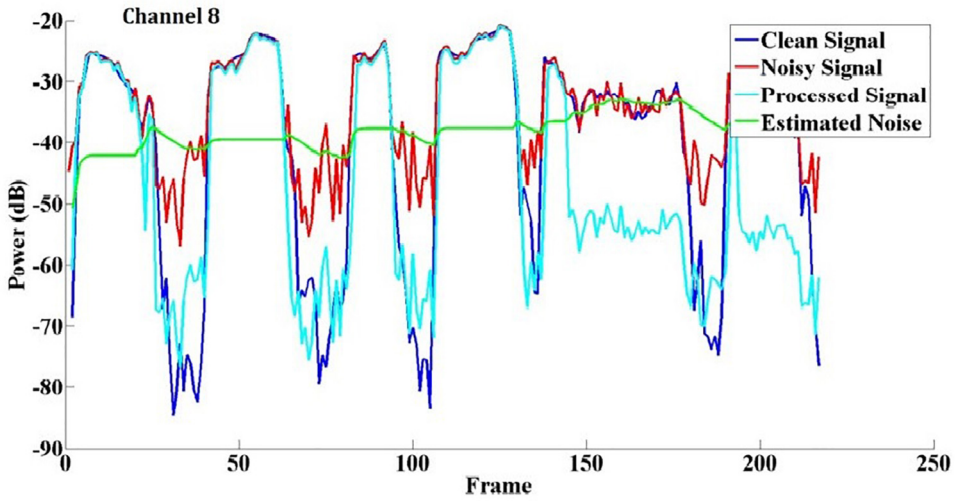

FIGURE 2 | An example of noise reduction at channel 8 in SSN at +5 dB SNR. The frame shift is 8 ms.

TABLE 1 | Demographic details of participants in experiment 1.

Participant	Age range at testing	Etiology	Profound deafness duration at implanted side (years)	Implanted side	CI experience (years)
N1	45–49	Sudden deafness	10	L	2
N2	25–29	Unknown	2	R	2
N3	20–24	Drug induced	17	R	4
N4	45–49	Unknown	4	R	2
N5	45–49	Drug induced	35	R	2
N6	45–49	Drug induced	28	R	4
N7	35–39	Drug induced	15	L	6
N8	40–44	Sudden deafness	0.25	L	1.5
N9	55–59	Sudden deafness	4	L	3
N10	40–45	Sudden deafness	8	R	6
N11	40–45	Drug induced	17	R	6

Abbreviations: CI, cochlear implant; F, female; L, left; M, male; R, right.

Hospital corpus consists of 12 lists each with 11 sentences, and each sentence includes six to eight key words. The House Research Institute corpus comprises 10 lists each with 10 phonetically balanced sentences, and each sentence contains seven words. All sentences were read by female speakers. Eleven of the 12 lists in the 301 corpus and all lists in the House corpus were used.

Because of the limited number of material lists, different lists from the PLA General Hospital and House Research Institute corpora were randomly assigned to blocks for each participant, with one list for each block. Special care was taken to ensure that the blocks of each algorithm used lists from the same corpus. In each block, sentences were presented in a random order, and a percentage word correctness score was calculated. Stimuli were presented in a soundproof room by a speaker located 1 m in front of the participant at a comfortable level (approximately 65 dBA). The tests were administered using QuickSTAR4TR software developed by Qianjie Fu (Emily Fu Foundation, 2019).

Statistical Analysis

Repeated-measures one-way analysis of variance (ANOVA) was used to analyze speech recognition in quiet. Repeated-measures three-way ANOVAs were performed to assess speech recognition in noise. Bonferroni adjustments were used for multiple comparisons.

Results

Subjective Evaluation Test

Figure 3 shows the results of the subjective ratings for NR1 and NR2.

For NR1 (i.e., eVoice), there were many positive ratings and few negative ones. Most participants gave positive ratings to Q2, Q4, Q5, and Q6, which indicated better listening experience with NR1 on than off in scenarios such as multitalker communication, at an intersection, and in a vehicle. For Q3, Q7, and Q8, most participants had neutral ratings, which corresponded to scenarios such as in a restaurant or supermarket and near an air conditioner

or fan. This result suggested comparable performance between NR1 on and off in these settings. There were a few participants who give positive ratings to Q3, Q7, and Q8 (better experience with NR1 turned on in listening scenarios such as a one-on-one conversation in a restaurant, by an air conditioner or fan, or in a busy supermarket). For listening in quiet, most participants reported that NR1 had no effect on a one-on-one conversation in quiet and gave positive ratings to Q1 (the NRA had no effect on one-on-one conversations in quiet rooms).

For NR2 (i.e., binary masking), the feedback was more variable. In general, ratings were almost evenly distributed between negative and positive for all eight questions except Q2 and Q8, which means that there were participants who thought NR2 was helpful in most listening scenarios. However, comparable numbers of participants thought it was not helpful or were neutral. For Q2 and Q8, most participants gave neutral ratings, which indicate that most thought the NR2 had no effect for multitalker communication in quiet or a one-on-one conversation in a supermarket.

TABLE 2 | Questionnaire used for subjective evaluation.

