Biomedical Signal Processing and Control 79 (2023) 104169

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Effects of number of maxima and electrical dynamic range on speech-in-noise perception with an "*n*-of-*m*" cochlear-implant strategy

Yefei Mo [a,d,1], Huali Zhou [b,c,1], Fanhui Kong [b], Zhifeng Liu [d], Xiaohong Liu [d], Hongming Huang [d], Yan Huang [d], Nengheng Zheng [b], Qinglin Meng [c,*], Peina Wu [d,a,*]

[a] School of Medicine, South China University of Technology, Guangzhou, Guangdong 510006, PR China
[b] College of Electronics and Information Engineering, Shenzhen University, Shenzhen, Guangdong 518060, PR China
[c] Acoustic Laboratory, School of Physics and Optoelectronics, South China University of Technology, Guangzhou, Guangdong 510610, PR China
[d] Department of Otolaryngology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, Guangdong 510080, PR China

ARTICLE INFO

Keywords:
Cochlear implant
Vocoder simulation
Electrical dynamic range
Number of maxima
Speech-in-noise perception

ABSTRACT

Cochlear implant (CI) recipients face great challenges in speech-in-noise recognition, partially due to the fact that only temporal envelopes from a limited number of bands are preserved in most CI signal processing strategies. In "*n*-of-*m*" strategies (e.g., the Advanced Combinational Encoder, ACE), the number of maxima (n_{max}) and electrical dynamic range (EDR) are two essential parameters that may affect the envelope representation and further influence speech perception. Speech recognition can be improved by optimizing parameter settings in CI programming. To investigate the effects of n_{max} and EDR on speech-in-noise perception, Mandarin speech reception thresholds (SRTs) in babble noise were measured in CI recipients using ACE. The n_{max} was set to 2, 4, 6, 8, and 16. The EDR was set to the base EDR (i.e., participants' clinical EDR) and 50 % EDR (i.e., 50 %-compressed base EDR). Results showed that: 1) there was no significant interaction effect between n_{max} and EDR, 2) SRTs with n_{max} = 2, 4, and 16 were significantly higher (or worse) than those with moderate n_{max} (6–8), 3) narrower EDRs significantly lead to higher SRTs. Simulation experiments using a Gaussian-Enveloped Tones Vocoder in normal-hearing listeners were also conducted and provided both supportive and additional observations to the CI results. This study suggests that, in CI programming, n_{max} and EDR are two independent influencing factors. Large n_{max} (e.g., 16) is not recommended as it may harm speech intelligibility in noisy environments, and inaccurate EDR measurements should be avoided.

1. Introduction

Hearing loss, a common sensory disorder, has become a global health problem due to its increasing prevalence and adverse impact on quality of life. According to the World Health Organization, 60.5 million people worldwide suffer from severe or higher grades of hearing loss [1]. Cochlear implants (CIs) are the most successful sensory neuroprosthesis, having restored functional hearing to 1 million people with severe-to-profound sensorineural hearing loss [2]. Sound coding strategies enable CIs to replace the normal cochlea and perform acoustic-electric conversions to directly stimulate the auditory nerve. Most current CI sound coding strategies extract temporal envelopes from a limited number of bands, and deliver the further compressed and quantified envelopes by the amplitude modulation of electric current pulse trains. As a result, CIs provides only coarse spectral information. Due to the high redundancy of speech, many CI users reach close-to-normal speech intelligibility in quiet environments [3,4]. However, they exhibit considerable difficulty understanding speech in noisy environments [5,6], which is detrimental to their quality of life.

Improving speech understanding in noise is often the top priority for CI recipients and there are two available signal processing approaches to tackle the noise issue. Firstly, an intuitive approach is developing pre-processing speech enhancement algorithms to increase signal-to-noise ratio (SNR) before the core CI strategy processing. These include

* Corresponding authors at: Department of Otolaryngology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, Guangdong 510080, PR China (Peina Wu) Acoustics Laboratory, School of Physics and Optoelectronics, South China University of Technology, Guangzhou, Guangdong 510610, PR China (Qinglin Meng).

E-mail address: linwupeina@hotmail.com (P. Wu).

[1] Yefei Mo and Huali Zhou contributed equally to this work and should be considered co-first authors.

https://doi.org/10.1016/j.bspc.2022.104169
Received 9 March 2022; Received in revised form 18 August 2022; Accepted 4 September 2022
Available online 13 September 2022
1746-8094/

single-channel noise reduction and multi-microphone processing for spatial filtering (i.e., "directional microphones" or "beamformers") [7,8]. In various biomedical engineering applications, in order to improve the SNR, a large variety of filters have been applied to extract signals hidden in noise or reduce noise [9,10]. Secondly, the core CI strategy could also be improved by parameter modification or novel signal processing design to enlarge the acoustic cue differences between signal and noise in the electric stimuli [11,12]. This is inspired by the fact in normal-hearing (NH) people that acoustic cue differences in voice characteristics and spatial locations between signal and noise can lead to significant masking release in noisy environments. Our study belongs to the second type and is concerned with the effects of two parameters on the speech-in-noise perception of CI users.

Signal coding strategy is the core of CIs to achieve acoustic-electric conversion, determining the output signal expression. The most typical CI strategy, i.e., the Continuous Interleaved Sampling (CIS) strategy, is purely based on temporal envelopes, and has all channels (electrodes) active with stimulation in each processing frame, while "n-of-m" strategies dynamically pick n channels with the highest energy for stimulation in each frame from all m physically available ones. The number of maxima (n_{max}, i.e., n) is equal to m in CIS and smaller than m in "n-of-m" strategies. For example, a default setting of 8-of-22 is used in the Advanced Combinational Encoder (ACE) strategy, implemented in the Nucleus® CI systems (Cochlear Limited, Australia) [13]. "n-of-m" strategies were proposed to enable faster stimulation rate, reduce inter-channel interaction, and increase battery life [14]. Previous research suggests that most subjects have optimal speech recognition performance when n_{max} is at least 9 [14,15]. However, it remains to be investigated whether a larger n_{max} is beneficial to provide sufficient detail for a speech signal embedded in competing noise.

Electrical dynamic range (EDR) is another parameter in all contemporary CI strategies that can affect the temporal envelope expression. EDR is defined as the difference between the CI user's perceptual threshold level (T level) and most comfortable level (C level) for electrical stimulation. The acoustic envelope amplitude is mapped to an electrical pulse amplitude using a loudness growth function in CI sound coding strategies [16,17]. The pulse amplitude is related to the loudness of the sound, i.e., increasing the current amplitude results in a louder sound percept. In most clinical fittings, the EDRs are comparable among electrodes for individual ears but variable among ears. Wider EDRs lead to deeper modulation both in the spectral and temporal envelopes.

Both parameters of n_{max} and EDR have influences on the electric representation of spectral modulation in the ACE stimuli. To the best of our knowledge, the possible joint or interactive effects of n_{max} and EDR on speech-in-noise perception have not been studied yet. This study aims at investigating the effect of n_{max} and EDR on speech-in-noise perception for CI users. Within-subject experiments were conducted with actual CI users using Cochlear ltd devices and with simulated CI users using a pulsatile vocoder, both using the ACE strategy. The vocoder simulation provided more opportunities to examine the effects of the two target parameters.

2. Related works

The effect of "m" (i.e., the number of total channels) on speech-in-noise perception has been extensively studied [18–21]. Significantly affecting speech perception, spectral envelopes represented by intensity variations (or modulations) among electrodes are key acoustic cues for phoneme identification. Previous studies investigating the effect of the number of total channels in the CIS strategy have found that speech perception in noise improved with increasing channel number, but typically saturated up to 8–10 channels [18–20]. More recently, Croghan et al. (2017) demonstrated that speech-in-noise perception might not saturate at 8–10 channels, but continues to improve beyond 8 to 22 channels for contemporary CI recipients, particularly those with good spectro-temporal resolution [21].

