

# Speech Understanding With Various Maskers in Cochlear-Implant and Simulated Cochlear-Implant Hearing: Effects of Spectral Resolution and Implications for Masking Release

Trends in Hearing  
Volume 22: 1–13  
© The Author(s) 2018  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/2331216518787276  
journals.sagepub.com/home/tia



Naomi B. H. Croghan<sup>1,2</sup> and Zachary M. Smith<sup>1,3</sup>

## Abstract

The purpose of this study was to investigate the relationship between psychophysical spectral resolution and sentence reception in various types of interfering backgrounds for listeners with cochlear implants and normal-hearing subjects listening to vocoded speech. Spectral resolution was measured with a spectral modulation detection (SMD) task. For speech testing, maskers included stationary speech-shaped noise (SSN), four-talker babble, multitone noise, and a competing talker. To explore the possible trade-offs between spectral resolution and susceptibility to different types of maskers, the degree of simulated current spread was varied within the vocoder group, achieving a range of performance for SMD and speech tasks. Greater simulated current spread was detrimental to both spectral resolution and speech recognition, suggesting that interventions that decrease current spread may improve performance for both tasks. Better SMD sensitivity was significantly correlated with improved sentence reception. In addition, differences in sentence reception across the four maskers were significantly associated with SMD across the combined group of cochlear-implant and vocoder subjects. Masking release (MR) was quantified as the signal-to-noise ratio difference in speech reception threshold between the SSN and competing talker. Several individual cochlear-implant subjects demonstrated substantial MR, in contrast to previous studies, and the degree of MR increased with better SMD thresholds across subjects. The results of this study suggest that alternative masker types, particularly competing talkers, are more sensitive than stationary SSN to differences in spectral resolution in the cochlear-implant population.

## Keywords

cochlear implants, speech perception, psychophysics, spectral resolution

Date received: 4 May 2017; revised: 1 June 2018; accepted: 4 June 2018

## Introduction

Cochlear implants (CIs), in part, restore the sense of hearing to people with moderate to profound sensorineural hearing loss. Although this technology successfully conveys the cues necessary for speech understanding in quiet, many CI listeners struggle to comprehend speech in the presence of interfering background sounds. Degraded spectral resolution, due to current spread in the cochlea, limited numbers of spectral channels, and other factors, may partially explain differences in speech-in-noise recognition between the CI and normal-hearing (NH) populations. Therefore, significant attention has been given to measuring spectral resolution

in CI listeners, particularly with tests using spectral-ripple stimuli (Anderson, Nelson, Kreft, Nelson, & Oxenham, 2011; Berenstein, Mens, Mulder, & Vanpoucke, 2008;

<sup>1</sup>Denver Research & Technology Labs, Cochlear Ltd., Centennial, CO, USA

<sup>2</sup>Department of Speech, Language, and Hearing Sciences, University of Colorado, Boulder, CO, USA

<sup>3</sup>Department of Physiology and Biophysics, School of Medicine, University of Colorado, Aurora, CO, USA

## Corresponding author:

Naomi B. H. Croghan, Denver Research & Technology Labs, Cochlear Ltd., 13059 E. Peakview Ave., Centennial, CO 80111, USA.

Email: [ncroghan@cochlear.com](mailto:ncroghan@cochlear.com)



Drennan, Won, Timme, & Rubinstein, 2016; Gifford, Hedley-Williams, & Spahr, 2014; Henry & Turner, 2003; Saoji, Litvak, Spahr, & Eddins, 2009; Won, Drennan, & Rubinstein, 2007).

Improvements in spectral-ripple discrimination have been shown as a result of technological developments aimed at increasing spectral resolution in CI subjects, such as focused stimulation (Smith, Parkinson, & Long, 2013). In addition, a number of studies with CI listeners have shown significant correlations between spectral-ripple discrimination or spectral-ripple detection and speech perception in quiet (Anderson et al., 2011; Anderson, Oxenham, Nelson, & Nelson, 2012; Berenstein et al., 2008; Drennan et al., 2016; Gifford et al., 2014; Henry & Turner, 2003; Henry, Turner, & Behrens, 2005; Litvak, Spahr, Saoji, & Fridman, 2007; Saoji et al., 2009; Won et al., 2011). However, when interfering background sounds are introduced, the relationship between performance on a variety of ripple tests and speech perception is more ambiguous. Some studies with CI listeners have reported significant correlations between performance in spectral-ripple tests and speech understanding in noise (Won et al., 2007, 2011), but other studies have shown no significant relationship (Anderson et al., 2011) or have shown a significant relationship for some speech and noise combinations but not for others (Anderson et al., 2012; Berenstein et al., 2008).

The equivocal findings surrounding spectral resolution and speech-in-noise perception may reflect the evidence that speech understanding in noise is affected by more than just energetic masking, which occurs when one signal prevents another signal from being audible due to overlapping energy in the same time and frequency (Brungart, Simpson, Ericson, & Scott, 2001). Differences in listeners' spectral resolution may change the effects of background sounds on the speech signal. For example, Stone et al. showed that the inherent random amplitude fluctuations that are present in "steady-state" noise, such as the speech-shaped noise (SSN) that is frequently used in speech experiments, have a large and detrimental effect on speech intelligibility for NH listeners, here referred to as modulation masking (Stone, Füllgrabe, Mackinnon, & Moore, 2011; Stone, Füllgrabe, & Moore, 2012). In contrast to NH subjects listening to unprocessed speech, modulation masking is absent or reduced for CI listeners and for NH individuals listening to spectrally smeared vocoded speech (Oxenham & Kreft, 2014). A possible explanation for this finding is that poor spectral resolution has the effect of smoothing random temporal-envelope fluctuations, thereby decreasing their negative impact and improving speech perception in SSN (Oxenham & Kreft, 2014).

Another form of modulation masking, here referred to as modulation interference, occurs when background sounds with slower amplitude modulations in the

temporal domain interfere with a listener's ability to process the syllabic structure of speech. Nelson, Jin, Carney, and Nelson (2003) showed evidence of this effect when they reported that CI subjects and NH subjects listening to vocoded speech experienced interference from noise maskers gated at rates of 2 to 4 Hz. A follow-up study by Nelson and Jin (2004) suggested that CI listeners' reliance on temporal-envelope cues due to reduced spectral resolution makes them more susceptible to envelope disruptions caused by gated noise signals. Modulation interference has been demonstrated in NH listeners as well, particularly in conditions with reduced redundancy of speech information due to limited spectral and temporal cues (Kwon & Turner, 2001) and when modulation is introduced on top of a steady-state masker without inherent fast temporal fluctuations (Stone & Moore, 2014).

In addition to peripheral effects like modulation masking and modulation interference, differences in spectral resolution may also influence higher-level auditory processes. For example, informational masking can occur when a competing talker interferes with a target talker due to difficulty perceptually segregating the two sources, even when audibility of the target signal is sufficient. For NH listeners, informational masking is greatest when the competing talker and the target talker share similar characteristics, such as gender (Brungart, 2001; Brungart et al., 2001). CI listeners have greater difficulty distinguishing between talkers due to reduced spectral resolution and other complicating factors such as impaired fundamental frequency (F0) processing (Stickney, Assmann, Chang, & Zeng, 2007), likely resulting in increased informational masking. Consistent with this hypothesis, Qin and Oxenham (2003) showed that speech reception in noise was poorer with a competing talker than with SSN for vocoded sounds, but the opposite result was found for unprocessed sounds. Stickney, Zeng, Litovsky, and Assmann (2004) showed similar effects in vocoder simulations and in CI subjects.

