Trends in Hearing
Volume 24: 1–11

DOI: 10.1177/2331216520948385
journals.sagepub.com/home/tia

Spectrotemporal Modulation Sensitivity in Cochlear-Implant and Normal-Hearing Listeners: Is the Performance Driven by Temporal or Spectral Modulation Sensitivity?

Ning Zhou[1], Susannah Dixon[1], Zhen Zhu[2], Lixue Dong[1], and Marti Weiner[1]

Abstract

This study examined the contribution of temporal and spectral modulation sensitivity to discrimination of stimuli modulated in both the time and frequency domains. The spectrotemporally modulated stimuli contained spectral ripples that shifted systematically across frequency over time at a repetition rate of 5 Hz. As the ripple density increased in the stimulus, modulation depth of the 5 Hz amplitude modulation (AM) reduced. Spectrotemporal modulation discrimination was compared with subjects' ability to discriminate static spectral ripples and the ability to detect slow AM. The general pattern from both the cochlear implant (CI) and normal hearing groups showed that spectrotemporal modulation thresholds were correlated more strongly with AM detection than with static ripple discrimination. CI subjects' spectrotemporal modulation thresholds were also highly correlated with speech recognition in noise, when partialing out static ripple discrimination, but the correlation was not significant when partialing out AM detection. The results indicated that temporal information was more heavily weighted in spectrotemporal modulation discrimination, and for CI subjects, it was AM sensitivity that drove the correlation between spectrotemporal modulation thresholds and speech recognition. The results suggest that for the rates tested here, temporal information processing may limit performance more than spectral information processing in both CI users and normal hearing listeners.

Keywords

spectrotemporal modulation, amplitude modulation detection, spectral ripple discrimination

Received 7 April 2020; revised 2 July 2020; accepted 17 July 2020

In modern cochlear implants (CIs), channel-specific slow amplitude variations (temporal envelope) of the acoustic signals are extracted and represented by amplitude modulation (AM) of constant-rate electrical pulse trains. The channel specificity of the envelope information is inherently poor, typically involving no more than 22 analysis bands. Temporal and spectral resolutions of CIs have been extensively studied as separate factors in CI users.

Numerous factors contribute to the coarse spectral resolution with CIs, and these factors can be anatomical, physiological, or related to surgical techniques. The electrode array is often not placed in ideal locations that facilitate place-specific stimulation of the auditory nerve. They can be at locations that are far from the central axis

---

[1]Department of Communication Sciences and Disorders, East Carolina University, Greenville, North Carolina, United States
[2]Department of Engineering, East Carolina University, Greenville, North Carolina, United States

Corresponding Author:
Ning Zhou, Department of Communication Sciences and Disorders, East Carolina University, 3310Y Health Sciences Building, Greenville, NC 27834, United States.
Email: zhoun@ecu.edu

of the modiolus due to abnormal anatomy or the design of the electrode array (Kawano et al., 1998; O'Connell et al., 2016). Some of the electrodes can penetrate to enter scala media or scala vestibuli (Finley et al., 2008). If the electrodes are in the right location, it is possible that there are not enough viable nerve fibers to be stimulated (reduced excitability or missing dendrites; Kawano et al., 1998; Nadol et al., 2011). These factors would result in a stimulation pattern in which the individual electrodes stimulate overlapping populations of neurons, and thus the channel-specific envelope information becomes smeared across spectral regions. Spectral resolution has been a major limiting factor for the success of the neural prosthesis and has been studied extensively in CI users (e.g., Friesen et al., 2001). The techniques for estimating spectral resolution of local regions include electrode discrimination (Busby & Clark, 2000; Zwolan et al., 1997), loudness summation (McKay et al., 2001), and the more complex but direct measures of psychophysical tuning curves (Nelson et al., 2011).

Global measures of spectral resolution involve discriminating or detecting spectral modulations in wideband signals (Henry & Turner, 2003; Won et al., 2007). In some of these tests, the upper limit of spectral resolution is measured for CI subjects, where the number of spectral ripples per octave (RPO; ripple density) increases until subjects can no longer discriminate stimuli of different ripple densities or stimuli of inverted ripple phases (e.g., Won et al., 2007). Alternatively, the ripple density is fixed, and the minimum modulation depth required to detect the spectral ripple from an unmodulated reference is measured (e.g., Gifford et al., 2018; Landsberger et al., 2019). The measures of ripple discrimination and detection have been shown to correlate with each other in CI listeners (Anderson et al., 2012). Previous research has shown a rather robust relationship between spectral ripple thresholds and outcomes with CIs, providing evidence that spectral resolution is important for CI function. The results in pediatric CI users are not conclusive (Gifford et al., 2018), but a consistent relationship has been found between ripple discrimination or detection and speech recognition in quiet and in noise (Dorman et al., 2012; Drennan et al., 2014; Gifford et al., 2014; Saoji et al., 2009; Won et al., 2007) as well as perception of music in adult CI listeners (Won et al., 2010). Ripple discrimination has also been shown to be sensitive to manipulations that aim to reduce channel interaction such as current steering (Berenstein et al., 2008), current focusing (Drennan et al., 2010), and electrode deactivation (Zhou, 2017).