However, there are limited studies investigating the effect of n_{max} (i.e., "n") on speech-in-noise perception. Upper-bound performance of sentence recognition has been observed in a "n-of-20" strategy when n was increased to 9 for NH participants listening to vocoded speech [15,22], and also in the ACE strategy when n_{max} was increased to 8 for CI recipients [23]. These studies were conducted about two decades ago and many aspects have changed, e.g., electrode design [24], surgical techniques [25], as well as cochlear implantation candidacy criteria [26]. These changes might impact the performance of CI users on speech-in-noise perception with different n_{max}. For example, the recent work of Berg *et al.* (2019) found that increasing n_{max} from 8 to 16 resulted in improved monosyllabic word recognition and sentence recognition in noise in perimodiolar electrode recipients [27]. Their findings suggested that CIs may be able to harness spectral information from more electrodes. Therefore, it is worth revisiting the effect of n_{max} on speech intelligibility in CI users. A larger n_{max} may activate more electrodes with lower signal amplitudes. This may enrich the spectral envelope, which may be useful for speech intelligibility.

The effect of EDR was studied in a between-subject protocol in some previous studies, and the magnitude of the EDR was found to be weakly associated with speech recognition [28,29]. This weak association may be due to the high individual variability of actual CI recipients. Considering the difficulty to examine its independent effects with a practical size of the participant cohort, some within-subject experiments were also performed. They observed a small but significant improvement in speech-in-noise recognition when EDRs were compressed, i.e., increasing the T level higher than measured perceptual threshold in all channels [30] and increasing T level (plus 5 % EDR) of specific electrodes with a poor modulation detection threshold at 50 % the EDR [31]. This may be attributed to the fact that when the T level is increased, soft sounds are mapped to higher current levels, so that the audibility of soft sounds is improved. However, several studies have demonstrated that inappropriate EDR compression exerted to a significant detriment in speech perception [16,32–34]. The effect of EDR in actual CI users is still to be explored.

The previous researches done so far have one common denominator, n_{max} and EDR can affect speech perception in CI recipients, but there is still no consensus on their optimal settings and potential joint or interaction effects on speech-in-noise perception. We hypothesized that larger nmax and wider EDR would yield better performance. To test this hypothesis, speech-in-noise performance was measured with different settings of n_{max} and EDR. Based on previous studies and the experience of our hearing center, n_{max}, which is often set to 8 by default, was set from 2 to 16 in this study. In contrast, true T levels and C levels for individual patients are generally measured accurately in a standard psychophysical procedure, and cannot be set to arbitrary values like n_{max}. Therefore, increased T levels were set resulting in narrower EDRs for CI users. These artificially increased T levels could not represent a patient with the same true T levels, but they allow us to get some implications for true narrower EDRs, and to observe the influence of poorly-fitted T levels which do possibly happen.

3. Material and methods

3.1. Study participants

Twelve adult Nucleus® CI recipients (seven males and five females, aged 29.5 ± 9.4 years) and ten normal-hearing (NH) listeners (three males and seven females, aged 24.1 ± 2.3 years) were recruited. Inclusion criteria for CI users included at least 1 year of CI experience, use of the ACE strategy, 22 active electrodes, and severe-to-profound sensorineural hearing loss in both ears (pure tone thresholds for 0.5, 1, 2, and 4 kHz > 65 dB HL; see Fig. 1) without hearing device intervention. One CI participant was unable to complete the task and was excluded. Table 1 provides demographic information for the CI group. All NH listeners are native Mandarin speakers with pure-tone thresholds

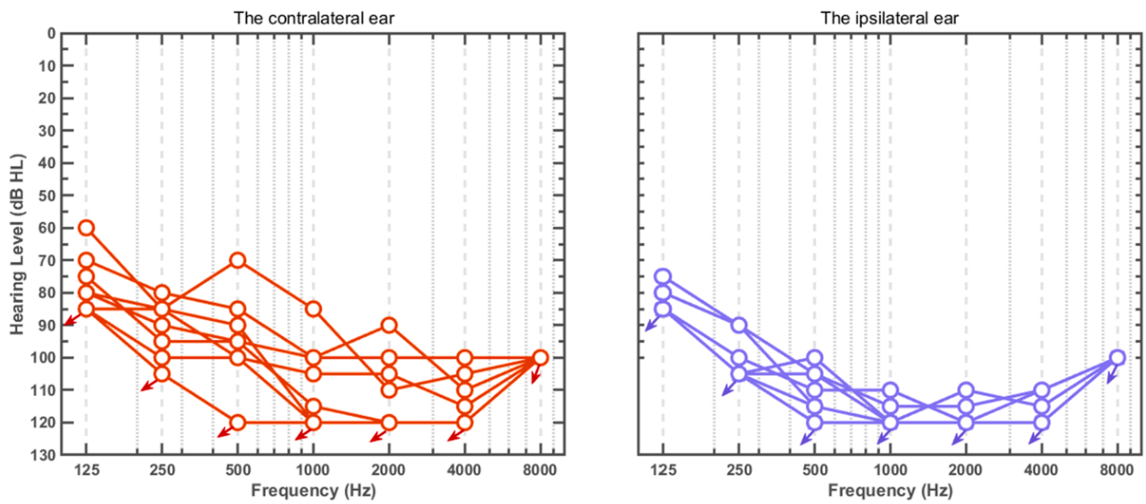

Fig. 1. Unaided pure-tone audiograms of CI participants. Arrows indicate no response at the audiometer's maximum output levels..

Table 1
Demographic distribution of the CI participants.

ID	Age (years)	Gender	Etiology of deafness	Duration of deafness (years)	Duration of CI use (years)	implant	Sound processor
S01	30	F	Ototoxicity	20	2	CI24RE	CP802
S02	21	M	Unknown	2	19	CI24M	CP810
S03	45	F	Unknown	20	10	CI24R(ST)	CP810
S04	23	F	Genetic	3	20	CI22M	CP910
S05	25	F	Unknown	5	5	CI24R(CA)	CP900
S06	24	F	Unknown	20	4	CI512	CP900
S07	25	M	Unknown	3	22	CI24M	CP1000
S08	27	M	Unknown	4	17	CI24M	CP950
S09	50	M	Unknown	2	3	CI24RE	CP802
S10	31	M	Malformation	8	10	CI24RE	CP810
S11	24	M	Meningitis	17	5	CI24RE	Freedom
mean	29.5	N/A	N/A	9.9	11.2	N/A	N/A

within 20 dB HL at octave frequencies from 125 to 8000 Hz in both ears), and were recruited from South China University of Technology. All participants signed informed consent statements and were paid for their participation. The study was approved by the Ethics Review Board of Shenzhen University.

3.2. Speech materials

Sentences from the Mandarin Chinese matrix (CMN matrix) test [35] spoken by a female speaker were used as the target speech. A 4-talker babble noise was used as the interferer. It was built using sentences from a female-speaker corpus from Mandarin speech perception (MSP) [36] and a male-speaker corpus from Mandarin Hearing in Noise Test (MHINT) [37]. First, two MSP lists (List 1 and List 2), each including 10 sentences, and one MHINT list (List 1) including 20 sentences divided into two 10-sentence lists, were selected to obtain four 10-sentence lists. Then, the 10 sentences within each of the four 10-sentence lists were concatenated and their root mean square (RMS) level was normalized. After that, the 2male-2female babble noise was built by the superposition of four pieces with a fixed duration of 10 s selected from each concatenated list with a random start point. In each trial, a piece of babble noise was selected with a random start point and was presented 300 ms before the target speech, ending 300 ms after.

3.3. Sound processor programs

All CI participants used the "8-of-22" ACE strategy with a 900-Hz stimulation rate in their MAPs, and were programmed in monopolar mode with the ball/needle and plate electrode (MP1 + 2) as the reference electrode. The default setting of C levels were mapped to 65 dB SPL and T levels were mapped to 25 dB SPL per channel. The slow-acting dynamic acoustic gains of ADRO® and Autosensitivity™ were enabled while the SmartSound® iQ function with SCAN was off. The participant's daily-used MAP was used as the base MAP. Then, EDR and n_{max} were modified based on the base MAP according to the experimental protocol. All new MAPs were then tested in the "live-mode", and the volume control was adjusted to achieve a loudness balance across MAPs. The volume control in the fitting software is realized by changing the C levels. During our experiment, the maximum C level change is about 6.4CL which was much smaller than the selected EDR.[2] CI participants' own voice or the experimenter's voice were additionally confirmed to be at a comfortable level.