Q1	The NRA has no effect on one-on-one conversations in quiet rooms
Q2	The NRA helps during multitalker (at least three talkers) conversations in quiet rooms
Q3	The NRA helps during one-on-one conversations in restaurants
Q4	The NRA noticeably suppresses noise or helps to converse with others when vehicles pass
Q5	The NRA helps during one-on-one conversations or yields clearer station announcements inside a crowded bus
Q6	The NRA helps during one-on-one conversations or provides clearer radio sound inside a car
Q7	The NRA helps during one-on-one conversations near an air conditioner or fan
Q8	The NRA helps during one-on-one conversations in a busy supermarket

NRA refers to the noise-reduction algorithm to be evaluated (i.e., NR1 in week 1 or NR2 in week 2).

Speech Intelligibility Test

Results of speech recognition in quiet are shown in **Figure 4**. A repeated-measures one-way ANOVA revealed no significant difference among the mean results (∼90%) of the three algorithms ($p = 0.452$).

Figure 5 shows the results of speech recognition in the SSN and babble noise. Statistical significance was determined using ANOVA with the percent correct scores as the dependent variable and the noise type (SSN or babble), SNR (5, 10, or 15 dB), and algorithm (APS, NR1, or NR2) as within-subject factors. Tests of within-subjects effects indicated a significant effect of noise type ($p = 0.022$), SNR ($p < 0.001$), and algorithm ($p = 0.002$), as well as significant interactions between noise type and SNR ($p < 0.001$). Pairwise comparisons revealed that the overall performance of NR1 was significantly better than APS ($p = 0.001$) and NR2 ($p = 0.016$), and there was no significant difference between APS and NR2 ($p = 0.612$). When noise type and SNR were fixed to determine the effect of algorithms at specific SNRs in a particular noise type, NR1 performed significantly better than NR2 at the 5-dB SNR in the SSN ($p = 0.010$) and also significantly better than APS ($p = 0.027$) at the 5-dB SNR in the babble

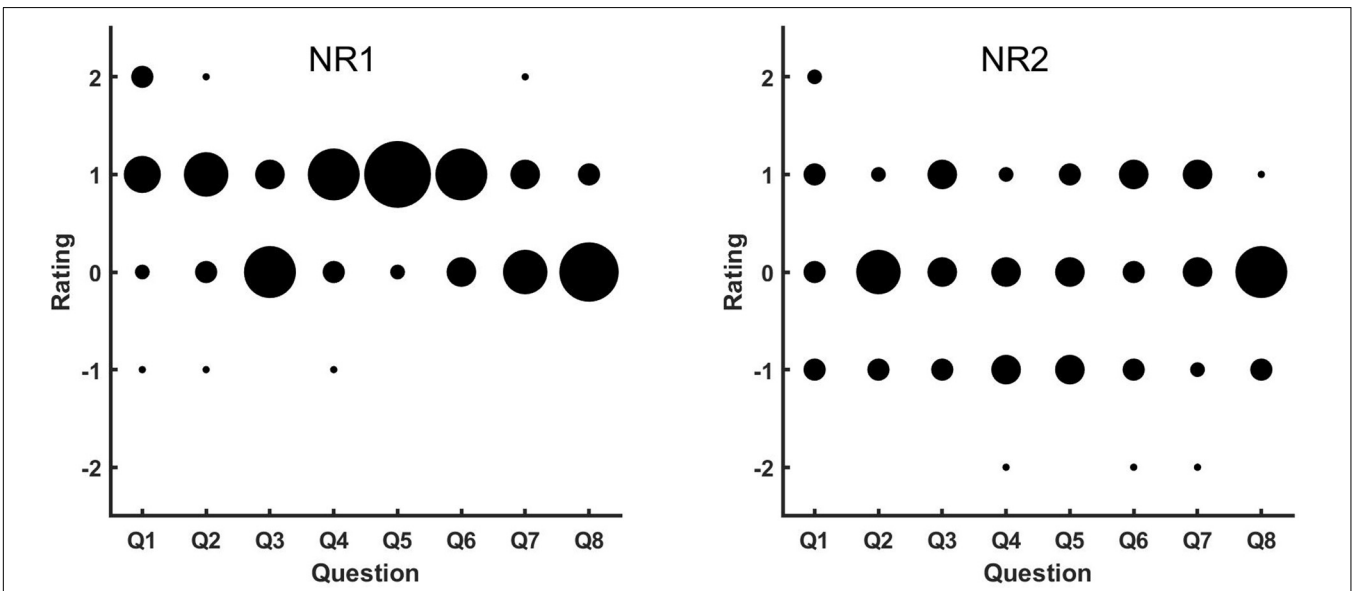

FIGURE 3 | Results of subjective evaluations of NR1 **(left panel)** and NR2 **(right panel)**. The abscissa lists all eight questions used for the subjective evaluation, and the ordinate is the rating given by the participants. Along the ordinate, "-2" represents strong disagreement on the question, and "2" represents strong agreement. The larger the number, the more positive the subjective evaluation is that the NR could help in different noisy scenarios and did not impact listening in quiet settings. The size of the circles represents the number of participants who gave the corresponding ratings, with larger circles indicating more participants.