Taken together, these findings point to a variety of possible effects of reduced spectral resolution on speech understanding in the presence of interfering backgrounds, suggesting that impaired spectral sensitivity due to CI processing may alter the effects of different maskers on speech recognition. The purpose of this study was to examine the relationship between spectral resolution and sentence reception in varying backgrounds. The overall goal was to identify which types of maskers would be the most sensitive to differences in spectral resolution in light of the complex effects of modulation masking, modulation interference, and informational masking discussed earlier. Four different interferers were included in this study. SSN containing random envelope fluctuations served as a baseline

condition because it is commonly used in speech perception experiments. A multitone masker, modeled after Oxenham and Kreft (2014), was used to deliver the same amount of energy per channel as the SSN but with no inherent fluctuations. Four-talker babble was included as a more realistic masker that also probed the effects of modulation interference and had some potential for informational masking. Finally, a competing talker condition was used to assess both modulation interference and a greater potential for informational masking.

Within the overarching goal of measuring speech reception with different types of maskers, an additional question of interest was to quantify the relationship between spectral resolution and masking release (MR). MR occurs when a fluctuating interferer or competing talker causes less masking than stationary SSN at the same signal-to-noise ratio (SNR). Historically, NH listeners consistently demonstrate MR using unprocessed stimuli (Festen & Plomp, 1990; Fu & Nogaki, 2005; Nelson et al., 2003; Qin & Oxenham, 2003), while MR is absent, reduced, or even negative in CI listeners and vocoder simulations (Fu & Nogaki, 2005; Nelson et al., 2003; Qin & Oxenham, 2003; Stickney et al., 2004). Degraded spectral resolution in CI hearing may affect several aspects of speech perception and could be partially responsible for the lack of MR that has been established in the literature (Fu & Nogaki, 2005; Gnansia, Péan, Meyer, & Lorenzi, 2009; Qin & Oxenham, 2003). For example, reduced spectral resolution may decrease modulation masking, thus improving speech reception in the presence of SSN. Simultaneously, reduced spectral resolution may increase informational masking, thus degrading speech reception with a competing talker and leading to smaller differences between these two types of maskers. Another possible factor is that reduced spectral resolution leads to poorer baseline speech reception with the speech-shaped masker, resulting in subjects operating at positive SNRs where MR is less likely to occur (Bernstein & Brungart, 2011; Bernstein & Grant, 2009; Oxenham & Simonson, 2009). Due to these potential effects of spectral resolution on the target-interferer combinations that lead to MR, our hypothesis for the current study was that greater amounts of MR would be associated with better sensitivity to spectral modulations.

One way to explore different degrees of spectral resolution in a controlled manner is through the use of vocoded sounds (e.g., Bingabr, Espinoza-Varas, & Loizou, 2008). In this study, we used a tone vocoder with different amounts of simulated current spread to directly manipulate spectral resolution within a group of NH listeners (referred to as the vocoder group hereinafter). In a group of CI listeners, the same speech recognition task was conducted as in the vocoder group. However, current spread was not directly manipulated

in the CI listeners but was assumed to vary across subjects due to individual differences in factors such as electrode placement and stimulation levels. To measure spectral resolution, a spectral modulation detection (SMD) task was conducted with both groups of subjects.

In summary, this study examined the relationship between spectral resolution and speech understanding with various target-masker combinations for vocoder and CI subjects. In the vocoder group, simulated current spread was manipulated to measure how variations in spectral resolution affect differences in sentence reception across maskers. The difference between the SSN masker and the competing talker was explored to quantify the relationship between spectral resolution and MR. The results of this study may help guide the choice of maskers that are most sensitive to differences in spectral resolution when evaluating speech reception with new technologies, such as focused stimulation, that attempt to increase spectral resolution with CIs.

## Methods

This study included a sentence-reception task with four different types of maskers and a psychophysical task measuring spectral resolution. Two subject groups were tested: NH subjects listening to a tone-excited vocoder with simulated current spread and CI subjects. All subjects signed an approved (Western Institutional Review Board<sup>®</sup>) informed consent form covering all aspects of the research protocol.

## Subjects

The vocoder group consisted of 7 adult NH listeners (6 females and 1 male) who were between 24 and 27 years of age, with a mean age of 24.7 years. Hearing screenings were conducted prior to participation to confirm NH in both ears (thresholds of 25 dB HL or better across audiometric frequencies from 250 to 8000 Hz). The CI group included 20 subjects (7 females and 13 males) implanted with Nucleus<sup>®</sup> devices. The CI subjects were older than the vocoder subjects, with an age range of 39 to 85 years and a mean age of 61.9 years. One additional CI subject was included in the original cohort but was excluded due to difficulty with the tasks. Subject demographics for the CI group are displayed in Table 1. Seventeen subjects used the default Advanced Combination Encoder (ACE) coding strategy, and three subjects used a low-rate research coding strategy called Fundamental Asynchronous Stimulus Timing (FAST), in which the stimulation rate varies and follows the signal F0 of voiced speech and other harmonic sounds. Although some CI subjects were bilaterally implanted or used a contralateral hearing aid for everyday listening, subjects were tested in a unilateral,

**Table 1.** CI Subject Demographics.

Subject	Age	Gender	Ear tested	Device	Strategy	Bilateral CI	Contralateral hearing aid	Duration of deafness (years)	Duration of CI use (years)	Etiology
R94	56	M	R	CI512	ACE	No	Yes	10	5	Unknown
R96	75	M	R	CI512	ACE	No	No	2	5	Genetic
R47	61	M	L	CI24RE (CA)	FAST	Yes	No	0.5	10	Meniere's
R119	45	M	L	CI512	ACE	Yes	No	1	4	Genetic
R56	69	F	L	CI24RE (CA)	ACE	No	Yes	5	8	Unknown
R116	83	M	L	CI24RE (CA)	ACE	Yes	No	7	3	Unknown
R04	68	M	L	CI24R (CA)	ACE	Yes	No	1	11	Otosclerosis
R127	53	F	L	CI512	ACE	Yes	No	17	4	Unknown
R45	53	F	L	CI24R (CA)	ACE	Yes	No	0.5	11	Cogan's
R125	52	M	L	CI422	ACE	No	No	2	2	Sudden SNHL
R15	76	F	L	CI24R (CA)	ACE	No	No	1	11	Unknown
R95	85	M	L	CI512	ACE	No	Yes	10	5	Failed middle ear implant
R19	79	M	R	CI24R (CA)	ACE	No	No	5	11	Noise exposure
R87	71	M	L	CI24RE (CA)	ACE	No	Yes	3	6	Unknown
R59	48	F	R	CI24RE (CA)	ACE	No	Yes	8	8	Muckle–Wells
R134	54	F	R	CI422	ACE	Yes	No	52	2	Malformation
R17	54	M	R	CI24R (CS)	ACE	No	No	1	12	Genetic
R76	66	M	R	CI24RE (CA)	ACE	Yes	No	10	7	Infections
R68	51	M	L	CI24RE (CA)	FAST	Yes	No	17	7	Unknown
R129	39	F	R	CI24RE (CA)	FAST	No	Yes	3	4	Otosclerosis

Note. ACE = Advanced Combination Encoder; CI = cochlear implant; FAST = Fundamental Asynchronous Stimulus Timing; SNHL = sensorineural hearing loss.

implant-only condition for this study. Bilaterally implanted subjects used their preferred ear for testing.