Previous research has raised questions of the extent to which the ripple tests measure CI listeners' ability to perceive the spectral shape of the stimulus (McKay et al., 2009). It has been argued that local intensity cues such as level of the spectral edges, which changes depending on the starting phase of the spectral modulation, or loudness differences between the phase-inverted ripple stimuli in single channels can help CI users perform the task but the thresholds do not necessarily reflect their spectral resolution. A number of groups have investigated the potential contribution of these factors by level roving (Won et al., 2011), modeling (Won et al., 2011), applying a Gaussinan-shaped spectral envelope to the stimuli (Supin et al., 1998), or using shallow spectral slopes on the spectral edges (Anderson et al., 2011). Results of these studies support the idea that the spectral ripples were not detected via local loudness cues but rather the thresholds reflected an across-channel processing of the spectral information. Anderson et al. (2011) also commented that the subjects have to have some spectral resolution for the intensity cues in single channels to be detected.

In contrast to the static nature of the spectral modulation used in these spectral resolution tests, some tests use spectral ripples that systematically vary over time across the frequencies. When spectral modulation changes over time, that is, gliding ripples, it creates AM in the signal. In some of these tests, subjects discriminate ripples that are of the same density but differ in the direction of the frequency sweep, that is, upward versus downward, as in Schroeder-phase discrimination (Drennan et al., 2008) or the spectrotemporal ripple for investigating processor effectiveness (STRIPES) test (Archer-Boyd et al., 2018). Importantly, in the upward or downward gliding ripples, the rate and modulation depth of the AM are identical in any given frequency region. If the task is to compare a target and reference that differ in the gliding ripple density, then subjects may use the difference in AM depth within channel to discriminate the target and reference, because modulation depth of the AM reduces with increasing ripple density. Spectrotemporal modulation sensitivity was also able to robustly predict CI users' speech recognition performance (Holden et al., 2016; Lawler et al., 2017). Mehraei et al. (2014) found that spectrotemporal modulation detection thresholds also predicted performance of hearing-impaired listeners in speech tasks. However, it is unclear whether it is the subjects' temporal modulation sensitivity or spectral resolution that contributes to the performance of the test. It is thus also unclear whether the correlation with speech recognition reflects the importance of AM sensitivity or spectral resolution for perceiving the speech stimuli. There is equally compelling evidence to indicate that AM sensitivity explains some variance in CI users' speech recognition performance (Fu, 2002; Garadat et al., 2012, 2013; Luo et al., 2008). There have been few studies that examined the relative importance of temporal modulation and spectral modulation sensitivity for discrimination/

detection of spectrotemporal modulation. Won et al. (2015) reported that detection of static spectral ripples in CI users was correlated with spectrotemporal modulation detection for stimulus of the same ripple density and 5 Hz AM but not for stimulus with higher AM rates. The contribution of AM detection to spectrotemporal modulation detection was weak, and, none of these correlations were significant after adjusting for multiple comparisons. Zheng et al. (2017) decomposed the spectrotemporal modulation transfer functions measured in CI users into the temporal and spectral dimensions and found that the joint spectrotemporal modulation sensitivity is heightened and cannot be predicted by the product of temporal and spectral modulation sensitivity. The authors speculated that the heightened joint sensitivity was a result of relying on the better and relatively intact AM sensitivity to overcome the poorer sensitivity for high ripple densities. Taken together, there has not been direct and strong evidence to show which sensitivity dominates the acuity to modulations covaried in both the temporal and spectral domains.

In this study, we examined the contribution of AM detection and static ripple discrimination to spectrotemporal modulation sensitivity using the spectrotemporally modulated ripple test (SMRT) developed by Aronoff and Landsberger (2013). If either AM detection or static ripple discrimination drives the spectrotemporal modulation threshold in CI users, it would be interesting to evaluate whether the response reflects a cue weighting behavior specific to the CI users, or if the same pattern extends to subjects receiving normal acoustic stimulation. To that end, we tested AM detection, static ripple discrimination using phase-inverted ripple stimuli, and spectrotemporal modulation sensitivity using SMRT, in both CI and normal-hearing (NH) subjects. For CI users, we also evaluated the correlation of speech recognition in noise with the spectrotemporal modulation thresholds while controlling for either temporal or spectral modulation sensitivity.