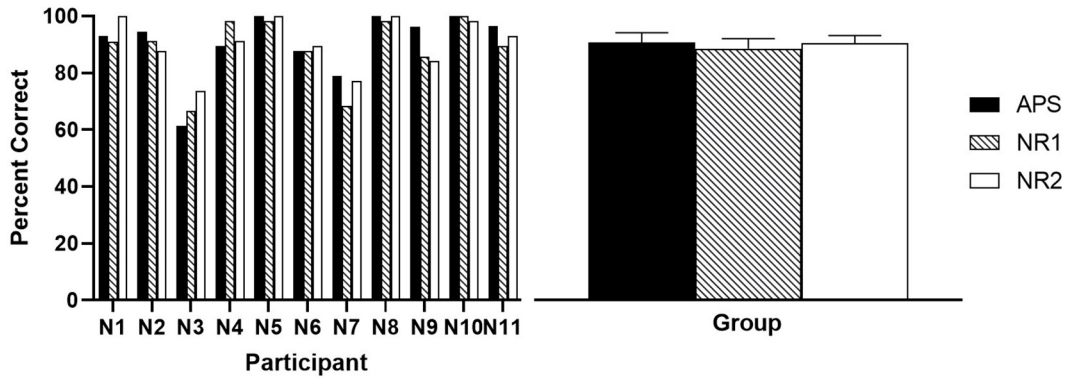

FIGURE 4 | Results of speech recognition score in quiet settings. The **left panel** shows the individual percent correct scores, and the **right panel** shows the group means, with error bars indicating the standard error of group means.

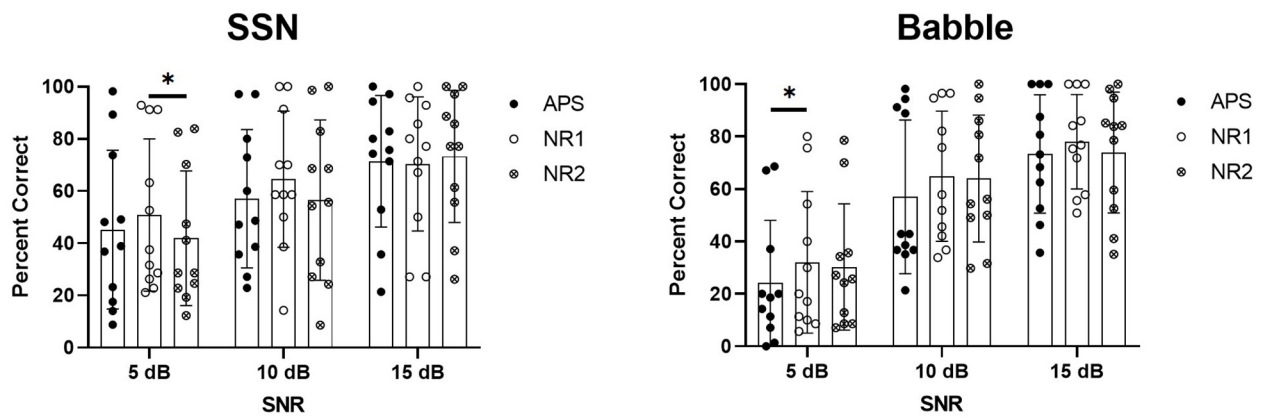

FIGURE 5 | Results of speech recognition score in SSN **(left panel)** and babble noise **(right panel)**. Results of each individual participant are plotted, and the bars show the mean values, with error bars indicating the standard deviations.

noise. In both the SSN and babble noise at the SNRs of 10 and 15 dB, there were no significant differences among the three algorithms. However, higher mean scores of NR1 could be observed against APS and NR2 at the 10-dB SNR in SSN (nearly eight percentage points), as well as at the 10-dB SNR (eight percentage points higher than APS) and 15-dB SNR (∼5 percentage points higher than APS and NR2) in the babble noise, although these improvements were not statistically significant.

Short Summary

In this experiment, we tested two NRAs: eVoice (NR1) and another that used binary masking (NR2). Both use the same noise estimation process but differ in the noise cancelation process. NR1 uses a smoothing gain function, whereas NR2 uses a binary masking. The subjective evaluation ratings show that NR1 was positively reviewed, whereas ratings of NR2 were almost evenly distributed from negative to positive, with a slight dominance of neutral responses. The speech recognition test results indicate overall better performance of NR1 compared to NR2 and APS. However, a significant benefit was only found at 5-dB SNR. The above results demonstrate that NR1 had better performance than NR2 for both speech recognition tests and subjective evaluations.

EXPERIMENT 2: SPEECH RECEPTION THRESHOLD TEST

Rationale

The hypothesized significant benefit of eVoice was not always supported by the results of the first experiment. One reason may be from the fixed SNR procedure and large performance variance in the cohort. From the results of Experiment 1 (left panel in **Figure 5**), we noticed that the ceiling effect could be observed in some participants at the SNR of 15 dB, and the floor effect could be observed at the SNR of 5 dB. Speech perception in noise varied dramatically among participants, even at the same SNR in the same noise. This indicates a limit of testing percent correct scores at fixed SNRs because this type of test is not able to exclude potential ceiling and floor effects. To overcome this limitation, we designed Experiment 2, which used an adaptive

speech reception threshold (SRT) test to measure the potential benefits of eVoice.