For the NH listeners, a pulsatile Gaussian-enveloped tones (GET) vocoder [38] was used to simulate the ACE strategy. Compared to the conventional channel vocoders using continuous sine wave or noise carriers, GET vocoders use discrete acoustic Gaussian-enveloped tones to simulate the pulsatile electrical pulses in a pulse-to-pulse mapping fashion. Therefore, it provides direct simulations of EDR and maxima selection, neither of which are considered in conventional vocoders.

The implementation of this GET vocoder is reported in detail in [39]

[2] The volume control may be adjusted across 10 steps (from 1 to 10). In general, clinical MAPs for CI users use default %EDR (i.e., 20%) and volume control (i.e., 6). Therefore, a participant with an EDR of 80CL would experience a 6.4CL increase in stimulation when adjusting the volume control from 6 to 10 (80CL * 20% *4 / 10= 6.4CL) [7].

and is briefly summarized here. The GET vocoder differs from the conventional channel vocoders [40] mainly in the synthesizing end. Conventional vocoders extract envelopes in each channel, and use them to amplitude modulate a set of continuous carriers (sine wave tones, or, band-limited noise). The acoustic-electric mapping process in actual CIs is skipped and the pulses are not simulated. The GET vocoder first goes through the whole sound coding processing stages as in actual CI strategies (e.g., ACE) to get the pulsatile stimuli (sequences of stimulation channel and corresponding current level), including information of acoustic-electric mapping (using the default setting of the acoustic-to-electric compression function in ACE) and maxima selection. Then the inverse electric-to-acoustic function was used to convert the current level of all electrical pulses to acoustic levels. In this study, electrical pulses in the ACE strategy were generated using the software kit of the CCi-MOBILE research processor [41,42] which implemented the ACE strategy in MATLAB (version v2.2c1, https://crss.utdallas.edu/CILab/software.html). Then in each channel, the electric pulses were simulated by the Gaussian-shaped pulses, whose peak levels were defined by the acoustic levels, to construct the modulators for further amplitude modulation. Sinusoidal carriers with random initial phase are then amplitude modulated by the modulators, and the modulated signals across channels are summed and presented to both ears, i.e., diotically. Detailed implementation of the GET vocoder in the current study exactly followed the description of the GET vocoder in [39].

3.4. Setup and test conditions

Each CI participant was tested with ten conditions (a factorial design of five n_{max} settings and two EDR settings). The five n_{max} include $n_{max} =$ 2, 4, 6, 8 and 16. The two EDR conditions were the base EDR (EDR used by CI users in their daily life) and the 50 % EDR. Similar as in [32], the 50 % EDR was programmed by keeping C levels unchanged while increasing T levels (i.e., programmed T level = base T level + 50 % EDR on each electrode). Fig. 2 shows the across-channel average EDRs for the two conditions of each CI participant. The narrowest base EDR was 43.5CL for S06 and the widest was 81.4CL for S03. During the process of building these 10 MAPs, the base MAP serves as a benchmark for loudness balance. For CI participants, the stimulus was presented through a loudspeaker (Yamaha HS5) connected to a sound card (Focusrite 2i4) at 1 m in front of the participants, and had an output level of about 65 dB SPL.

Each NH participant was tested with nineteen conditions, including eighteen vocoded conditions (a factorial design of six n_{max} settings and three EDR settings) and one non-vocoded condition (i.e., original speech). The six n_{max} settings were $n_{max} = 2, 4, 6, 8, 12$ and 16. The three

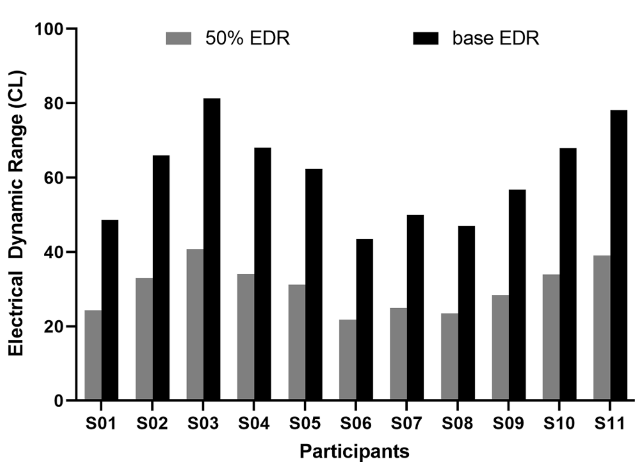

Fig. 2. The across-channel average width of the electrical dynamic range in current level (CL) units for the base condition and the 50% compression condition for the 11 participants.

EDR settings were 30, 100, and 150CL centered at 150CL (i.e., 135 to 165CL, 100 to 200CL, and 75 to 225CL). The central value of 150CL was selected based on the MAP records in our clinic. Fig. 3A-L shows examples of CI electrodogram and vocoded spectrogram using different n_{max} and EDR settings. As n_{max} increased, the density of pulses of all channels increased. In addition, the enlarged view of one channel with the same n_{max} in Fig. 3M-O illustrates better envelope representation with a larger EDR. For NH listeners, the vocoded stimulus was diotically delivered through a pair of headphones (HD650) connected to a sound card (Focusrite 2i4). Stimuli were presented at a comfortable level for each participant.

The experiments were conducted in a soundproof room with an indoor background noise below 30 dBA. Speech reception thresholds (SRTs), defined as the SNR required for 50 % correct performance, were measured using a 1-down-1-up adaptive procedure [43] in a closed-set format. The same test procedure was used for both CI and NH participants. Each participant took a 20–30-minute pre-test training with two lists (40 sentences) that were not used in the formal tests. In the formal test, two runs for each condition were conducted to check for consistency, and the average of the two runs was used to determine the final results. Therefore, each CI user finished 20 runs (10 conditions × 2 runs/condition) within one day, and each NH participant finished 38 runs (19 conditions × 2 runs/condition) in two sessions separated with an interval of at least twenty-four hours. All experiments were carried out in a double-blind fashion and all conditions were presented in random order for each run.

3.5. Statistical analyses

The Shapiro-Wilk (S-W) normality test was made, followed by a two-way repeated-measures analysis of variance (RM ANOVA). Sphericity was tested using Mauchly's test to ensure the assumption of sphericity was not violated, and multivariate models were applied if these assumptions were not met. *Post hoc* analysis was performed using Fisher's least significant difference (LSD) correction. An independent samples *t*-test was used to compare the results of CI users and vocoder simulations. $p < 0.05$ was considered statistically significant using an error rate of $\alpha = 0.05$.

4. Results

4.1. CI participants

Fig. 4 shows the individual SRTs in different n_{max} settings for CI participants with base EDR (panel A) and 50 % EDR (panel B), as well as the comparison of mean SRTs in different conditions (panel C). In panel A and B, each point was a mean of two measured SRTs, with the difference between the two measured SRTs was 1.4 ± 1.3 dB. All results were confirmed to be in compliance with normal distribution by the S-W test. A two-way RM ANOVA was completed with n_{max} and EDR as the independent variables and the SRTs as the dependent variable. There was a significant main effect of n_{max} [F (4,7) = 11.592, $p < 0.01$] and of EDR [F (1,10) = 6.545, $p < 0.05$]. No n_{max} × EDR interactions were found to be significant [F (4,7) = 1.532 , $p > 0.05$].