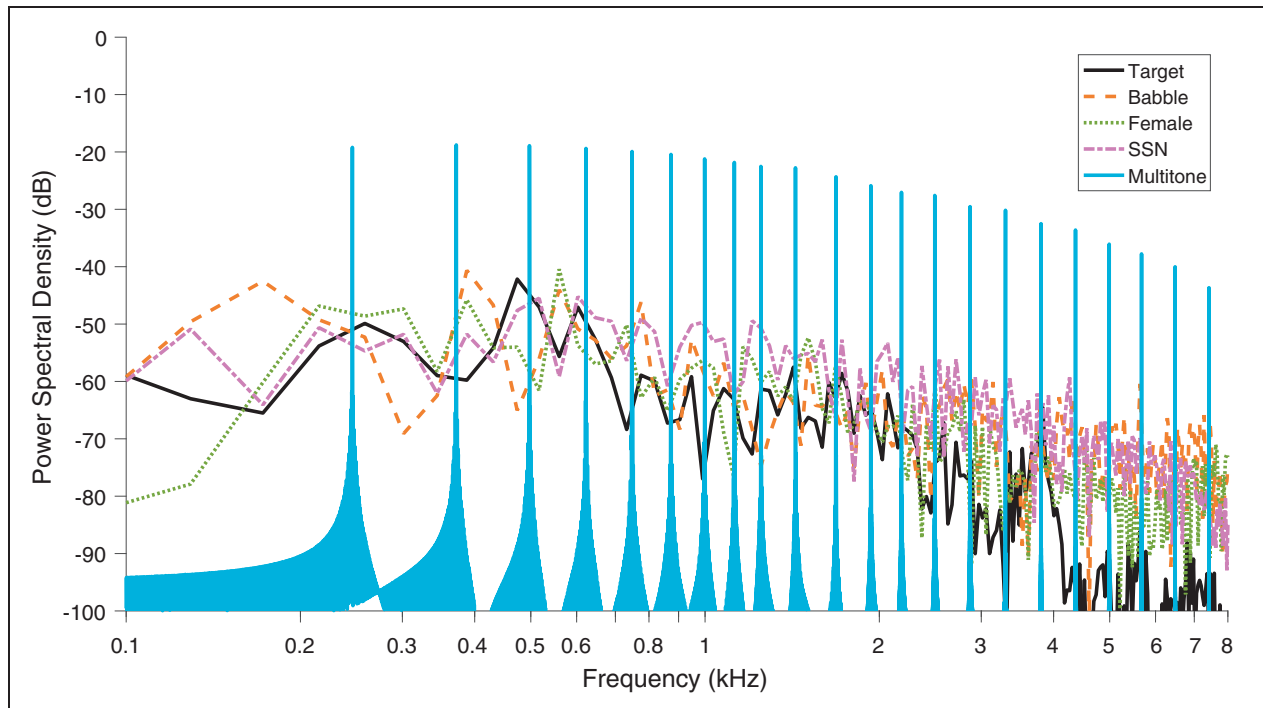
### Stimuli and Procedures

**Speech reception threshold.** A speech reception threshold (SRT) was measured in four different target-interferer conditions using the AuSTIN software (Dawson, Hersbach, & Swanson, 2013). In each condition, the level of the target sentence was fixed at 65 dB SPL, and the level of the masker varied according to the Hearing in Noise Test adaptive rule (Nilsson, Soli, & Sullivan, 1994) with a large step size of 4 dB and a small step size of 2 dB. The SRT was defined as the SNR in dB that provided 50% correct for word identification and was estimated by fitting a logistic function.

The target speech for the SRT test comprised Bench–Kowal–Bamford (BKB) sentences (Bench, Kowal, & Bamford, 1979) with a male talker, taken from the BKB-SIN CD (Etymotic Research, 2005). The average F0 of the target male talker was 118.5 Hz. Target speech was always presented with one of the four different maskers: four-talker babble (“Babble”), female competing talker (“Female”), multitone noise (“Multitone”), or

speech-shaped noise (“SSN”). The masker started 3 s prior to the target speech and ended 0.3 s after the target speech. One complete BKB sentence list pair, with a total of either 16 or 20 sentences, was presented in each condition. BKB lists and masker types were presented in random order across subjects.

The four-talker babble was taken from the BKB-SIN CD, and paired sentence and babble tracks were played together to maintain list equivalency. The female competing talker was taken from the Starkey open-access masker stimuli (Starkey Hearing Technologies, 2013) and used the Rainbow Passage for source material (Fairbanks, 1960). The average F0 of the Female talker was 211.0 Hz. For each trial, a random section of the competing-talker file was selected and presented with the target sentence. The SSN was the calibration noise included on the BKB-SIN CD. Multitone noise was constructed according to Oxenham and Kreft (2014). Briefly, a pure tone was generated at the center frequency of each CI or vocoder analysis channel. For CI subjects, the tones matched the center frequencies used in each subject's individual map (from 19 to 22 tones). For vocoder subjects, the tones matched the default center frequencies used in the 22-channel vocoder (described in detail later).



**Figure 1.** Long-term spectra of the four different maskers and the target speech. SSN = speech-shaped noise.

The level of each tone of the Multitone noise was set to match the frequency-specific long-term energy of the SSN at the output of the vocoder or CI envelope detection block for the corresponding channel. The Female talker and Multitone noise were both root mean square equalized to the level of the SSN. Figure 1 shows the long-term spectra of the four noise types and the target speech.

**Spectral modulation detection.** To evaluate the relationship between psychophysical spectral resolution and speech recognition with the four different maskers, an SMD task was conducted with both groups of subjects. The carrier stimulus was a pink-shaped noise spanning 6 octaves from 120 to 7680 Hz. Spectral modulations were introduced in the frequency domain at a ripple density of 0.5 ripples/octave. Figure 2 shows the amplitude spectra of example reference and target stimuli. The test was a cued four-interval three-alternative forced choice task. In each trial, the unmodulated reference was presented first, followed by two unmodulated stimuli and one modulated (target) stimulus in random order. The subject's task was to identify the target stimulus using a graphical user interface. Ripple phase was randomized across trials. Stimuli were presented at a base level of 50 dB SPL and were roved at a level of  $\pm 3$  dB. Ripple depth was varied in a 2-down 1-up adaptive procedure with a large step size of 1.0 dB (2 reversals) and a small step size of 0.5 dB (6 reversals) on a  $\log_2$

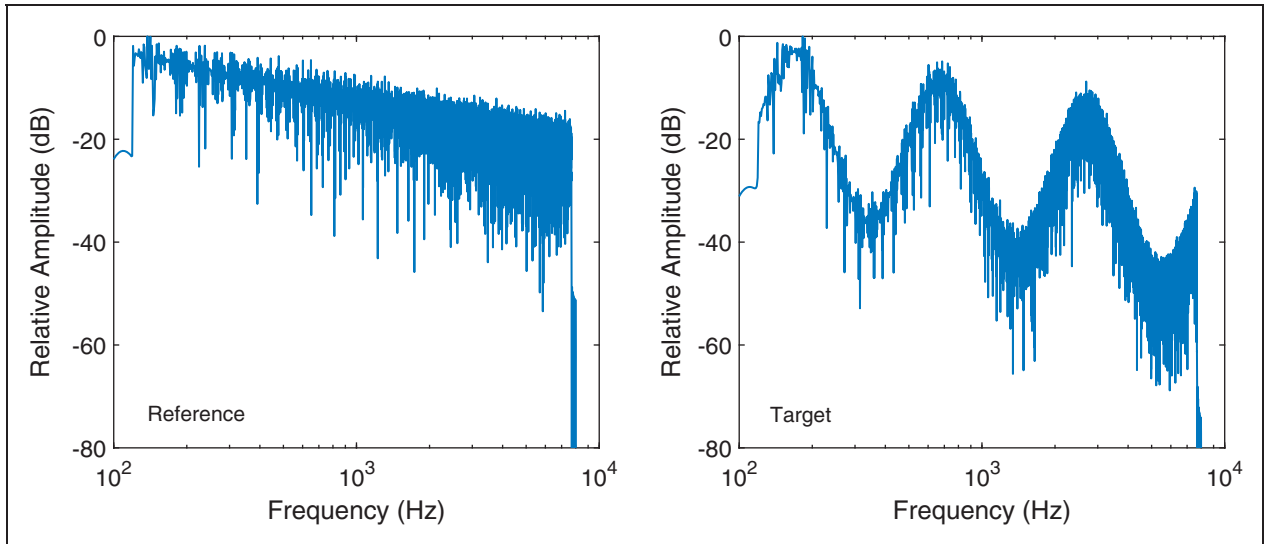
(dB ripple depth) scale. The average ripple depth of the last four reversals was taken to estimate the 71% correct point on the psychometric function (Levitt, 1971). Each CI listener completed one practice run of the SMD test and then finished four test runs. Vocoder subjects completed 12 SMD runs (four runs in each of the vocoder current spread conditions, as described later), in addition to one practice run using the condition with the least simulated current spread. The average of the four runs was taken for every condition.