In the first experiment, we found clearly that NR1 (i.e., eVoice) provided better performance than NR2 (i.e., the ideal binary one) in the subjective test, although little improvement was observed in the speech-in-noise recognition test. To further explore the potential of eVoice and to save experiment time, only NR1 was evaluated in the second experiment.

Methods

Participants

Eight experienced CI users were recruited for this experiment (five females and three males, aged from 23 to 62 years with a mean of 43.6 years). All spoke Mandarin Chinese as their native language. They were all postlingually deafened adults unilaterally implanted with a CS-10A implant and used Enduro™ devices as their clinical processors, programed with the APS coding strategy with a remote control option to switch eVoice on or off. Demographic data for individual participants are presented in **Table 3**. Participants were compensated for their time and traveling expenses. All provided informed consent before the experiment, and all procedures were approved by the local institution's ethical review board.

Materials and Procedures

An adaptive staircase SRT in noise test was administered to further evaluate the performance of eVoice. This SRT measurement method was adopted from our previous studies (Meng et al., 2016, 2019) with two minor changes: (1) the stimulus presentation time was reduced from three at most to two at most, and (2) the correctness judgment threshold was changed from 50% words in a sentence to 80% words. The first was done to reduce experiment time. The second was for tracking a higher threshold, which is more indicative for a true understanding. Therefore, we were actually tracking a threshold around which the subjects have a 50% chance to obtain 80% correctness.

The Mandarin Hearing in Noise Test (MHINT) corpus (Wong et al., 2007) recorded by a single male speaker was used. There are 12 lists for formal tests and 2 lists for practice, with 20 sentences in each list, and 10 words in each sentence. In this experiment, 10 of 12 formal test lists were used as target speech in the formal tests, and both practice lists were used in the training stage to familiarize participants with the test procedures.

The SRTs for each condition with and without eVoice were tested. For each condition, two types of background noise were used: SSN and babble noise, which were generated using the method described in section "Speech Stimuli and Tasks" of Experiment 2 in Meng et al. (2019). The SRT for each condition–background combination was tested twice using two different MHINT lists, and the results were averaged between the two lists as the final SRT. Speech intelligibility for each condition in quiet was also measured using one MHINT list. Therefore, a total of 10 lists were used for testing (two backgrounds × two conditions × two lists per combination + two lists for speech intelligibility in quiet). The order of lists and conditions was randomized across participants. Prior to the formal test, two practice lists were used to familiarize participants with the test procedures of the SRT and the speech intelligibility in quiet test. During the test, each sentence was presented at most twice on the request of the participants; participants were instructed to repeat words that form a sentence with a meaning, and no feedback was given.

The SNR in each trial was adapted by changing the level of target speech with fixed background noise. Participants were instructed to repeat as many words as they could, and the target level was decreased if no less than eight of the words were repeated correctly; otherwise, the target level was increased. The step size was 8 dB before the second reversal, followed by 4 dB before the fourth reversal and 2 dB for the remaining reversals. The arithmetic mean of the SNRs of the last eight sentences was calculated and recorded as the final SRT.

It is worth mentioning that the babble noise used in this study consisted of voices of the same talker as the target speech (Meng et al., 2019), which is extremely challenging for any NRA. Additional information about the procedures and materials can be obtained from Meng et al. (2016, 2019).

Results

The eight CI users listed in **Table 3** participated in this experiment, but N17 was found to have auditory neuropathy. Therefore, N17 data were excluded from the analyses.

Results of speech recognition in quiet are shown in **Figure 6**. The group mean scores were 93.1 and 93.3% for eVoice-off and eVoice-on, respectively. A two-tailed paired-samples t-test showed no significant difference between the two conditions ($t(6) = -0.162$, $p = 0.877$).

Figure 7 shows the results of the SRTs in the SSN (left panel) and babble noise (right panel). In the SSN, every participant had lower SRTs with eVoice-on than with eVoice-off. The group mean SRTs were 7.9 and 5.7 dB for eVoice-off and eVoice-on, respectively. This 2.2-dB difference was a statistically significant improvement ($t(6) = 6.892$, $p < 0.001$).

In the babble noise, group mean SRTs of 10.9 and 10.7 dB were observed for eVoice-off and eVoice-on, respectively. A two-tailed paired-samples t test revealed no significant difference between the two conditions ($t(6) = 0.249$, $p = 0.812$).

Short Summary

The aim of this experiment was to quantify the benefit introduced by eVoice for speech intelligibility and exclude potential ceiling and floor effects. Speech intelligibility was measured using an adaptive SRT test with two different backgrounds: SSN and babble noise. There was no significant difference in speech recognition rates in quiet settings. This result indicates that eVoice would not affect speech perception in quiet. eVoice yielded an SRT decrease of 2.2 dB in SSN, whereas no significant effect was found in SRTs in babble noise.

TABLE 3 | Demographic details of participants in experiment 2.