Post hoc multiple comparisons with LSD corrections revealed significant performance differences between $n_{max} = 2$ and all other n_{max} conditions ($p < 0.05$ for all comparisons). The worst performance was observed at $n_{max} = 2$ (mean = 4.2 dB). The best performance was observed at $n_{max} = 6$ (mean = − 0.9 dB) and 8 (mean = − 0.7 dB), with SRTs significantly lower than those of $n_{max} = 2$ (mean = 4.2 dB), 4 (mean = 0.0 dB) and 16 (mean = 0.2 dB) ($p < 0.05$ for all comparisons). Additionally, there was a small but significant difference between base EDR and 50 % compression EDR (mean SRT gap = –0.9 dB).

Berg *et al.* (2019) which found that perimodiolar electrodes may have a positive effect on speech perception with increased n_{max}. Among the 11 CI participants in our experiment, five (black lines in Fig. 4A)

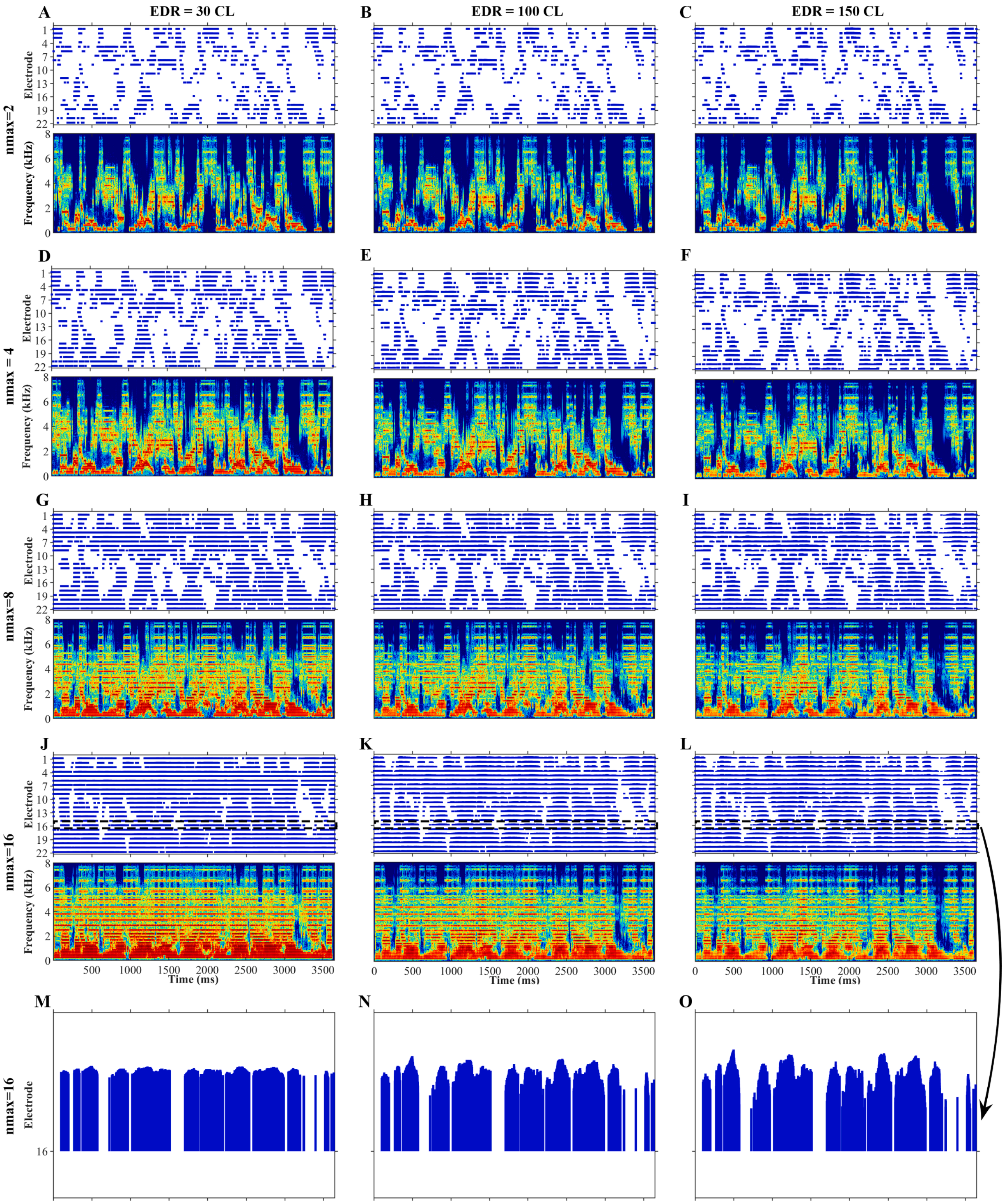

Fig. 3. **(A-L).** Example electrodograms and spectrograms of the same target speech and babble noise for different combination settings at an SNR of 3 dB. The EDR settings are shown on the top and n_{max} settings are shown on the left. **(M−O).** An enlarged view of Channel 16 of the electrodograms with $n_{max} = 16$.

were using lateral wall electrodes, and the remaining six (gray lines in Fig. 4A) were using perimodiolar electrodes. It can be observed from Fig. 4A that participants with these two different types of electrodes demonstrated different patterns when n_{max} was increased from 8 to 16 at the base EDR. A significant performance difference were found between $n_{max} = 8$ (mean SRT = − 1.1 dB) and $n_{max} = 16$ (mean SRT = 0.4 dB) in perimodiolar electrode recipients with base EDR ($p < 0.05$), and also in lateral wall electrode recipients with 50 % EDR (mean SRT = − 1.2 dB vs 0.4 dB) ($p < 0.05$). Both of these two comparisons showed opposite results to the finding of Berg *et al.* (2019). No significant difference was found for the other two conditions (Table 2).

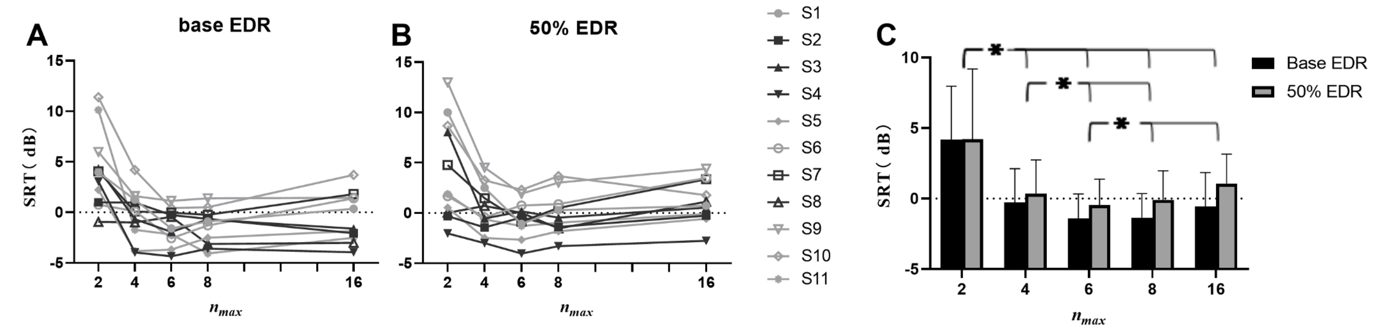

Fig. 4. SRTs in different n_{max} settings for CI participants. **A** and **B** show the individual SRTs for and 50 % compression EDR, respectively. Gray legends represent CI recipients with perimodiolar electrodes and black legends represent lateral wall electrodes. **C** illustrates group means and standard deviations. * indicates statistical significance (p < 0.05).

Table 2
Paired samples *t*-tests for n_{max} = 8 and 16.

		n_{max} = 8	n_{max} = 16	*t*	*P*
base EDR	lateral wall electrodes (n = 5)	−1.7 ± 1.6	−1.8 ± 2.2	0.17	0.874
	perimodiolar electrodes (n = 6)	−1.1 ± 2.0	0.4 ± 2.3	−3.21	0.024*
50 % EDR	lateral wall electrodes (n = 5)	−1.2 ± 1.4	0.4 ± 2.2	−3.47	0.026*
	perimodiolar electrodes (n = 6)	0.8 ± 2.2	1.6 ± 2.0	−1.27	0.262

[*] represents $p < 0.05$ with statistical significance.