To analyze the SMD data, ripple depth in linear amplitude was converted to modulation depth as defined by Equation 1:

$$md = 20 \times \log_{10} \left( \frac{p_{lin} - v_{lin}}{p_{lin} + v_{lin}} \right) \quad (1)$$

where  $md$  is modulation depth in dB,  $p_{lin}$  is the peak of the modulation in linear amplitude, and  $v_{lin}$  is the valley of the modulation in linear amplitude. Threshold modulation depth in dB was used for all further analysis.

**Vocoder.** For the vocoder group, all sounds were processed by a 22-channel tone-excited vocoder. The vocoder used a 128-point fast Fourier transform filterbank closely replicating the Advanced Combination Encoder filterbank, and the channel frequency boundaries matched clinical defaults for Nucleus<sup>®</sup> processors. For each subband, the Hilbert envelope was extracted.



**Figure 2.** Example reference and target stimuli used for the SMD task.

To vary spectral resolution for each listener, different degrees of cochlear current spread were modeled by convolving channel envelopes with a current spread decay function dropping off at  $-2$ ,  $-4$ , or  $-8$  dB/channel prior to modulating tones at the respective center frequencies and finally summing across all channels. The  $-2$  dB/channel condition represented the greatest amount of simulated current spread (poorest spectral resolution), and the  $-8$  dB/channel condition represented the least amount of simulated current spread (best spectral resolution). The simulated current spread conditions were selected based on pilot experiments. For speech testing, speech and masker signals were combined at the desired SNR prior to vocoder processing. All vocoder tests were conducted in all three current spread conditions, in order from the least difficult to the most difficult level of simulated current spread.

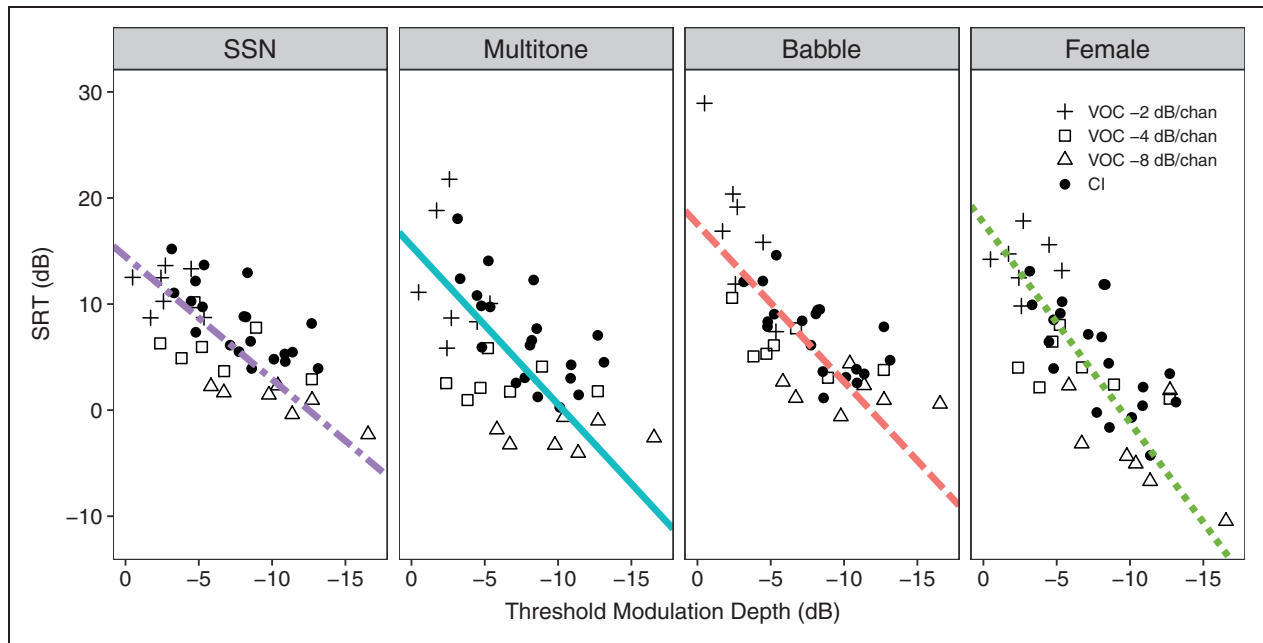
**Apparatus.** For CI subjects, stimuli were presented in the sound field in a unilateral, implant-only condition (i.e., no hearing aids were worn in the contralateral ear). Subjects with bilateral CIs used their preferred ear ( $N=9$ ). Testing was conducted using a lab-owned CP810 sound processor programmed with the subject's clinical map in a standard microphone-directionality program with no advanced input signal processing. The volume and sensitivity settings of the sound processor were fixed across tests. Sounds were routed from a desktop computer to an external EDIROL UA-1000 USB audio interface and presented via a Magnepan MMGW loudspeaker. The subject was seated 1 m from the speaker at  $0^\circ$  azimuth. For vocoder subjects, sounds were routed from a laptop computer to an external Mbox USB audio interface and presented diotically

over Sennheiser HD555 headphones. All testing was conducted in a sound-treated booth.

## Results

Individual SRT and SMD results are shown in Figure 3, with SRT plotted as a function of SMD threshold. In each subplot, each filled circle represents one CI subject. For vocoder subjects, different symbols are used to represent the three different current spread conditions. Plotted regression lines were fit to all subjects and conditions within each masker. To determine whether speech reception was significantly correlated with psychophysical spectral resolution across subjects, the relationship between SRT and SMD was analyzed with a linear mixed-effects (LME) model. The LME model was tested with the *lme4* package version 1.1-13 (Bates, Maechler, Bolker, & Walker, 2015) in R version 3.4.3 (R Core Team, 2017). With SRT as the outcome variable, masker type and SMD threshold were entered into the LME model as fixed effects along with their interaction. Group was also included as a fixed effect to explore differences between the vocoder and CI subjects, and age was included as a fixed control variable. Subject was entered as a random effect with varying intercepts and slopes. The full LME model accounted for 79.7% of the variance in SRT scores.