Participant	Age range at testing	Etiology	Profound deafness duration at implanted side (years)	Implanted side	CI experience (years)	Enduro experience (years)
N12	20–24	Drug induced	1	R	6	1
N13	50–54	Unknown	0.4	L	6	0.5
N14	30–34	Sudden deafness	1	R	1.5	1
N15	45–49	Unknown	5	L	7	3
N16	50–54	LAVS	1	L	6	1
N9	60–64	Sudden deafness	4	L	6	3
N10	45–49	Sudden deafness	8	R	9	3
N17	35–39	Unknown	7	L	8	0.5

Abbreviations: CI, cochlear implant; F, female; L, left; LAVS, large vestibular aqueduct syndrome; M, male; R, right.

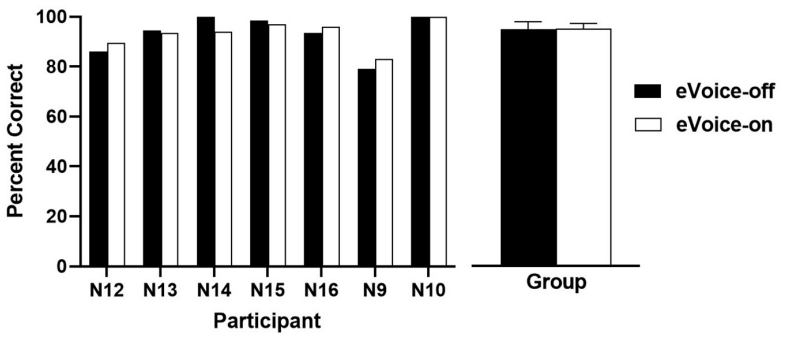

FIGURE 6 | The speech recognition scores in quiet with eVoice-off and eVoice-on. The **left panel** shows the individual scores, and the **right panel** shows the group means, with error bars showing the standard errors of group means.

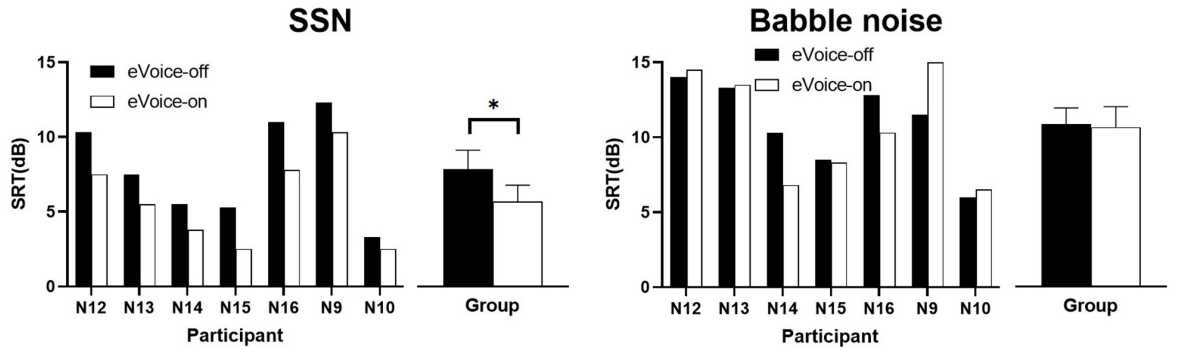

FIGURE 7 | Results of SRT in the SSN **(left panel)** and babble noise **(right panel)**. Individual SRTs are shown on the left, and the group mean SRTs are shown on the right. Error bars show the standard error of group means. The significant difference is illustrated by the asterisk ($p < 0.05$).

DISCUSSION

In this study, we examined eVoice, the first noise-suppression technique in Nurotron® CIs. eVoice is a single-channel NRA implemented within the APS strategy in the Enduro processor. Two experiments were conducted to evaluate this algorithm. First, the performance of eVoice was compared with another binary-masking method in a speech recognition test and also underwent a subjective evaluation in Experiment 1 (N = 11). The eVoice performed slightly better than the binary-masking NRA. Then, the more indicative adaptive SRT test was conducted to quantify the noise reduction effect of eVoice on speech intelligibility in Experiment 2 (N = 7). Comparing eVoice on and off, there was a 2.2-dB SRT benefit in stationary noise and no difference in quiet and non-stationary noise.

Compared to other single-channel NRAs implemented in CI strategies, eVoice has comparable performance with those reported in the literature. For example, a single-channel NRA implemented in the ACE strategy was found to have an SRT benefit of up to 2.14 dB in stationary noise (Dawson et al., 2011). The ClearVoice implemented in the HiRes 120 strategy used a time window of 1.3 s for noise estimation and yielded a percent

correct score increase of up to 24 percentage points (Buechner et al., 2010). This may translate to a 1.3- to 3.4-dB SRT decrease according to the literature that for typical speech materials, a 1-dB SRT decrease leads to 7- to 19-percentage-point increase in the percent correct score (Moore, 2007). However, significant benefits in non-stationary noise are seldom reported in the literature, which may indicate a limit of traditional single-channel NRAs. More advanced techniques should be developed to improve speech perception in non-stationary noise for CI users.