4.2. Vocoder simulations in NH participants

Results of vocoder simulations and non-vocoded original tests in NH participants are shown in Fig. 5A. For the non-vocoded condition, the mean SRT was − 9.7 dB, which was significantly better than all vocoded SRTs. The mean SRTs for the vocoded conditions were mostly higher than 0 dB. Data normality was confirmed by the S-W test. A two-way RM ANOVA found a significant main effect of n_{max} [F (5, 5) = 75.9, p < 0.001] and of EDR [F (2, 18) = 480.1, p < 0.001]. Furthermore, the n_{max} × EDR interaction was also significant [F (10, 90) = 44.811 , p < 0.001].

Results of LSD multiple-comparisons are shown in Fig. 5B-D. n_{max} = 2 derived significantly higher mean SRTs than n_{max} = 4 for all three EDR conditions (p < 0.05 for all comparisons). Default n_{max} = 8 for the ACE strategy derived significantly lower mean SRTs than at n_{max} = 12 and 16 for almost all EDR conditions (p < 0.05), except n_{max} = 8 vs n_{max} = 12 at

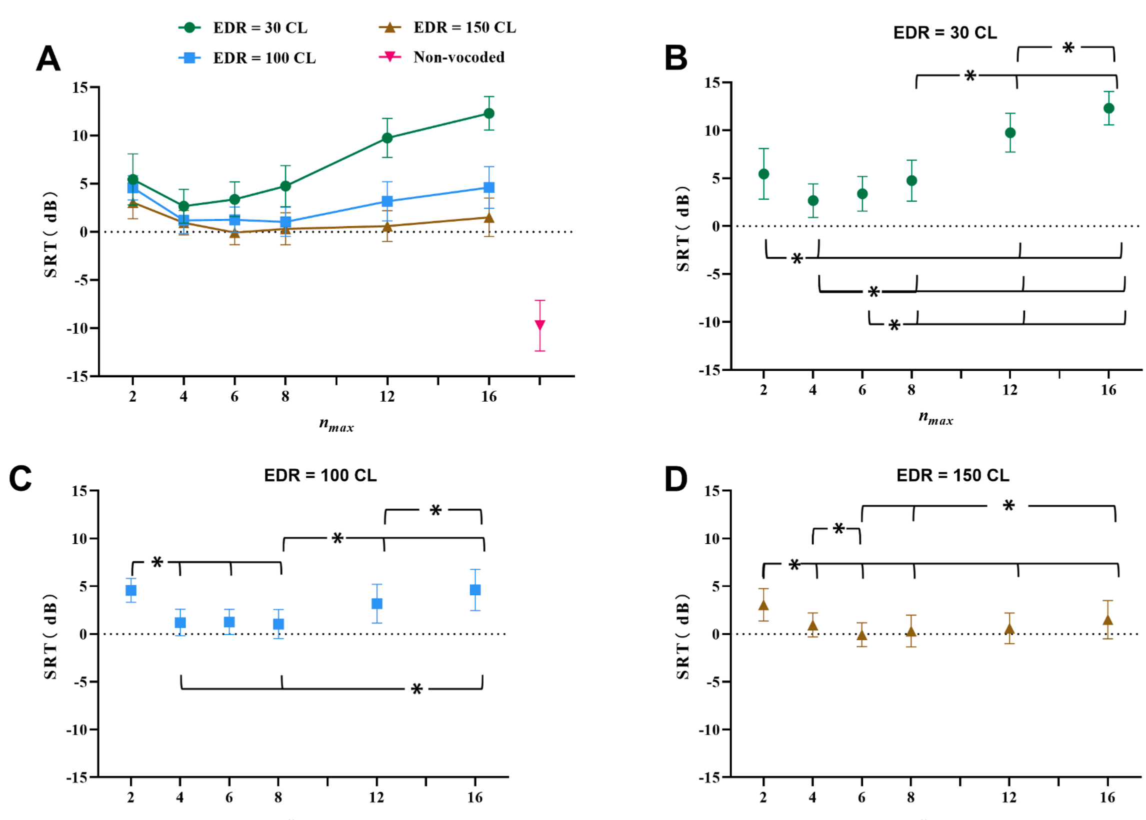

Fig. 5. Mean SRTs for different listening settings in the vocoder test. **A.** SRTs for all 19 conditions. **B-D.** SRTs across different n_{max} at EDR = 30, 100 and 150CL. Error bars represent standard deviation (SD). Symbols for statistical significance levels: *p < 0.05 (refers to LSD multiple-comparison).

EDR = 150CL (p = 1.00). Additionally, n_{max} = 8 yields mean SRTs comparable to n_{max} = 4 in EDR = 100 and 150CL condition (gap = –0.2 dB and –0.6 dB, $p > 0.05$ for two comparison), but significantly higher than that in EDR = 30CL (gap = 2.1 dB, $p < 0.05$).

Wider EDR derived significantly lower SRTs for almost all n_{max} conditions ($p < 0.05$) (Fig. 5A). The only exceptions with non-significant differences were 30CL vs 100CL at n_{max} = 2 and 100CL vs 150CL at n_{max} = 2, 4, and 8 conditions ($p > 0.05$). The results also showed that when the EDR is narrower, the SRTs were more susceptible to changes in n_{max}, e.g., the difference between the maximum and minimum mean SRTs are 12.3 dB and 2.7 dB for EDR = 30 and 150CL, respectively.

4.3. Actual CIs vs vocoder simulations

The results of actual CI users (Fig. 4C) and vocoder simulations (Fig. 5A) in different conditions are summarized in Fig. 6. All curves showed convex shapes. Among the five EDR conditions of either CI or simulation, CI users with base EDR performed best in all conditions except at n_{max} = 2, while in all CI simulation conditions, NH listeners performed the worst when EDR = 30CL. Note that the curves of EDR = 150CL (NH) and 50 % EDR (CI) match quite well with each other. Independent-sample t tests were used to compare the independent groups with each n_{max} between EDR = 150CL (NH) and either EDR (CI) condition. As shown in Table 3, there is no significant difference between the results of CI and CI simulations except at n_{max} = 8 and n_{max} = 16 in the comparison of EDR = 150CL (NH) and base EDR (CI). The acute mean CI simulation results at EDR = 150CL coincide with those of the experienced CI users at base EDR and 50 % EDR. Their differences are not significant and no>0.6 dB with $n_{max} > 2$.

5. Discussion

5.1. Effect of the number of maxima (n_{max})

In previous studies investigating the effects of active electrode number and number of maxima (i.e., n_{max}), the factor of EDR was not controlled or even reported. In our study, compared to n_{max} = 6 and 8, the suboptimal n_{max} (corresponding to the higher mean SRTs) were larger n_{max} (i.e., 16) and smaller n_{max} (i.e., $n_{max} < 4$) in different EDR conditions for both CI users and vocoder simulations using in NH listeners (Fig. 4C and Fig. 5). The results that n_{max} = 2 and 4 would derive poorer performance was consistent with the results of [15], which is intuitive given the coarse spectral resolution (Fig. 3A-C). Even so, the mean SRTs with n_{max} = 2 for all conditions was around 5 dB, which is a surprisingly good performance result.