To test for statistical significance,  $p$  values were obtained using the Kenward–Roger approximation (Halekoh & Højsgaard, 2014) to compare the full model with the effect of interest against a small model without the effect of interest. Across groups, the effect of SMD was significant,  $F(1, 151.2) = 213.79$ ,  $p < .001$ , such that listeners with better SMD performance had better

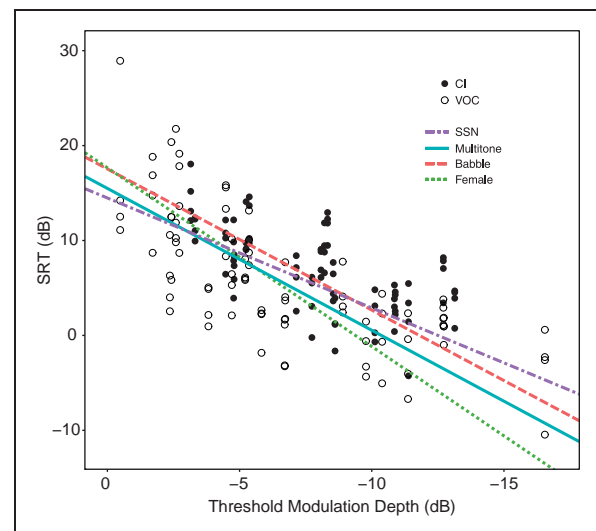


**Figure 3.** Speech reception in the four different masker conditions as a function of SMD threshold for vocoder and CI subjects. Each vocoder subject has three points plotted in each panel to represent the different simulated current spread conditions. Each CI subject has one point per masker type.

CI = cochlear implant; SRT = speech reception threshold; SSN = speech-shaped noise; VOC = vocoder.

sentence reception. The effect of masker type was also significant,  $F(3, 18.0) = 4.35$ ,  $p = .018$ , indicating that subjects performed better with some maskers than with others. The average SRT for Babble was 7.87 dB, for SSN was 7.11 dB, for Multitone noise was 5.56 dB, and for the Female talker was 4.97 dB. In addition to the significant main effects, the interaction between SMD and masker type was also significant,  $F(3, 52.3) = 3.63$ ,  $p = .019$ , suggesting that SRT differences across the four maskers were significantly associated with performance on the psychophysical test of spectral resolution. Post hoc comparisons using the Tukey test showed that the Female competing talker had a significantly steeper slope than the SSN,  $t(1, 53.2) = 3.35$ ,  $p = .008$ , but no other differences in slopes were significant. Figure 4 shows the slopes all four masker types plotted across both groups of subjects. The difference between the CI and vocoder group was not significant,  $F(1, 27.0) = 0.14$ ,  $p = .712$ , suggesting that the vocoder conditions included in this study were generally appropriate for modeling CI performance. In addition, the effect of age was not significant,  $F(1, 30.7) = 0.45$ ,  $p = .510$ .

The main LME analysis showed a significant interaction between masker type and SMD. One of the aims of this study was to quantify SRT differences across maskers and investigate how those differences may be related to spectral resolution. Therefore, three additional LME analyses were conducted using SSN as the baseline condition and exploring differences between



**Figure 4.** Speech reception as a function of SMD threshold, with regression lines plotted for the four different maskers. Vocoder subjects are shown in open symbols, and CI subjects are shown in filled symbols. In the vocoder group, each subject has three points for the three different simulated current spread conditions. CI = cochlear implant; SRT = speech reception threshold; SSN = speech-shaped noise; VOC = vocoder.

SSN and the other maskers. In each analysis, the outcome measure was the SRT difference in dB between SSN and the masker of interest. Each model included SMD and group as fixed effects along with their

interaction, and subject was included as a random effect with varying intercepts. Age was included as a control variable in each model. Tests for statistical significance were conducted in the same manner as the main analysis described earlier.

Figure 5 shows the results of the follow-up analyses. The first follow-up analysis considered the difference between SSN and the Multitone noise, displayed in the top two panels of Figure 5, with vocoder subjects shown on the left and CI subjects shown on the right. The first analysis accounted for 53.4% of the variance in the SSN–Multitone difference and showed a significant effect of SMD,  $F(1, 158.8) = 18.55$ ,  $p < .001$ , while the effect of Group,  $F(1, 24.1) = 0.47$ ,  $p = .501$ , and the interaction,  $F(1, 32.5) = 0.19$ ,  $p = .667$ , were not significant. In the middle row, the SRT difference between the SSN and Babble maskers is plotted against SMD. The second analysis revealed significant effects of SMD,  $F(1, 120.9) = 16.76$ ,  $p < .001$ , and Group,  $F(1, 26.0) = 6.75$ ,  $p = .015$ , but the SMD by Group interaction was not significant,  $F(1, 46.1) = 3.08$ ,  $p = .086$ . This model accounted for 39.2% of the variance. The final analysis examined the difference between SSN and the Female competing talker, also known as the amount of MR, as shown in the bottom row of Figure 5. This model accounted for 85.7% of the variance, with a significant effect of SMD,  $F(1, 151.0) = 249.77$ ,  $p < .001$ . Although the overall Group difference was not significant,  $F(1, 23.9) = 2.84$ ,  $p = .105$ , the interaction between SMD and Group was significant,  $F(1, 26.4) = 4.92$ ,  $p = .035$ , suggesting that with decreasing SMD thresholds, MR increased more for vocoder listeners than for CI listeners. Despite this significant interaction between groups, the correlation between MR and SMD was still significant when tested in the CI group alone ( $r = -.30$ ,  $p = .006$ ).

While the earlier analysis shows a significant relationship between MR and SMD, indicating increasing MR with better SMD thresholds, another possible factor affecting the amount of MR is the baseline SRT at which each subject is operating (measured with the SSN masker). To further illustrate the effects of MR and its possible relationship to baseline SRT, Figure 6 shows SRTs for the SSN and Female maskers for individual vocoder and CI subjects. In each panel, subjects are rank-ordered according to their baseline SRT, where lower thresholds indicate better performance. The length and style of each line connecting the SRTs for the two maskers indicates the relationship between the maskers for each subject. A solid line is used when the Female SRT is lower than the SSN SRT, demonstrating MR, and a dashed line is used in the opposite situation, demonstrating negative MR. For the vocoder group, separate repeated-measures analyses of variance revealed that simulated current spread had a significant effect on