This article is significant from the implantees' and the audiologists' perspectives. For a new system with a quickly growing number of users, this report on eVoice is useful for understanding the system and the new noise reduction method. For a new NRA in CIs, two questions are of great concern to users and audiologists: (1) whether this NRA really works in various types of noises and (2) to what extent users can benefit from it. Our results demonstrate that eVoice can improve speech intelligibility in stationary noise and does not affect speech perception in quiet and non-stationary noise. This is because eVoice is a monaural SNR estimation–based algorithm that assumes that the noise is relatively stationary compared with speech. We found that some users of the Enduro processor might have not noticed the existence of this NRA, and their audiologists can advise or remind them to turn eVoice on to improve their speech perception performance in noise.

Another significant contribution of this article is to inspire people to rethink noise management for CI systems. Researchers should consider the assumptions about directionality and complex non-linear patterns that can be computationally modeled by signal processing or machine learning (e.g., Bianco et al., 2019; Gong et al., 2019). Previous studies and present work provide considerable support for optimizing and updating noise-suppression techniques to improve speech-in-noise recognition for CI users.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Medical Ethics Committee of Shenzhen University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors conceived and designed the analysis and wrote the manuscript. HZ and NW collected the data.

FUNDING

This work was supported by the National Natural Science Foundation of China (11704129, 11574090, and 61771320), Natural Science Foundation of Guangdong, China (2020A1515010386 and 2018B030311025), and Shenzhen Science and Innovation Funds (JCYJ 20170302145906843).

ACKNOWLEDGMENTS

We are grateful to all the CI participants for their patience and cooperation during this study. We would like to thank Wenhe Tu, Sui Huang, Peiyao Wang, Carol Peng, Shanxian Gao, and Jiaqi Zhang from Nurotron for their help during data collection. We also thank the two reviewers for helpful comments.

REFERENCES

Berenstein, C. K., Mens, L. H., Mulder, J. J., and Vanpoucke, F. J. (2008). Current steering and current focusing in cochlear implants: comparison of monopolar, tripolar, and virtual channel electrode configurations. *Ear. Hear.* 29, 250–260.

Bianco, M. J., Gerstoft, P., Traer, J., Ozanich, E., Roch, M. A., and Gannot, S. (2019). Machine learning in acoustics: theory and applications. *J. Acoust. Soc. Am.* 146, 3590–3628.

Bonham, B. H., and Litvak, L. M. (2008). Current focusing and steering: modeling, physiology, and psychophysics. . *Hear. Res.* 242, 141–153. doi: 10.1016/j.heares.2008.03.006

Buechner, A., Brendel, M., Saalfeld, H., Litvak, L., Frohne-Buechner, C., and Lenarz, T. (2010). Results of a pilot study with a signal enhancement algorithm for HiRes 120 cochlear implant users. *Otol. Neurotol.* 31, 1386–1390.

Buechner, A., Dyballa, K. H., Hehrmann, P., Fredelake, S., and Lenarz, T. (2014). Advanced beamformers for cochlear implant users: acute measurement of speech perception in challenging listening conditions. *PLoS One* 9:e95542. doi: 10.1371/journal.pone.0095542

Chen, F., Hu, Y., and Yuan, M. (2015). Evaluation of noise reduction methods for sentence recognition by mandarin-speaking cochlear implant listeners. *Ear Hear.* 36, 61–71.

Chung, K., Zeng, F. G., and Waltzman, S. (2004). Using hearing aid directional microphones and noise reduction algorithms to enhance cochlear implant performance. *Acoust. Res. Lett. Onl.* 5, 56–61.

Dawson, P. W., Mauger, S. J., and Hersbach, A. A. (2011). Clinical evaluation of signal-to-noise ratio–based noise reduction in nucleus cochlear implant recipients. *Ear. Hear.* 32, 382–390.

Dhanasingh, A., and Jolly, C. (2017). An overview of cochlear implant electrode array designs. *Hear. Res.* 356, 93–103. doi: 10.1016/j.heares.2017.10.005

Emily Fu Foundation. (2019). *MSP*. http://msp.emilyfufoundation.org/msp_about.html (accessed August 15, 2019).

Fu, Q. J., Zhu, M., and Wang, X. (2011). Development and validation of the mandarin speech perception test. *J. Acoust. Soc. Am.* 129, EL267–EL273. doi: 10.1121/1.3590739

Gao, N., Xu, X. D., Chi, F. L., Zeng, F. G., Fu, Q. J., Jia, X. H., et al. (2016). Objective and subjective evaluations of the nurotron venus cochlear implant

system via animal experiments and clinical trials. *Acta Otolaryngol.* 136, 68–77.

Goehring, T., Keshavarzi, M., Carlyon, R. P., and Moore, B. (2019). Using recurrent neural networks to improve the perception of speech in non-stationary noise by people with cochlear implants. *J. Acoust. Soc. Am.* 146:705. doi: 10.1121/1.5119226

Gong, S., Wang, Z., Sun, T., Zhang, Y., Smith, C. D., Xu, L., et al. (2019). "Dilated FCN: Listening Longer to Hear Better," in *2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Piscataway, NJ: IEEE, 254–258.