The finding that a large n_{max} (i.e., n_{max} = 16) yielded poorer performance contradicts previous studies. Both Dorman *et al.*, (2002) and Plant *et al.*, (2002) reported performance saturation when n_{max} reaches about 8 [15,23], while Berg *et al.* (2019) suggested that performance was significantly better with n_{max} = 16 maxima than with n_{max} = 8 for

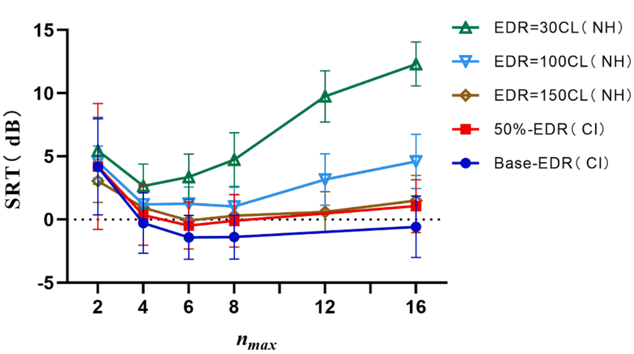

Fig. 6. Mean SRTs for different listening conditions among CI users and simulation in NH listeners.

perimodiolar electrode CI recipients [27]. Berg *et al.* (2019) also pointed out that there was a significant negative correlation between electrode-to-modiolus distance and the degree of benefit afforded by n_{max} = 16 [27]. In our study, 6 CI recipients with perimodiolar electrodes (i.e., with relatively smaller electrode-to-modiolus distance) and 5 CI recipients with lateral wall electrodes (i.e., with a relatively larger electrode-to-modiolus distance) were included. However, no benefit of n_{max} = 16 was found in participants with small electrode-to-modiolus distance. Instead, participants performed better at n_{max} = 8 than at n_{max} = 16. This finding implies that more individual factors of CI users should be considered and regression and correlation analyses may be needed with a larger sample size in future studies.

In theory, there are several factors that could contribute to a larger n_{max} not necessarily yielding better performance for CI users in our study. Firstly, the increased n_{max} may introduce more noise [44], and CI users have poorer abilities to segregate the target speech from the background noise. The higher the n_{max}, the greater the likelihood of selecting a noise dominant channel, especially in a low SNR condition. The combined effects make it more difficult for CI listeners with higher n_{max} to perceive speech. Additionally, the increased channel interaction with n_{max} increase might be a contributing factor. Finally, because of poor spectral resolution and high inter-channel interference, the lateral wall electrodes used by many CI recipients may prevent taking advantage of the more spectral information that comes with n_{max} = 16.

5.2. Effect of electrical dynamic range (EDR)

In our study, narrower EDR within each patient's psychophysically measured range was always associated with poorer performance in both CI and vocoder groups (Fig. 6). The EDR compression results in reduced speech recognition by significantly increasing the T level for all channels, as was shown in prior studies among CI users [28,32]. Although it has also been shown that significantly better sentence recognition in noise is observed during EDR compression, this is mainly due to the small amount of T levels adjustment in their study [30,31]. In signal processing, spectral modulation depth increases with increasing EDR. The wider the EDR, the better representation of spectral peaks (Fig. 3G-L) and spectral envelope contrast (Fig. 3H-O), and the smaller the interference in the speech gap (Fig. 3G-L). Spectral contrast enhancement may help with background-noise reduction and improvement of speech recognition [45]. It is reasonable to presume that the mechanism underlying poor recognition in narrower EDR could be the blurring of spectral contrast due to shallower envelope modulation. In addition, in tonal languages like Mandarin, pitch represented by F0-related periodicity plays a critical role in determining word meaning. However, the F0 representation in the ACE strategy may become more indistinct as the amplitude modulation becomes shallower (i.e., EDRs become narrower), which is related to the weakening of signal periodicity.

Some studies have shown that higher T levels can improve the aided threshold and audibility of soft sounds [12,30,31]. However, compressing EDR by increasing the overall T level in all channels will amplify both noise and speech, which may reduce the SNR and impair speech recognition. The increased risk of interaction across channels should be noted with increasing stimulation level (T level). Furthermore, the increase in EDR by decreasing T levels was found to result in degraded speech recognition [33,46]. Thus, by setting the T levels above or below the "true" perception threshold, CI recipients cannot use sufficient spectral information even with wider EDR.

In order to maximize the use of EDR depending on the auditory nerve, it is crucial to measure the "true" threshold accurately in CI programming. The fitted EDR is within the range of 40–60CL for most CI users [28] and the majority of CI user's EDR showed only small amounts of change in levels in long-term post-implantation [47]. Simulation level might correlate with nerve survival and distance from the electrode to the modiolus [28,48]. Considering the potential advantage of a wide EDR, CI device and surgery techniques are suggested to be improved to

Table 3
Independent samples *t*-tests for CI and CI simulations.

	EDR = 150CL (NH) (n = 10)	Base-EDR (CI) (n = 11)	*t*	*P*	EDR = 150CL (NH) (n = 10)	50 %-EDR (CI) (n = 11)	*t*	*P*
$n_{max} = 2$	3.1 ± 1.7	4.2 ± 3.8	−0.85	0.404	3.1 ± 1.7	4.2 ± 4.0	−0.69	0.497
$n_{max} = 4$	1.0 ± 1.3	−0.3 ± 2.4	1.44	0.166	1.0 ± 1.3	0.4 ± 2.4	0.70	0.493
$n_{max} = 6$	−0.1 ± 1.3	−1.4 ± 1.7	2.00	0.060	−0.8 ± 1.3	−0.5 ± 1.9	0.57	0.578
$n_{max} = 8$	0.3 ± 1.7	−1.4 ± 1.7	2.28	0.034*	0.3 ± 1.7	−0.1 ± 2.1	0.50	0.626
$n_{max} = 16$	1.5 ± 2.0	−0.6 ± 2.4	2.14	0.046*	1.5 ± 2.0	1.1 ± 2.1	0.50	0.624

* represents $p < 0.05$ with statistical significance.

preserve a wider EDR for users in the future. Meanwhile, audiologists should accurately assess T/C levels of CI recipients at all electrodes to avoid the negative impact of inappropriate EDR on speech perception.

5.3. Effectiveness of CI simulation with GET vocoder

Overall trends of SRTs as a function of EDR or n_{max} = 2, 4, 6, 8, and 16 are similar between the actual and GET simulated CI listeners (Fig. 6). And the SRT-n_{max} function for CIs in GET vocoder simulations in EDR = 150CL is almost in perfect agreement with that of 50 % compression EDR and base EDR (Table 2.

In light of the significant interaction between EDR and n_{max}, the optimal n_{max} is smaller than 8 (i.e., 4 and 6) for EDR = 30CL, but around 8 for EDR = 100 and 150CL. Conversely, this significant interaction was not found in CI users. Furthermore, the SRT-n_{max} curves of vocoders with EDR = 30 and 100CL are steeper than those of CIs, which may be explained by individual variability in speech perception of actual CI users. It is possible that the GET simulation results with a narrow EDR (i.e., 30CL) are a prediction for some poor performance CI users. Narrower EDRs introduce more spectral-temporal interactions and shallower modulations which may also be limitations for poor performance users.

One reason for the worse results (i.e., higher SRT) for GET simulation than our actual CI results using comparable EDRs may come from training effects. In current study, training was only carried out in two-of-eighteen conditions before the acute test, but our actual CI participants all have at least two years of listening experience and generally have good speech communication abilities. Simulation results may get closer to CI results after adequate training to the NH listeners. Previous studies have shown that NH listeners can get significant improvements in recognizing vocoded speech after systematic training [49–52].

5.4. Limitations

There were several limitations in our experiments:

1) The narrow EDRs of CI users in this experiment were obtained by increasing the T level, and cannot fully represent "true" narrow EDRs. Recruiting more participants with different EDRs for between-subject experiments will be required for future studies.
2) For the vocoder simulation, the simulated EDRs were also not simulating the true absolute thresholds and most comfortable levels, which means the loudness dynamic ranges are not equivalent between the actual and simulated CI listeners.
3) To get better (or closer to actual CIs) simulation results, two directions could be tried. The first is adding systematic adaptive training before formal testing. The second is optimizing the vocoder parameters. Further research is needed to validate and optimize this technique.
4) The effects of the two target parameters on CI speech quality perception in quiet conditions or high SNR conditions was not examined.