Materials and Method

Subjects and Hardware

A total of 22 postlingually deafened and implanted ears were tested in the study. Four of them were implanted with Advanced Bionics (AB) devices (Advanced Bionics, Valencia, CA) and the rest of the group was implanted with Cochlear Nucleus® devices (Cochlear Corporation, Englewood, CO). Five subjects were bilaterally implanted and both ears were tested. In addition, 27 young NH subjects participated in the study. They were screened in both ears for NH (<20 dB HL) at octave frequencies between 250 and 8000 Hz.

All psychophysical tests were acoustic tests performed in the sound field in a double-walled sound-treated booth. The acoustic stimuli were played from a loudspeaker placed 1 m away from the head of the subject at 0 azimuth, at a presentation level of 65 dB (A). For CI subjects, the acoustic tests were conducted using the subjects' own speech processor set at the daily-use program. None of the CI subjects had residual hearing in the implanted ear and those who had residual hearing in the contralateral ear were plugged for all psychophysical testing. The use of human subjects was approved by the institutional review board at East Carolina University. Demographic information of the CI subjects is shown in Table 1.

Spectrotemporal Modulation Discrimination Test

The SMRT (Aronoff & Landsberger, 2013) was used to assess subjects' spectrotemporal modulation sensitivity. The stimuli consisted of 202 sine waves spaced every 0.03 octaves within the frequency range of 100 to 6400 Hz. The amplitude of the pure tones was modulated by a half-wave rectified sinewave with a starting phase that systematically changes across the carriers. The shift in the starting phase of the modulators created a drift in the spectral ripples over time. The temporal repetition rate of the ripples was 5 Hz the ripple density increased in the stimulus, the depth of the temporal modulation became shallower (Lawler et al., 2017). Thus as the ripples increased, both the spectral and temporal envelopes of the stimulus became flatter. The task was to discriminate the target stimulus, which adapted in number of RPO, from the reference stimulus containing 20 RPO. A three-alternative forced-choice (3AFC) paradigm was used, where the target stimulus appeared in one of the three intervals chosen at random. The duration of the stimuli was 500 ms, including 100-ms onset and offset linear ramps. The interstimulus interval was 1 s. The starting RPO in the target stimulus was 0.5 and it increased in a step size of 0.2 RPO based on the subject's response using a 2-up 1-down rule. Although no formal training (feedback) was provided, the first repetition of the test was used to familiarize the subjects with the test procedure and stimuli, and the threshold for that repetition was not used.

Static Spectral Ripple Discrimination Test

The static spectral ripple discrimination test used techniques reported in Won et al. (2007). The stimuli were 500-ms long complex signals (150-ms onset and offset ramps) consisting of 2,000 sine waves logarithmically spaced in the 100 Hz to 5000 Hz range. The spectrum of the stimulus was modulated by a full-wave rectified sinusoidal envelope on a logarithmic scale. The modulation depth was kept at 30 dB. Ripple density started at 0.125 RPO and adapted in a ratio of 1.414 using a 2-up 1-down rule. In the inverted version of the stimulus, the starting

Table 1. Subject Demographics.

Subject	Ear	Gender	Age	CI use (years)	Duration of deafness (years)	Implant type	Processor type	Speech processing strategy
S1	L	M	80.24	17.2	0.6	CI24R (CS)	CP1000	ACE
S1	R	M	80.24	11.2	6.0	CI24RE (CA)	CP1000	ACE
S3	L	F	68.88	12.9	3.4	CI24RE (CA)	CP920	ACE
S3	R	F	68.88	14.4	1.8	CI24RE (CA)	CP920	ACE
S4	L	F	56.22	7.7	4.6	CI24RE (CA)	CP810	ACE
S7	R	F	73.58	8.4	27.8	CI24RE (CA)	CP1000	ACE
S10	L	F	69.03	18.1	0.8	CI24R (CS)	CP1000	ACE
S10	R	F	69.03	6.4	12.4	CI24RE (CA)	CP1000	ACE
S18	L	F	67.32	4.6	3.6	CI422	CP910	ACE
S19	L	F	72.53	11.9	4.3	CI24RE (CA)	CP1000	ACE
S22	R	F	74.37	6.8	0.4	CI24RE(CA)	CP920	ACE
S25	L	F	62.08	11.6	0.7	CI24RE (CA)	CP900	ACE
S25	R	F	62.08	10.8	1.4	CI24RE (CA)	CP900	ACE
S27	R	M	59.66	13.2	0.0	CI24RE	CP920	ACE
S28	R	F	76.53	12.5	0.8	HiFocus 1J	Naida CI Q70	HiRes Optima-S
S31	L	M	69.60	3.7	1.5	CI422	Kanso	ACE
S32	L	M	68.77	2.4	6.4	HiFocus ms	Naida CI Q90	HiRes Optima-S
S33	R	F	68.82	2.0	30.2	HiFocus 1J	Naida CI Q90	HiRes optima-P
S34	L	F	72.58	1.5	3.7	HiFocus ms	Naida CI Q90	HiRes Optima-S
S36	L	M	79.61	0.7	21.9	CI522	CP1000	ACE
S37	L	M	74.21	4.8	34.4	CI422	CP910	ACE
S37	R	M	74.21	16.2	0.4	CI24R (CS)	CP910	ACE