Hagen, R., Radeloff, A., Stark, T., Anderson, I., Nopp, P., Aschbacher, E., et al. (2019). Microphone directionality and wind noise reduction enhance speech perception in users of the MED-EL SONNET audio processor. *Cochlear Implants Int.* 21, 53–65. doi: 10.1080/14670100.2019.1664529

Hersbach, A. A., Arora, K., Mauger, S. J., and Dawson, P. W. (2012). Combining directional microphone and single-channel noise reduction algorithms: a clinical evaluation in difficult listening conditions with cochlear implant users. *Ear. Hear.* 33, e13–e23. doi: 10.1097/aud.0b013e31824b9e21

Holden, L. K., Brenner, C., Reeder, R. M., and Firszt, J. B. (2013). Postlingual adult performance in noise with HiRes 120 and clear voice low. Medium, and high. *Cochlear Implants Int.* 14, 276–286. doi: 10.1179/1754762813Y.0000000034

Hu, Y., Loizou, P. C., Li, N., and Kasturi, K. (2007). Use of a sigmoidal-shaped function for noise attenuation in cochlear implants. *J. Acoust. Soc. Am.* 122, EL128–EL134.

Jeschke, M., and Moser, T. (2015). Considering optogenetic stimulation for cochlear implants. *Hear. Res* 322, 224–234.

Kam, A. C. S., Ng, I. H. Y., Cheng, M. M. Y., Wong, T. K. C., and Tong, M. C. F. (2012). Evaluation of the clear voice strategy in adults using HiResolution Fidelity 120 sound processing. *Clin. Exp. Otorhinolaryngol.* 5(Suppl.1), S89–S92.

Kasturi, K., and Loizou, P. C. (2007). Use of S-shaped input-output functions for noise suppression in cochlear implants. *Ear. Hear.* 28, 402–411.

Kral, A., Dorman, M. F., and Wilson, B. S. (2019). Neuronal development of hearing and language: cochlear implants and critical periods. *Annu. Rev. Neurosci.* 42, 47–65.

Lai, Y. H., Tsao, Y., Lu, X., Chen, F., Su, Y. T., Chen, K. C., et al. (2018). Deep learning-based noise reduction approach to improve speech intelligibility for cochlear implant recipients. *Ear. Hear.* 39, 795–809.

Li, X., Nie, K., Imennov, N. S., Won, J. H., Drennan, W. R., Rubinstein, J. T., et al. (2012). Improved perception of speech in noise and Mandarin tones with acoustic simulations of harmonic coding for cochlear implants. *J. Acoust. Soc. Am.* 132, 3387–3398.

Loizou, P. C. (1999). Signal-processing techniques for cochlear implants. *IEEE Eng. Med. Biol. Mag.* 18, 34–46. doi: 10.1109/51.765187

Loizou, P. C. (2006). Speech processing in vocoder-centric cochlear implants. *Adv. Otorhinolaryngol.* 64, 109–143.

Loizou, P. C. (2007). *Speech Enhancement, Theory and Practice (Illustrated ed.)*. Boca Raton, FL: CRC Press.

Martin, R. (2001). Noise power spectral density estimation based on optimal smoothing and minimum statistics. *IEEE T. Audio Speech.* 9, 504–512.

Mauger, S. J., Arora, K., and Dawson, P. W. (2012a). Cochlear implant optimized noise reduction. *J. Neural Eng.* 9:065007. doi: 10.1088/1741-2560/9/6/065007

Mauger, S. J., Dawson, P. W., and Hersbach, A. A. (2012b). Perceptually optimized gain function for cochlear implant signal-to-noise ratio based noise reduction. *J. Acoust. Soc. Am.* 131, 327–336.

Mc Laughlin, M., Lu, T., Dimitrijevic, A., and Zeng, F. G. (2012). Towards a closed-loop cochlear implant system: application of embedded monitoring of peripheral and central neural activity. *IEEE Trans. Neural Syst. Rehabil. Eng.* 20, 443–454.

McAulay, R., and Malpass, M. (1980). Speech enhancement using a soft-decision noise suppression filter. *IEEE Trans. signal Process.* 28, 137–145.

Meng, Q., Wang, X., Cai, Y., Kong, F., Buck, A. N., Yu, G., et al. (2019). Time-compression thresholds for mandarin sentences in normal-hearing and cochlear implant listeners. *Hear. Res.* 374, 58–68.

Meng, Q., Zheng, N., and Li, X. (2016). Mandarin speech-in-noise and tone recognition using vocoder simulations of the temporal limits encoder for cochlear implants. *J. Acoust. Soc. Am.* 139, 301–310.

Middlebrooks, J. C., and Snyder, R. L. (2007). Auditory prosthesis with a penetrating nerve array. *J. Assoc. Res. Otolaryngol.* 8, 258–279.

Moberly, A. C., and Reed, J. (2019). Making sense of sentences: top-down processing of speech by adult cochlear implant users. *J. Speech Lang. Hear. Res.* 62, 2895–2905.

Moore, B. C. J. (2007). *Cochlear Hearing Loss*, 2nd Edn. Chichester: Wiley.