6. Conclusion

For both actual and GET-simulated CI listeners, recognition of speech-in-noise with negative (<0 dB) SNRs is a great challenge, while NH listeners could still recognize ~ 50 % of speech at much lower SNRs (e.g., ~ −10 dB in our experiment) even without spatial cues. Our results also indicate that pulsatile GET vocoder could generate good predictions and implications for SRT in noise with CIs.

Electrical dynamic range and number of maxima in an "*n*-of-*m*" strategy are independently associated with speech-in-noise perception by influencing the spectro-temporal envelope modulation transmission. In the ACE strategy, additional information provided by increasing n_{max} from 8 to 16 may degrade speech in noise recognition with CIs. The EDR compression, caused by inaccurate measurements and settings in ACE strategy, is detrimental to temporal envelope presentation and speech-in-noise recognition in CI recipients.

Funding

This work was supported by the Guangdong Basic and Applied Basic Research Foundation Grant (2020A1515010386), Science and Technology Program of Guangzhou (202102020944, 202102020689), Guangdong Medical Science and Technology Research Foundation Project (B2020153), National Natural Science Foundation of China (61771320), Guangdong Key Area R&D Project (No. 2018B030338001), Cochlear Limited, and Tencent Corporation.

CRediT authorship contribution statement

Yefei Mo: Methodology, Investigation, Data curation, Writing – original draft. **Huali Zhou:** Methodology, Software, Data curation. **Fanhui Kong:** Investigation, Writing – review & editing, Supervision. **Zhifeng Liu:** Investigation. **Xiaohong Liu:** Investigation, Resources. **Hongming Huang:** Supervision, Writing – review & editing. **Yan Huang:** Funding acquisition, Writing – review & editing. **Nengheng Zheng:** Writing – review & editing, Funding acquisition. **Qinglin Meng:** Conceptualization, Methodology, Writing – review & editing, Resources, Project administration, Funding acquisition. **Peina Wu:** Conceptualization, Methodology, Writing – review & editing, Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We would like to thank all the research volunteers that generously donated their time to participate in this study. We thank Drew Cappotto for helping on proof-reading of the early manuscripts.

References

[1] World Health Organization, World report on hearing, World Health Organization (2021). https://www.who.int/publications/i/item/world-report-on-hearing.

[2] F.-G. Zeng, Celebrating the one millionth cochlear implant, JASA Express Lett 2 (2022) 077201. https://doi.org/10.1121/10.0012825.

[3] B.S. Wilson, M.F. Dorman, The surprising performance of present-day cochlear implants, IEEE Trans. Biomed. Eng. 54 (2007) 969–972, https://doi.org/10.1109/TBME.2007.893505.

[4] B.S. Wilson, M.F. Dorman, Cochlear implants: A remarkable past and a brilliant future, Hear. Res. 242 (2008) 3–21, https://doi.org/10.1016/j.heares.2008.06.005.

[5] B.L. Fetterman, E.H. Domico, Speech recognition in background noise of cochlear implant patients, Otolaryngol. Head Neck Surg. 126 (2002) 257–263, https://doi.org/10.1067/mhn.2002.123044.

[6] P.C. Loizou, Y. Hu, R. Litovsky, G. Yu, R. Peters, J. Lake, P. Roland, Speech recognition by bilateral cochlear implant users in a cocktail-party setting, J. Acoust. Soc. Am. 125 (2009) 372–383. https://doi.org/10.1121/1.3036175.

[7] A. Spriet, L. Van Deun, K. Eftaxiadis, J. Laneau, M. Moonen, B. Van Dijk, A. Van Wieringen, J. Wouters, Speech understanding in background noise with the two-microphone adaptive beamformer BEAM™ in the Nucleus Freedom™ cochlear implant system, Ear Hear. 28 (2007) 62–72, https://doi.org/10.1097/01.aud.0000252470.54246.54.

[8] H. Zhou, N. Wang, N. Zheng, G. Yu, Q. Meng, A New Approach for Noise Suppression in Cochlear Implants: A Single-Channel Noise Reduction Algorithm1, Front. Neurosci. 14 (2020), https://doi.org/10.3389/fnins.2020.00301.

[9] S. Aydın, Tikhonov regularized solutions for improvement of signal-to-noise ratio in case of auditory-evoked potentials, Med. Biol. Eng. Compu. 46 (2008) 1051–1056, https://doi.org/10.1007/s11517-008-0385-0.

[10] S. Aydın, A new combination: scale-space filtering of projected brain activities, Med. Biol. Eng. Compu. 47 (2009) 435–440, https://doi.org/10.1007/s11517-009-0450-3.

[11] A. Kan, Q. Meng, The Temporal Limits Encoder as a Sound Coding Strategy for Bilateral Cochlear Implants, IEEE/ACM Trans Audio Speech Lang Process 29 (2021) 265–273, https://doi.org/10.1109/TASLP.2020.3039601.

[12] T.B. Nunn, D. Jiang, T. Green, P.J. Boyle, D.A. Vickers, A systematic review of the impact of adjusting input dynamic range (IDR), electrical threshold (T) level and rate of stimulation on speech perception ability in cochlear implant users, Int. J. Audiol. 58 (2019) 317–325, https://doi.org/10.1080/14992027.2018.1564844.

[13] Ltd. COCHLEAR^{TM}, Clinical Guidance Document, D1 128020 ISS6, COCHLEARTM, Ltd. (2017).

[14] J. Wolfe, Cochlear implants: audiologic management and considerations for implantable hearing devices, CA, Plural Publishing, San Diego, 2020.

[15] M.F. Dorman, P.C. Loizou, A.J. Spahr, E. Maloff, A comparison of the speech understanding provided by acoustic models of fixed-channel and channel-picking signal processors for cochlear implants, Journal of speech, language, and hearing research : JSLHR 45 (2002) 783–788, https://doi.org/10.1044/1092-4388(2002/063).

[16] P.C. Loizou, M. Dorman, J. Fitzke, The effect of reduced dynamic range on speech understanding: implications for patients with cochlear implants, Ear Hear. 21 (2000) 25–31.

[17] W. Nogueira, A. Büchner, T. Lenarz, B. Edler, A. Psychoacoustic, "NofM"-Type Speech Coding Strategy for Cochlear Implants, EURASIP J. Adv. Signal Process. 2005 (2005), https://doi.org/10.1155/ASP.2005.3044.

[18] L.M. Friesen, R.V. Shannon, D. Baskent, X. Wang, Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants, J. Acoust. Soc. Am. 110 (2001) 1150–1163, https://doi.org/10.1121/1.1381538.

[19] C. Garnham, M. O'Driscoll, R. Ramsden, S. Saeed, Speech understanding in noise with a Med-El COMBI 40+ cochlear implant using reduced channel sets, Ear Hear. 23 (6) (2002) 540–552.

[20] R. Shannon, Q.-J. Fu, J. Galvin Iii, The number of spectral channels required for speech recognition depends on the difficulty of the listening situation, Acta Otolaryngol. Suppl. 124 (0) (2004) 50–54.

[21] N.B.H. Croghan, S.I. Duran, Z.M. Smith, Re-examining the relationship between number of cochlear implant channels and maximal speech intelligibility, J. Acoust. Soc. Am. 142 (6) (2017) EL537.

[22] P.-A. Cucis, C. Berger-Vachon, H. Thaï-Van, R. Hermann, S. Gallego, E. Truy, Word Recognition and Frequency Selectivity in Cochlear Implant Simulation: Effect of Channel Interaction, J. Clin. Med. 10 (4) (2021) 679.

[23] K.L. Plant, L.A. Whitford, C.E. Psarros, A.E. Vandali, Parameter selection and programming recommendations for the ACE and CIS speech-processing strategies in the Nucleus 24 cochlear implant system, Cochlear implants international 3 (2002) 104–125, https://doi.org/10.1179/cim.2002.3.2.104.

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

[25] J.G. Naples, M.J. Ruckenstein, Cochlear Implant, Otolaryngol. Clin. North Am. 53 (2020) 87–102, https://doi.org/10.1016/j.otc.2019.09.004.