Note. CI = cochlear implant; M = male; F = female; L = left; R = right; ACE = Advanced Combination Encoder; CA = Contour Advanced; CS = Contour.

phase of the rectified sinusoidal spectral envelope was set to $\pi/2$ instead of 0 radians. The subject's task was to discriminate the standard stimulus from its inverted version in a 3AFC paradigm. The interstimulus interval was 1 s. A training trial was conducted before formal testing, and feedback was given. The threshold was the average of the RPOs at the last 6 reversal points out of a total of 12 reversals. Note that CI processing may introduce nonlinear spectral distortions in the ripple stimuli.

Temporal Modulation Detection Thresholds

Amplitude modulation detection thresholds (MDTs) were measured using amplitude modulated broad-band noises. The AM was fixed at 4 Hz and modulation depth of the AM was adapted to measure the minimum depth required to detect the AM (dB with regard to 100% modulation depth). The stimuli were 500 ms long including 100-ms onset and offset ramps. A 2AFC paradigm was used, where the subject was instructed to choose the one interval that had the *warble* sound. The stimulus–stimulus interval was 1 s. Modulation depth started at 50% and adapted following subjects' responses. Step size was 5 dB for the first reversal, 2 dB for the next 3 reversals, and 1 dB for the rest of the 12 reversals. Threshold was taken as the average modulation depth at the last 6 reversal points of a total of 12 (dB with regard to 100% modulation depth). A training trial was conducted before formal testing, and feedback was given. It should be noted that the broadband AM signal is more likely to trigger automatic gain control in CI processing than within-channel AM. Thus, the MDTs measured here may underestimate the possible contribution of within-channel AM to spectrotemporal modulation discrimination.

Speech Recognition

CI subjects were measured for the 50%-correct speech reception thresholds (SRTs). Thresholds were measured using City University of New York (CUNY) sentences (Boothroyd et al., 1985) in a speech-shaped noise amplitude modulated at 4 Hz. The level of the sentences was fixed at 65 dB (A) and the noise level was adapted. Signal to noise ratio (SNR) started at 20 dB and adapted in a step size of 2 dB following a 1-down 1-up rule tracking 50% correct of the psychometric function. A response was considered correct if all key words were correctly recalled. SRT was measured twice and the thresholds were averaged.

Results

CI Subjects

Figure 1 shows that there was a highly significant correlation between the spectrotemporal modulation

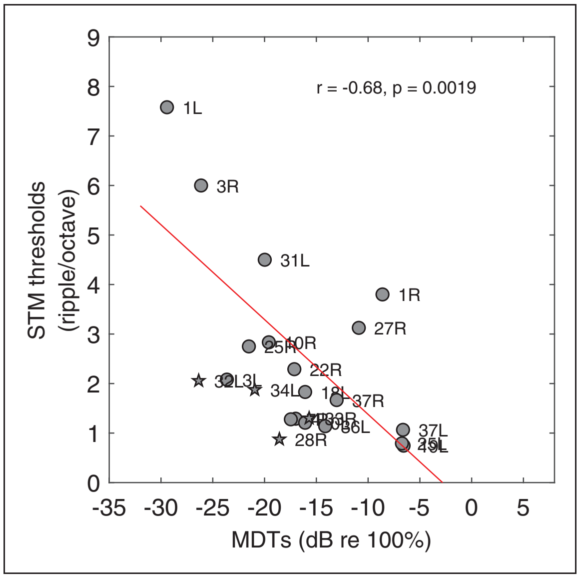

Figure 1. Correlations Between STM Thresholds and MDTs in CI Subjects. The line indicates linear fit to the data. Data from AB subjects shown in pentagrams were not included in data analysis. STM = spectrotemporal modulation; MDTs = modulation detection thresholds.