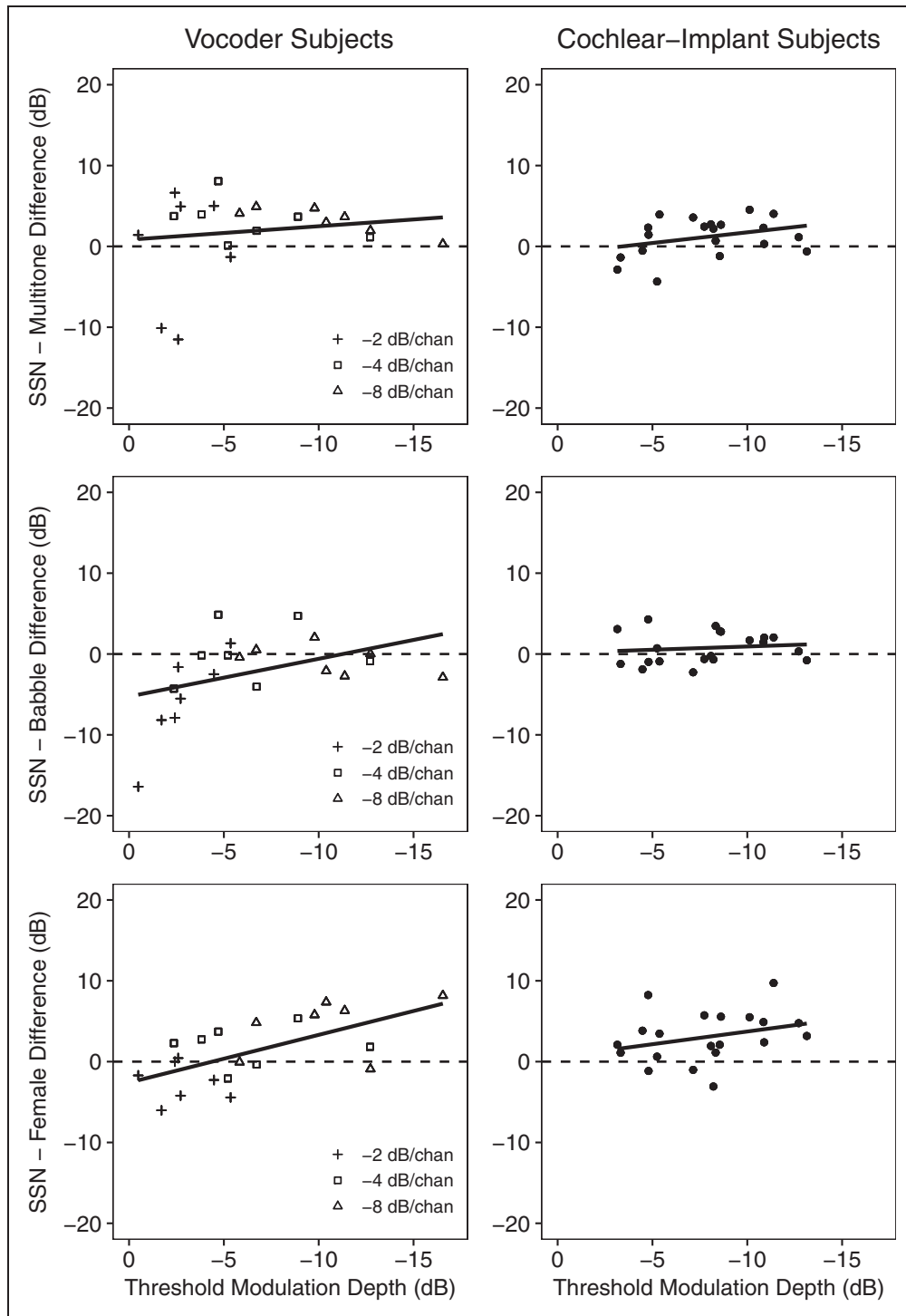
both sentence reception (controlling for masker type and its interaction with current spread),  $F(2, 12) = 451.4$ ,  $p < .001$ , and SMD threshold,  $F(2, 12) = 50.8$ ,  $p < .001$ . The effect of reduced simulated current spread on sentence reception can be seen in the left panel of Figure 6 as a decrease in baseline SRT. Five out of seven vocoder subjects showed negative MR in the condition with the greatest simulated current spread. As the baseline SRT decreases, the amount and consistency of MR increases so that five out of seven vocoder subjects showed substantial MR in the condition with the least simulated current spread. Because baseline SRT and SMD threshold both improve with reduced simulated current spread, it is possible that the correlations observed between the magnitude of MR and SMD might reflect the different SNRs at which the vocoder conditions were tested, rather than solely the effects of spectral resolution.

The right panel of Figure 6 shows that CI subjects who were the best performers, with baseline SRTs below about 5 dB, had relatively large and consistent MR. This finding agrees with the vocoder data at low SNRs. However, a great deal of variability was observed for CI subjects with baseline SRTs above 5 dB. Some subjects showed little to no MR, while other subjects showed large MR despite having higher baseline SRTs. In contrast to the vocoder group, only three of the CI subjects showed any amount of negative MR. This finding was surprising given the consistent effects of informational masking and modulation interference demonstrated in the vocoder group with the largest amount of simulated current spread.

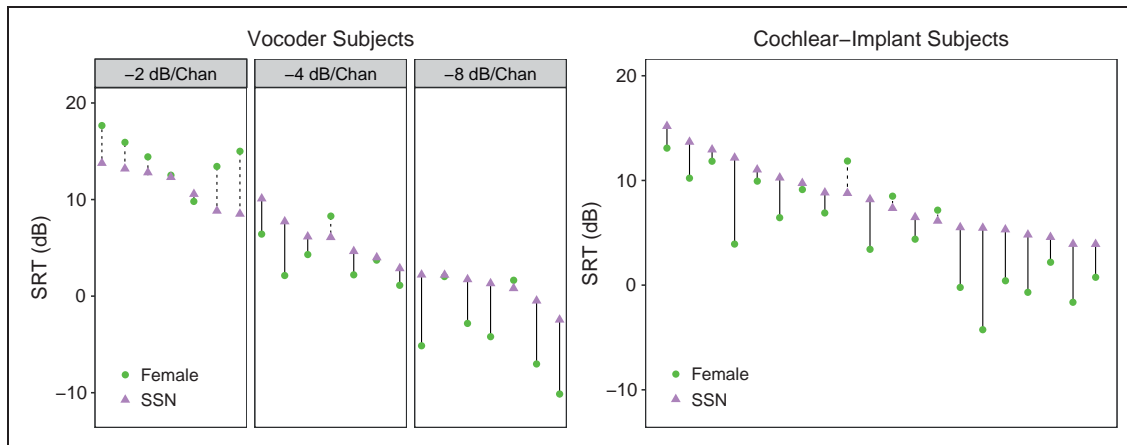
## Discussion

The results of this study demonstrate that sentence reception improves with improved SMD sensitivity for four different types of interferers across CI and vocoder subjects. This finding is consistent with a number of previous studies that have shown significant correlations between spectral-ripple tests and speech recognition in noise in the CI population (e.g., Won et al., 2007, 2011). However, the relationship between speech reception and SMD varied across masker types. Performance in the SSN masker condition was the least dependent on spectral resolution, followed by the Multitone noise and Babble, while the Female condition was the most dependent on spectral resolution. Interestingly, when previous studies have reported no significant relationship between spectral-ripple tests and speech recognition in noise, the masker used was frequently SSN (Anderson et al., 2011, 2012; Berenstein et al., 2008), suggesting that perhaps a significant relationship may have been less elusive with a different type of interferer. Indeed, Berenstein et al. (2008) showed significant correlations between spectral-ripple performance and speech reception in





**Figure 5.** Differences in speech reception between the baseline SSN masker and the three other interferers, plotted against SMD threshold. The difference between SSN and the Multitone noise is shown in the top row, the difference between SSN and Babble is shown in the middle row, and the difference between SSN and the Female competing talker is shown in the bottom row. The difference between SSN and Female represents the effects of MR. SSN = speech-shaped noise.



**Figure 6.** MR for individual subjects. Data points show SRTs for the SSN and Female maskers and are rank-ordered according to each subject's baseline SRT with the SSN masker. The length and style of the line connecting the SSN data point to the Female masker data point represents the degree of MR (solid lines) or negative MR (dashed lines). Results from the vocoder subjects are shown for each of the three simulated current spread conditions (left panel), whereas there is only a single SRT pair of data for each CI subject (right panel). SRT = speech reception threshold; SSN = speech-shaped noise.

fluctuating noise, while the correlations for speech reception in SSN were not significant. Potentially, the type of masker may have a role in determining whether or not significant relationships are found between spectral-ripple tests and speech understanding in competing backgrounds.

The differences in the relationship between SRT and SMD across maskers may be explained, at least in part, by the effects of modulation masking, modulation interference, and informational masking. In the vocoder group, spectral resolution was intentionally degraded via simulated current spread. This likely had multiple effects, including poorer SMD performance as well as smoothing of the random temporal-envelope fluctuations in the SSN masker. With reduced inherent noise fluctuations, modulation masking would also have been reduced, causing a trade-off between the negative effects of poorer spectral resolution and the positive effects of reduced modulation masking (Oxenham & Kreft, 2014; Stone et al., 2011, 2012). This trade-off may be responsible for the relatively shallow slope of the SSN compared with the other maskers (see Figure 4). Although spectral resolution was not directly manipulated in the CI group, the similar SRT values between vocoder and CI subjects suggest that differences in SMD capture a comparable trade-off between groups.