[26] R.H. Gifford, M.F. Dorman, J.K. Shallop, S.A. Sydlowski, Evidence for the expansion of adult cochlear implant candidacy, Ear Hear 31 (2010) 186–194, https://doi.org/10.1097/AUD.0b013e3181c6b831.

[27] K.A. Berg, J.H. Noble, B.M. Dawant, R.T. Dwyer, R.F. Labadie, R.H. Gifford, Speech recognition as a function of the number of channels in perimodiolar electrode recipients, J. Acoust. Soc. Am. 145 (3) (2019) 1556–1564.

[28] F. de Graaff, B.I. Lissenberg-Witte, M.W. Kaandorp, P. Merkus, S.T. Goverts, S. E. Kramer, C. Smits, Relationship Between Speech Recognition in Quiet and Noise and Fitting Parameters, Impedances and ECAP Thresholds in Adult Cochlear Implant Users, Ear Hear. 41 (2020) 935–947, https://doi.org/10.1097/AUD.0000000000000814.

[29] S.Y. Kim, S.-K. Jeon, S.H. Oh, J.H. Lee, M.-W. Suh, S.-Y. Lee, H.-J. Lim, M.K. Park, Electrical dynamic range is only weakly associated with auditory performance and speech recognition in long-term users of cochlear implants, Int. J. Pediatr. Otorhinolaryngol. 111 (2018) 170–173, https://doi.org/10.1016/j.ijporl.2018.06.016.

[30] M.W. Skinner, L.K. Holden, T.A. Holden, M.E. Demorest, Comparison of two methods for selecting minimum stimulation levels used in programming the Nucleus 22 cochlear implant, Journal of Speech, Language, and Hearing Research 42 (1999) 814–828, https://doi.org/10.1044/jslhr.4204.814.

[31] N. Zhou, B.E. Pfingst, Effects of site-specific level adjustments on speech recognition with cochlear implants, Ear Hear. 35 (1) (2014) 30–40.

[32] P.A. Busby, K. Arora, Effects of Threshold Adjustment on Speech Perception in Nucleus Cochlear Implant Recipients, Ear Hear. 37 (2016) 303–311, https://doi.org/10.1097/AUD.0000000000000248.

[33] K.V.C. Martins, M.V.S. Goffi-Gomez, R.K. Tsuji, R.F. Bento, Do the minimum and maximum comfortable stimulation levels influence the cortical potential latencies or the speech recognition in adult cochlear implant users? Hear. Res. 404 (2021), 108206 https://doi.org/10.1016/j.heares.2021.108206.

[34] F.-G. Zeng, J.J. Galvin, Amplitude mapping and phoneme recognition in cochlear implant listeners, Ear Hear. 20 (1) (1999) 60–74.

[35] H. Hu, X. Xi, L.L.N. Wong, S. Hochmuth, A. Warzybok, B. Kollmeier, Construction and evaluation of the Mandarin Chinese matrix (CMNmatrix) sentence test for the assessment of speech recognition in noise, Int. J. Audiol. 57 (2018) 838–850, https://doi.org/10.1080/14992027.2018.1483083.

[36] Q.-J. Fu, M. Zhu, X. Wang, Development and validation of the Mandarin speech perception test, The Journal of the Acoustical Society of America 129, J. Acoust. Soc. Am. 129 (6) (2011) EL267–EL273.

[37] L.L.N. Wong, S.D. Soli, S. Liu, N. Han, M.-W. Huang, Development of the Mandarin Hearing in Noise Test MHINT, Ear. Hear. 28 (2007) 70S–74S.

[38] Q. Meng, G. Yu, Y. Wan, F. Kong, X. Wang, N. Zheng, Effects of Vocoder Processing on Speech Perception in Reverberant Classrooms, Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) 2018 (2018) 761–765.

[39] Q. Meng, H. Zhou, T. Lu, F.-G. Zeng, Pulsatile Gaussian-Enveloped Tones (GET) Vocoders for Cochlear-Implant Simulation, medRxiv (2022). https://www.medrxiv.org/content/10.1101/2022.02.21.22270929v3.

[40] M.F. Dorman, P.C. Loizou, D. Rainey, Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs, J. Acoust. Soc. Am. 102 (1997) 2403–2411, https://doi.org/10.1121/1.419603.

[41] H. Ali, N. Mamun, A. Bruggeman, R.C.M. Chandra Shekar, J.N. Saba, J.H.L. Hansen, The CCi-MOBILE Vocoder, J. Acoust. Soc. Am. 144 (2018) 1872–1872. https://doi.org/10.1121/1.5068238.

[42] J.H.L. Hansen, H. Ali, J.N. Saba, C.M.C. Ram, N. Mamun, R. Ghosh, A. Brueggeman, CCi-MOBILE: Design and Evaluation of a Cochlear Implant and Hearing Aid Research Platform for Speech Scientists and Engineers, in: IEEE-EMBS International Conference on Biomedical and Health Informatics. IEEE-EMBS International Conference on Biomedical and Health Informatics 2019, 2019, https://doi.org/10.1109/BHI.2019.8834652.

[43] T. Brand, B. Kollmeier, Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests, J. Acoust. Soc. Am. 111 (2002) 2801–2810, https://doi.org/10.1121/1.1479152.

[44] O.U.R. Qazi, B. van Dijk, M. Moonen, J. Wouters, Understanding the effect of noise on electrical stimulation sequences in cochlear implants and its impact on speech intelligibility, Hear. Res. 299 (2013) 79–87, https://doi.org/10.1016/j.heares.2013.01.018.

[45] N. El Boghdady, F. Langner, E. Gaudrain, D. Baskent, W. Nogueira, Effect of Spectral Contrast Enhancement on Speech-on-Speech Intelligibility and Voice Cue Sensitivity in Cochlear Implant Users, Ear. Hear 42 (2020) 271–289, https://doi.org/10.1097/AUD.0000000000000936.

[46] T. Rader, P. Doms, Y. Adel, T. Weissgerber, S. Strieth, U. Baumann, A method for determining precise electrical hearing thresholds in cochlear implant users, Int. J. Audiol. 57 (2018) 502–509, https://doi.org/10.1080/14992027.2017.1412519.

[47] E.A. Gajadeera, K.L. Galvin, R.C. Dowell, P.A. Busby, Investigation of Electrical Stimulation Levels Over 8 to 10 Years Postimplantation for a Large Cohort of Adults Using Cochlear Implants, Ear. Hear. 38 (2017) 736–745, https://doi.org/10.1097/AUD.0000000000000466.

[48] C.J. Long, T.A. Holden, G.H. McClelland, W.S. Parkinson, C. Shelton, D.C. Kelsall, Z.M. Smith, Examining the electro-neural interface of cochlear implant users using psychophysics, CT scans, and speech understanding, J. Assoc. Res. Otolaryngol. 15 (2) (2014) 293–304.

[49] M.H. Davis, I.S. Johnsrude, A. Hervais-Adelman, K. Taylor, C. McGettigan, Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences, J. Exp. Psychol. Gen. 134 (2005) 222–241, https://doi.org/10.1097/AUD.0000000000000814.

[50] M.J. Goupell, G.T. Draves, R.Y. Litovsky, Recognition of vocoded words and sentences in quiet and multi-talker babble with children and adults, PloS one 15 (2020) e0244632. https://doi.org/10.1371/journal.pone.0244632.

[51] J.J. Huyck, R.H. Smith, S. Hawkins, I.S. Johnsrude, Generalization of Perceptual Learning of Degraded Speech Across Talkers, Journal of speech, language, and hearing research : JSLHR 60 (2017) 3334–3341, https://doi.org/10.1044/2017_JSLHR-H-16-0300.

[52] X. Chen, Y. Mo, K. Ouyang, M. Shi, H. Zhou, Y. Shi, W. Xiao, S. Shang, Q. Meng, N. Zheng, Internet Streaming Audio Based Speech Reception Threshold Measurement in Cochlear Implant Users, ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2022, pp. 9012-9016. Nogueirahttps://doi.org/10.1109/ICASSP43922.2022.9747404.