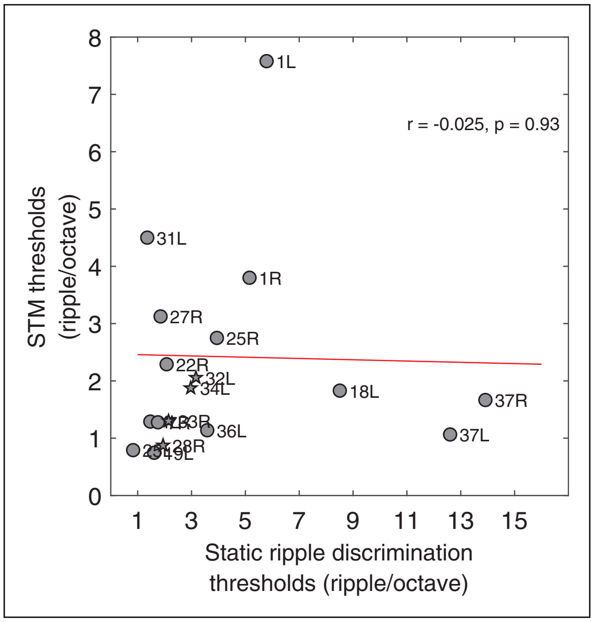

Figure 2. Correlation Between STM Thresholds and the Static Ripple Discrimination Thresholds in CI Subjects. The line indicates linear fit to all data. Data from AB subjects shown in pentagrams were not included in data analysis. STM = spectrotemporal modulation.

thresholds and MDTs ($r = -.68$, $p = .002$) in CI listeners, but there was no relationship between the spectrotemporal modulation thresholds and static spectral ripple thresholds ($r = .03$, $p = .93$; Figure 2). Note that in order to maintain the homogeneity of the group and remove confounding factors related to differences in CI processing, the four AB subjects were not included in the data analyses, but their data are shown in the figures (pentagrams in Figures 1–3). As can be seen from the figures, results from the AB subjects followed the trend of the data from the Nucleus device users despite the difference in CI processing. The correlation coefficients were standardized for comparing the strength of relationship with the spectrotemporal modulation thresholds. Standardized coefficient was statistically greater for the correlation with the MDTs, than with the static ripple thresholds (−0.68 vs. 0.009), as revealed by a *Z* test ($z = 2.11$, $p = .01$). The relative importance of temporal and spectral modulation sensitivity for spectrotemporal modulation thresholds was evaluated via linear regression, where the modulation threshold was the dependent variable, and MDTs and static ripple thresholds were entered stepwise to the model. MDTs explained a significant proportion of the variance in spectrotemporal modulation thresholds, $F(1,12) = 10.91$, $p = .006$, $R^2 = 0.43$, but adding the static ripple thresholds did not explain further unique variance ($t = 0.34$, $p = .74$) and thus did not improve the model, $F(2,11) = 5.11$, $p = .02$, $R^2 = 0.38$. The regression analysis was equivalent to conducting partial correlations with the spectrotemporal modulation thresholds controlling

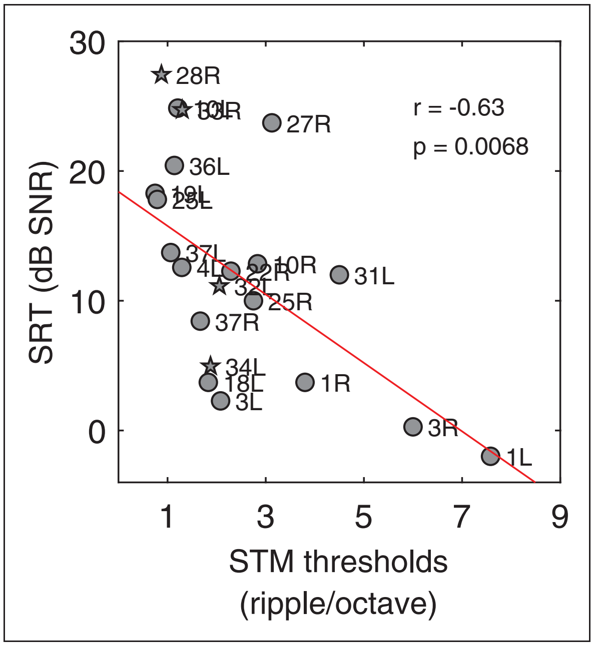

Figure 3. Correlations Between STM Thresholds and Speech Reception Thresholds (SRTs) in CI Subjects. The line indicates linear fit to the data. Data from AB subjects shown in pentagrams were not included in data analysis. STM = spectrotemporal modulation.