The Babble interferer and Female competing talker are distinct from the SSN and Multitone noise in that they both contain speech and therefore have slower amplitude modulations that may cause modulation interference. Due to reduced spectral resolution, CI subjects and vocoder subjects may be more susceptible to temporal-envelope disruptions (Kwon & Turner, 2001; Nelson et al., 2003; Nelson & Jin, 2004).

Modulation interference is one possible contributor to this reduced speech recognition with both types of speech maskers. In addition to modulation interference, the Female competing talker was also the most likely masker to elicit informational masking (Brungart, 2001). Subjects may have had more difficulty segregating the target talker from the competing talker when spectral resolution was poorer, resulting in greater informational masking. Potentially, the combined effects of modulation interference and informational masking led to the Female talker being the most sensitive masker to differences in spectral resolution.

Another way of examining variation across masker types is to measure the difference between the baseline performance with SSN and each of the other interferers as a function of SMD. This analysis showed a significant relationship for all three of the alternative maskers (Multitone, Babble, and Female), providing further evidence that other types of interferers beyond the commonly used SSN should be considered for studies investigating spectral resolution with CI subjects. Of the other three masker types, the difference between the SSN and the Female competing talker showed the strongest relationship with spectral resolution, accounting for the largest proportion of the variance across the three follow-up analyses. This not only suggests that the competing talker was the most sensitive masker, but it also addresses the second question investigating the relationship between spectral resolution and MR (defined as the SRT difference between the SSN and Female maskers).

Consistent with the hypothesis, MR was significantly correlated with spectral resolution across subjects. This finding agrees with previous results showing greater MR in vocoder simulations with better spectral resolution

and decreasing MR with poorer spectral resolution (Fu & Nogaki, 2005; Gnansia et al., 2009; Qin & Oxenham, 2003). In addition, a significant interaction showed that the vocoder group gained larger MR than the CI group did for a given amount of SMD improvement. At the worst levels of spectral sensitivity, vocoder subjects demonstrated poorer performance with the competing talker than with the SSN. This result is consistent with previous studies showing negative MR for vocoder and CI subjects, likely due to the effects of informational masking and modulation interference (Qin & Oxenham, 2003; Stickney et al., 2004). Surprisingly, very few subjects in the CI group demonstrated negative MR. Despite the expectation that CI subjects would have difficulty differentiating between the target and the competing talker, MR was observed not only for subjects with good spectral resolution but also for several CI subjects with relatively poor spectral resolution. This finding differs from previous studies showing reduced or absent MR in CI subjects (Fu & Nogaki, 2005; Nelson et al., 2003; Stickney et al., 2004). Overall, the competing talker was the least effective masker, which is consistent with the large amount of MR observed in the CI group. The spectral differences across maskers may also have contributed to differences in overall difficulty, particularly because the Female masker had less energy at low and high frequencies compared with the SSN.

Several other factors might help explain differences between this study and earlier CI data as well as differences between the vocoder and CI groups. Advancements in CI technology over time may have improved general performance on speech and psychophysical tasks, leading to MR in the current CI group when it has been absent previously. In addition, the vocoder conditions included in this study captured a slightly wider range of performance compared with the current sample of CI subjects, which could have led to greater increases in MR with improved spectral sensitivity and more negative MR with poorer spectral sensitivity. Another possibility is that the vocoder simulation of current spread mimics the SRT values of the CI group but does not adequately match the MR effects. Potentially, CI listeners with poorer SRTs may be more affected by factors other than current spread, such as limited neural survival, that may not affect MR to the same extent as speech perception, leading to greater MR at higher SRTs than the vocoder group. Future research could consider this possibility. Finally, vocoder subjects were younger than CI subjects, had NH, and were not experienced listeners of spectrally reduced speech. Although vocoder and CI data differed in some respects, the overall similarity of results suggests that modeling current spread with a tone vocoder is a useful tool for exploring differences in spectral resolution that might apply to CI technologies.

An additional consideration that is important when interpreting these results is the relationship between MR and the baseline SNR of the SSN. Previous studies with NH and hearing-impaired listeners have shown that MR decreases with increasing SNR, and little to no MR occurs for baseline SNRs above 0 dB (Bernstein & Brungart, 2011; Bernstein & Grant, 2009; Oxenham & Simonson, 2009). Thus, MR may be confounded with SNR, such that differences across processing conditions or subjects may depend on baseline performance. In one study examining this possibility, Bernstein and Brungart (2011) investigated the effects of spectral smearing on MR for NH listeners. At matched performance levels, the unprocessed speech produced significantly greater MR compared with the spectrally smeared speech, which is consistent with the current findings. However, when the baseline SNR was lowered for the spectrally smeared speech using a reduced word set size, the differences between the spectrally smeared and unprocessed conditions disappeared. Based on these findings, the authors postulated that the lack of MR traditionally observed in the hearing-impaired population is not due to impaired spectral processing per se, but rather due to the increased SNR necessary to operate at performance levels that are similar to NH listeners (Bernstein & Brungart, 2011). In this study, we did not control for SNR differences between different levels of simulated current spread for the vocoder subjects, so the correlation between MR and SMD might reflect the effects of SNR rather than spectral resolution. However, in both the vocoder and CI subject groups, baseline SRTs for the SSN masker were nearly all above 0 dB (see Figure 6), where MR is not expected to occur based on NH and hearing-impaired data (Bernstein & Brungart, 2011; Bernstein & Grant, 2009; Oxenham & Simonson, 2009). Therefore, although the baseline SNR is an important consideration for any study exploring MR, the current results are novel in that they show large MR for CI listeners operating at positive SNRs where vocoder subjects do not show MR.

The current study has demonstrated that the relationship between speech-in-noise understanding and spectral resolution differs based on the type of masker used. This finding may help elucidate why previous studies have reported inconsistent results regarding the relationship between speech understanding in noise and spectral-ripple tests. In addition, we have shown that some CI listeners exhibit substantial MR, contrary to previous reports, and that MR is significantly associated with spectral resolution, although the potential influence of baseline SNR cannot be ruled out. The current results suggest that an improvement in spectral resolution within an individual CI listener (e.g., via improved electrode placement or focused stimulation) might lead to lower SRTs and larger MR that more closely resemble

effects seen with NH listeners. When conducting future CI research on speech understanding in noise, the current findings highlight that the selection of appropriate background stimuli is an important experimental consideration, particularly for studies investigating the effects of spectral resolution.

### Acknowledgments

We thank Hannah Glick for assistance with data collection, Sara Duran and two anonymous reviewers for valuable comments on earlier versions of this article, and our research participants for their time and efforts.

### Authors' Note

Portions of this work were presented at the 2015 Conference on Implantable Auditory Prostheses, Lake Tahoe, CA and the 171st Meeting of the Acoustical Society of America, Salt Lake City, UT.

### Declaration of Conflicting Interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Naomi B. H. Croghan and Zachary M. Smith are employees of Cochlear Ltd.

### Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

### References

- Anderson, E. S., Nelson, D. A., Kreft, H., Nelson, P. B., & Oxenham, A. J. (2011). Comparing spatial tuning curves, spectral ripple discrimination, and speech perception in cochlear implant users. *Journal of the Acoustical Society of America*, *130*(1), 364–375. doi:10.1121/1.3589255
- Anderson, E. S., Oxenham, A. J., Nelson, P. B., & Nelson, D. A. (2012). Assessing the role of spectral and intensity cues in spectral ripple detection and discrimination in cochlear-implant users. *Journal of the Acoustical Society of America*, *132*(6), 3925–3934. doi:10.1121/1.4763999
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01
- Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bench-Kowal-Bamford) sentence lists for partially-hearing children. *British Journal of Audiology*, *13*(3), 108–112. doi:10.3109/03005367909078884
- Bernstein, C. K., Mens, L. H. M., Mulder, J. J. S., & Vanpoucke, F. J. (2008). Current steering and current focusing in cochlear implants: Comparison of monopolar, tripolar, and virtual channel electrode configurations. *Ear & Hearing*, *29*(2), 250–260. doi:10.1097/AUD.0b013e3181645336
- Bernstein, J. G. W., & Brungart, D. S. (2011). Effects of spectral smearing and temporal fine-structure distortion on the fluctuating-masker benefit for speech at a fixed signal-to-noise ratio. *Journal of the Acoustical Society of America*, *130*(1), 473–488. doi:10.1121/1.3589440
- Bernstein, J. G. W., & Grant, K. W. (2009). Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, *125*(5), 3358–3372. doi:10.1121/1.3110132
- Bingabr, M., Espinoza-Varas, B., & Loizou, P. C. (2008). Simulating the effect of spread of excitation in cochlear implants. *Hearing Research*, *241*(1–2), 73–79. doi:10.1016/j.heares.2008.04.012
- Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *Journal of the Acoustical Society of America*, *109*(3), 1101–1109. doi:10.1121/1.1345696
- Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *Journal of the Acoustical Society of America*, *110*(5), 2527–2538. doi:10.1121/1.1408946
- Dawson, P. W., Hersbach, A. A., & Swanson, B. A. (2013). An adaptive Australian sentence test in noise. *Ear & Hearing*, *34*(5), 592–600. doi:10.1097/AUD.0b013e31828576fb
- Drennan, W. R., Won, J. H., Timme, A. O., & Rubinstein, J. T. (2016). Nonlinguistic outcome measures in adult cochlear implant users over the first year of implantation. *Ear & Hearing*, *37*(3), 354–364. doi:10.1097/AUD.0000000000000261
- Etymotic Research. (2005). *The BKB-SIN speech-in-noise test, version 1.03* (Compact Disk). Elk Grove Village, IL: Author.
- Fairbanks, G. (1960). *Voice and articulation drillbook* (2nd ed., p. 127). New York, NY: Harper & Row.
- Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, *88*(4), 1725–1736. doi:10.1121/1.400247
- Fu, Q.-J., & Nogaki, G. (2005). Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing. *Journal of the Association for Research in Otolaryngology*, *6*(1), 19–27. doi:10.1007/s10162-004-5024-3
- Gifford, R. H., Hedley-Williams, A., & Spahr, A. J. (2014). Clinical assessment of spectral modulation detection for adult cochlear implant recipients: A non-language based measure of performance outcomes. *International Journal of Audiology*, *53*(3), 159–164. doi:10.3109/14992027.2013.851800
- Gnansia, D., Péan, V., Meyer, B., & Lorenzi, C. (2009). Effects of spectral smearing and temporal fine structure degradation on speech masking release. *Journal of the Acoustical Society of America*, *125*(6), 4023–4033. doi:10.1121/1.3126344
- Halekoh, U., & Højsgaard, S. (2014). A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models – The R package pbkrtest. *Journal of Statistical Software*, *59*(9), 1–30. doi:10.18637/jss.v059.i09
- Henry, B. A., & Turner, C. W. (2003). The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *Journal of the Acoustical Society of America*, *113*(5), 2861–2873. doi:10.1121/1.1561900

- Henry, B. A., Turner, C. W., & Behrens, A. (2005). Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *Journal of the Acoustical Society of America*, *118*(2), 1111–1121. doi:10.1121/1.1944567
- Kwon, B. J., & Turner, C. W. (2001). Consonant identification under maskers with sinusoidal modulation: Masking release or modulation interference? *Journal of the Acoustical Society of America*, *110*(2), 1130–1140. doi:10.1121/1.1384909
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, *49*(2B), 467–477. doi:10.1121/1.1912375
- Litvak, L. M., Spahr, A. J., Saoji, A. A., & Fridman, G. Y. (2007). Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *Journal of the Acoustical Society of America*, *122*(2), 982–991. doi:10.1121/1.2749413
- Nelson, P. B., & Jin, S.-H. (2004). Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners. *Journal of the Acoustical Society of America*, *115*(5), 2286–2294. doi:10.1121/1.1703538
- Nelson, P. B., Jin, S.-H., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *Journal of the Acoustical Society of America*, *113*(2), 961–968. doi:10.1121/1.1531983
- Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America*, *95*(2), 1085–1099. doi:10.1121/1.408469
- Oxenham, A. J., & Kreft, H. A. (2014). Speech perception in tones and noise via cochlear implants reveals influence of spectral resolution on temporal processing. *Trends in Hearing*, *18*, 1–14. doi:10.1177/2331216514553783
- Oxenham, A. J., & Simonson, A. M. (2009). Masking release for low- and high-pass-filtered speech in the presence of noise and single-talker interference. *Journal of the Acoustical Society of America*, *125*(1), 457–468. doi:10.1121/1.3021299
- Qin, M. K., & Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *Journal of the Acoustical Society of America*, *114*(1), 446–454. doi:10.1121/1.1579009
- R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from [www.R-project.org](http://www.R-project.org).
- Saoji, A. A., Litvak, L., Spahr, A. J., & Eddins, D. A. (2009). Spectral modulation detection and vowel and consonant identifications in cochlear implant listeners. *Journal of the Acoustical Society of America*, *126*(3), 955–958. doi:10.1121/1.3179670
- Smith, Z. M., Parkinson, W. S., & Long, C. J. (2013). Multipolar current focusing increases spectral resolution in cochlear implants. *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013*, 2796–2799. doi:10.1109/EMBC.2013.6610121
- Starkey Hearing Technologies. (2013). *Open access stimuli for the creation of multi-talker maskers*. Retrieved from [www.starkeypro.com/research/research-resources/open-access-stimuli](http://www.starkeypro.com/research/research-resources/open-access-stimuli).
- Stickney, G. S., Assmann, P. F., Chang, J., & Zeng, F.-G. (2007). Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences. *Journal of the Acoustical Society of America*, *122*(2), 1069–1078. doi:10.1121/1.2750159
- Stickney, G. S., Zeng, F.-G., Litovsky, R., & Assmann, P. (2004). Cochlear implant speech recognition with speech maskers. *Journal of the Acoustical Society of America*, *116*(2), 1081–1091. doi:10.1121/1.1772399
- Stone, M. A., Füllgrabe, C., Mackinnon, R. C., & Moore, B. C. J. (2011). The importance for speech intelligibility of random fluctuations in “steady” background noise. *Journal of the Acoustical Society of America*, *130*(5), 2874–2881. doi:10.1121/1.3641371
- Stone, M. A., Füllgrabe, C., & Moore, B. C. J. (2012). Notionally steady background noise acts primarily as a modulation masker of speech. *Journal of the Acoustical Society of America*, *132*(1), 317–326. doi:10.1121/1.4725766
- Stone, M. A., & Moore, B. C. J. (2014). On the near non-existence of “pure” energetic masking release for speech. *Journal of the Acoustical Society of America*, *135*(4), 1967–1977. doi:10.1121/1.4868392
- Won, J. H., Clinard, C. G., Kwon, S., Dasika, V. K., Nie, K., Drennan, W. R., . . . Rubinstein, J. T. (2011). Relationship between behavioral and physiological spectral-ripple discrimination. *Journal of the Association for Research in Otolaryngology*, *12*(3), 375–393. doi:10.1007/s10162-011-0257-4
- Won, J. H., Drennan, W. R., & Rubinstein, J. T. (2007). Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *Journal of the Association for Research in Otolaryngology*, *8*(3), 384–392. doi:10.1007/s10162-007-0085-8