Nie, K., Stickney, G., and Zeng, F. (2005). Encoding frequency modulation to improve cochlear implant performance in noise. *IEEE Trans. Biomed. Eng* 52, 64–73. doi: 10.1109/tbme.2004.839799

Nogueira, W., Nagathil, A., and Martin, R. (2018). Making music more accessible for cochlear implant listeners: recent developments. *IEEE Signal Proc. Mag.* 36, 115–127.

Ping, L., Wang, N., Tang, G., Lu, T., Yin, L., Tu, W., et al. (2017). Implementation and preliminary evaluation of 'C-tone': a novel algorithm to improve lexical tone recognition in Mandarin-speaking cochlear implant users. *Cochlear Implants Int.* 18, 240–249.

Rangachari, S., and Loizou, P. C. (2006). A noise-estimation algorithm for highly non-stationary environments. *Speech Commun.* 48, 220–231.

Rebscher, S., Zhou, D. D., and Zeng, F. G. (2018). Development and clinical introduction of the nurotron cochlear implant electrode array. *J. Int. Adv. Otol.* 14, 392–400.

Ren, C., Yang, J., Zha, D., Lin, Y., Liu, H., Kong, Y., et al. (2018). Spoken word recognition in noise in Mandarin-speaking pediatric cochlear implant users. *Int. J. Pediatr. Otorhinolaryngol.* 113, 124–130. doi: 10.1016/j.ijporl.2018.07.039

Rubinstein, J. T. (2004). How cochlear implants encode speech. *Curr. Opin. Otolaryngol. Head Neck Surg.* 12, 444–448.

Shannon, R. V. (2014). "Adventures in Bionic Hearing," in *Perspectives on Auditory Research, Springer Handbook of Auditory Research*, Vol. 50, eds A. Popper and R. Fay (New York, NY: Springer).

Tamati, T. N., Ray, C., Vasil, K. J., Pisoni, D. B., and Moberly, A. C. (2019). High- and low-performing adult cochlear implant users on high-variability sentence recognition: differences in auditory spectral resolution and neurocognitive functioning. *J. Am. Acad. Audiol.* doi: 10.3766/jaaa18106

Van Hoesel, R. (2007). Peak-derived timing stimulation strategy for a multi-channel cochlear implant. *J. Acoust. Soc. Am.* 123:4036. doi: 10.1121/1.2942439

Wang, D., and Hansen, J. (2018). Speech enhancement for cochlear implant recipients. *J. Acoust. Soc. Am.* 143:2244. doi: 10.1121/1.5031112

Wang, N., Tang, G., and Fu, Q. J. (2017). "Implementation and evaluation of a single-cannel noise reduction method in cochlear implants," in *2017 Conference on Implantable Auditory Prostheses*, Lake Tahoe, CA.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature.* 352, 236–238.

Wong, L. L., Soli, S. D., Liu, S., Han, N., and Huang, M. W. (2007). Development of the mandarin hearing in noise test (MHINT). *Ear. Hear.* 28, 70S–74S. doi: 10.1121/1.3590739

Wouters, J., McDermott, H. J., and Francart, T. (2015). Sound coding in cochlear implants: From electric pulses to hearing. *IEEE Signal Proc. Mag.* 32, 67–80.

Xi, X., Ching, T. Y., Ji, F., Zhao, Y., Li, J. N., Seymour, J., et al. (2012). Development of a corpus of mandarin sentences in babble with homogeneity optimized via psychometric evaluation. *Int. J. Audiol.* 51, 399–404. doi: 10.3109/14992027.2011.642011

Xu, Y., Luo, C., Zeng, F. G., Middlebrooks, J. C., Lin, H. W., and You, Z. (2018). Design, fabrication, and evaluation of a parylene thin-film electrode array for

cochlear implants. *IEEE Trans. Biomed. Eng.* 66, 573–583. doi: 10.1109/tbme.2018.2850753

Yang, L. P., and Fu, Q. J. (2005). Spectral subtraction-based speech enhancement for cochlear implant patients in background noise. *J. Acoust. Soc. Am.* 117, 1001–1004.

Zeng, F. G. (2004). Trends in Cochlear implants. *Trends Amlif* 8, 1–34.

Zeng, F. G., Rebscher, S., Harrison, W., Sun, X., and Feng, H. (2008). Cochlear implants: system design, integration, and evaluation. *IEEE Rev. Biomed. Eng.* 1, 115–142. doi: 10.1109/rbme.2008.2008250

Zeng, F. G., Rebscher, S. J., Fu, Q. J., Chen, H., Sun, X., Yin, L., et al. (2015). Development and evaluation of the Nurotron 26-electrode cochlear implant system. *Hear. Res.* 322, 188–199.

Zierhofer, C. M. (2003). *Electrical Nerve Stimulation Based on Channel Specific Sampling Sequences*. Patent No. EP1207938B1. Munich: European Patent Office.

Conflict of Interest: Nurotron provided some compensation to the subjects and accommodation to HZ during Experiment 2. NW was employed by the company Nurotron Biotechnology Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Zhou, Wang, Zheng, Yu and Meng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.