Table 2. Correlations Between Variables.

		MDTs		Static ripple		SRT (dB SNR)	
		R/*ρ*	*p*	*R*/*ρ*	*p*	*R*	*p*
All subjects (Pearson)	STM thresholds	−0.68	*.002*[a]	−0.03	.93	−0.63	*.007*[a]
	STM thresholds controlling for MDTs			0.10	.74	−0.38	.15
	STM thresholds controlling for static ripple discrimination	−0.69	*.009*[a]			−0.69	*.013*[a]
All subjects (Spearman rank order)	STM thresholds	−0.67	*.003*[a]	0.21	.46		
Data winsorizing (Pearson)	STM thresholds	−0.59	*0.009*[a]	0.28	.34		

Note. MDT = modulation detection threshold; STM = spectrotemporally modulated; SRT = speech reception threshold; SNR = signal to noise ratio.
[a]Significant after Holm–Bonferroni correction.

for the other independent variable. Statistics of the correlations can be found in Table 2. Note that multiple comparisons were Holm–Bonferroni corrected.

To faithfully represent the ripples in the spectrum, a CI processor must provide sufficient spectral sampling of the stimulus which would require that the number of channels per octave equal to at least twice the RPO in the stimulus. This corresponds roughly to a theoretical limit of 2.5 RPO for the Cochlear devices that use 22 channels. The ripples should no longer be discriminable if the ripple period is less than the bandwidth of the filters, which corresponds to a theoretical limit of 5 RPO (O'Neill et al., 2019). The limit may be lower for other devices that use fewer channels and different bandwidth allocations. Some subjects showed static ripple thresholds or spectrotemporal modulation thresholds better than what would be expected from these limits and some even performed at NH-like levels (S1, S18, and S37). Thresholds from these subjects were confirmed by additional repeats and they showed high trial-to-trial consistency. It is unclear how these subjects performed at such high levels. There might have been loudness differences in the target and reference, as their levels were not roved. Two approaches were taken to address this issue. Assuming that the thresholds better than the theoretical limits represent better performance and can be rank ordered, Spearman's correlations were used to reexamine the relationship with the spectrotemporal modulation thresholds. The conclusions were the same as those drawn from the Pearson's correlations: The spectrotemporal modulation thresholds were significantly correlated with MDTs ($\rho = -0.67$, $p = .003$) but not with static ripple thresholds ($\rho = 0.21$, $p = .46$). Assuming that the thresholds above the limit do represent better performance but the magnitude cannot be ordered (8 RPO not necessarily better than 7 RPO), the spectrotemporal modulation and static ripple thresholds that were greater than 2.5 RPO were replaced with the value of 2.5 and the analysis was repeated (Winn & O'Brien, 2019). Results remained the same: spectrotemporal modulation thresholds were correlated with MDTs ($r = -.59$, $p = .009$) but not with the static ripple thresholds ($r = .28$, $p = .34$). These correlations are also summarized in Table 2.

Figure 3 shows the relationship between speech recognition in noise measured in the CI subjects, and the spectrotemporal modulation thresholds ($r = -.63$, $p = .007$). Table 2 shows the same relationship while partialing out either temporal or spectral modulation sensitivity. SRTs were correlated with the spectrotemporal modulation thresholds while controlling for the static ripple thresholds ($r = -.69$, $p = .013$) but were not correlated with the spectrotemporal modulation thresholds while controlling for MDTs ($r = -.38$, $p = .15$).

NH Subjects

Results from the NH subjects showed patterns comparable to those of the CI subjects. Spectrotemporal modulation thresholds were significantly correlated with MDTs ($r = -.48$, $p = .012$) but were not correlated with the static ripple thresholds ($r = .37$, $p = .059$). One subject had an unusually poor static ripple threshold (highlighted with square) inconsistent with the pure tone thresholds (Figure 4). Removing the subject resulted in worse correlation between spectrotemporal modulation and the static ripple thresholds ($r = .14$, $p = .50$). Comparing the strength of the correlations, the standardized coefficients were larger for MDTs than for static ripple thresholds (−0.475 vs. 0.139), but the difference was not statistically significant as revealed by the Z test ($z = 0.29$, $p = .09$).

Discussion

In this study, CI subjects' sensitivity to modulation imposed in both the time and frequency domains was examined using a spectrotemporal modulation discrimination test, i.e., SMRT (Aronoff & Landsberger, 2013). The test was originally designed for removing possible artifacts in the static ripple tests as measures of spectral resolution, by systematically varying the spectral modulation over time. In doing so, AM is created within

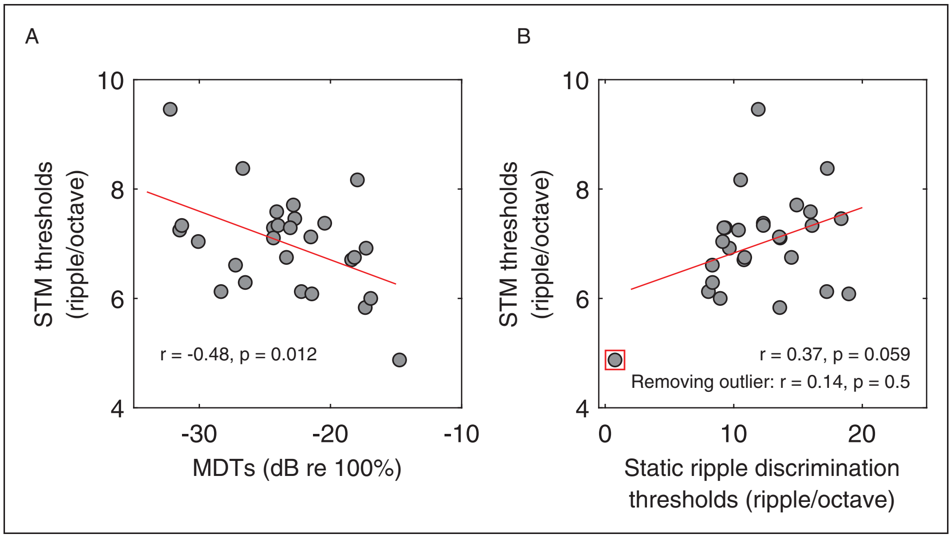

Figure 4. Correlations Between STM Thresholds and MDTs (Panel A) and Static Ripple Discrimination Thresholds (Panel B) in NH Subjects. Lines indicate linear fit to the data. STM = spectrotemporal modulation; MDTs = modulation detection thresholds.

channels. The modulation depth of the AM reduces as the ripple density increases in the SMRT stimulus. Narne et al. (2016) showed that NH subjects performed better in discriminating the gliding ripples than static ripples from a 20 RPO reference (essentially unmodulated). In the gliding ripples, they identified AM fluctuations at the high-frequency output of the simulated gammatone auditory filters. When the AM was masked by a notched noise, the performance between the gliding and static ripple discrimination then became equivalent. Using electrodograms, Archer-Boyd et al. (2018) demonstrated that there was salient within-channel AM in the SMRT stimuli in the output from a CI electrode. These results suggest that subjects can use either the difference in the AM depth or ripple density between the target and reference to perform the task. The question addressed in this study was: If modulations are present in both the time and frequency domains, is it the subjects' temporal modulation sensitivity or spectral modulation sensitivity that drives the discrimination performance? The second question was: If one sensitivity dominates the performance, is it specific to the CI population?

In examining the relationship between the spectrotemporal modulation thresholds, MDTs, and the static ripple thresholds in the CI subjects, the general pattern of the data showed that there was a stronger correlation between spectrotemporal modulation thresholds and MDTs, compared with the relationship between spectrotemporal modulation and the static ripple thresholds. The static ripple thresholds failed to account for the variance in spectrotemporal modulation thresholds that could not be explained by AM detection. These results suggest that the CI listeners weighed the temporal cues more heavily than the spectral cues to perform the task. It was interesting to observe that the subjects with exceptionally good static ripple discrimination thresholds, those in the NH range, such as S18L, S37L, and S37R, had spectrotemporal modulation thresholds below 2 RPO. It appeared that these subjects relied on the AM cues and thus were limited by their AM detection sensitivity in discriminating the spectrotemporal modulation. It was interesting that whatever information that these subjects used to perform at an NH-comparable level in the static ripple test, were not used to discriminate the moving ripples in the spectrotemporally modulated stimuli. Furthermore, subjects who had good AM sensitivity also seemed to have weighed more heavily the AM depth differences in the spectrotemporally modulated stimuli as a cue and produced thresholds consistent with their MDTs, rather than their static ripple thresholds (e.g., S31L). The correlation between SMRT and AM detection has been reported before (De Jong et al., 2018). The result could be interpreted as a correlation between spectral and temporal resolution, since SMRT was developed as a modified spectral resolution measure. Based on the results of this study, particularly the lack of correlation between SMRT and the static ripple thresholds, the correlation between SMRT and AM detection should perhaps be interpreted as AM sensitivity dominating the performance of SMRT. This pattern was largely the same when including all subjects or winsorizing the outliers, suggesting that the correlations or lack thereof when including the outliers were robust to spectral aliasing.

There were few CI studies in the past that have considered the relative contribution of temporal and spectral modulation sensitivity to detection or discrimination of modulation in both domains. Won et al. (2015) did not find any strong evidence that spectrotemporal modulation detection, that is, the minimum modulation depth required to detect the STM, was correlated with either the subjects' ability to detect AM or static spectral ripples. Zheng et al. (2017) also reported that the spectrotemporal modulation transfer functions cannot be predicted by the temporal, or spectral modulation transfer functions. In fact, the spectral transfer function, decomposed from a spectrotemporal modulation transfer function, was better than the measured values, especially for stimuli with higher ripple density. They speculated that this could reflect a compensatory mechanism of the CI-acclimated brain, in which the subjects rely heavily on the relatively intact AM cues to compensate for the relatively poor sensitivity for dense ripples in the spectrum, thus the spectrotemporal modulation functions were heightened compared with what would be expected from the measured sensitivity in either domain. There is also evidence from the hearing-impaired population that indicates that listeners who relied to a greater extent on temporal cues for speech recognition were those lacking the ability to discriminate the fine-grained spectral information (Souza et al., 2020). The present results from the CI users could also be consistent with such a cue availability theory. However, some subjects in this study performed exceptionally well on the static ripple test but their gliding ripple performance was poorer and limited by their AM sensitivity. These results suggest that the CI subjects' reliance on AM cues cannot simply be attributed to the spectral cues being less available. One could argue that the spectral cues, although sufficient for discriminating the static ripples, may be more distorted than the AM cue in the spectrotemporally modulated stimulus, therefore, as long as AM is present, CI listeners would prefer to use the less degraded AM information. These hypotheses were then evaluated by examining the same relationship between spectrotemporal modulation discrimination, AM detection, and static ripple discrimination in NH subjects. Results from the NH subjects showed similar patterns as those measured in CI subjects, although the pattern was somewhat weaker. Specifically, the spectrotemporal modulation thresholds were correlated with MDTs, but not with the static ripple thresholds, and static ripple thresholds did not account for unexplained variance in the spectrotemporal modulation thresholds when controlling for AM detection, consistent with the CI data. However, when considering the strength of the correlations of either variable with the spectrotemporal modulation thresholds (one was statistically significant and the other was not), the correlation coefficients when transformed to Z scores were not statistically different. Therefore, there is some weak evidence to indicate that this preference for using the timing cue over spectral cue, perhaps unconscious, is not specific to the CI population and that there is a central decision mechanism common to both CI and NH listeners.

It remains to be tested whether the contribution of AM detection seen in this study, would extend to the other types of spectrotemporal tests. The recently developed STRIPES test requires subjects to discriminate upward and downward gliding ripples (Archer-Boyd et al., 2018). Unlike SMRT, where the AM information in a given channel is different between the target and the reference, the AM at a specific frequency region, specifically the AM depth and rate between the upward and downward gliding ripples of the STRIPES are the same. Archer-Boyd et al. (2018) showed with vocoded STRIPES stimuli that reducing the cut-off frequency of the low-pass envelope filter from 300 Hz to 3 Hz was equivalent in terms of discrimination performance to lowering the number of channels from 16 to 12. These data suggest that although the AM within a specific channel cannot be used for the task, as it is the same between the target and the reference, the subjects must still detect the within-channel AM and compare them across channels to perceive the difference in the gliding direction. Archer-Boyd et al. showed that if the modulation depth of the AM was considerably reduced, as in the 3-Hz envelope, the AM patterns across channels would be more difficult to perceive. Thus, it would be interesting to examine if and the extent to which AM detection also plays a role in the STRIPES performance.

Previously, performance in the SMRT has been associated with CI users' speech recognition (Holden et al., 2016; Lawler et al., 2017). Relationship between performance in the SMRT and speech recognition in noise was replicated in this study and results confirmed that better SMRT thresholds were associated with better speech recognition performance in noise. The relative contribution of temporal and spectral modulation sensitivity to this relationship was examined. Results showed that when partialing out the effect of AM detection, SMRT performance no longer correlated with SRTs. However, when partialing out the effect of static ripple discrimination, the correlation between SMRT performance and speech recognition was largely unaffected compared with not controlling for static ripple discrimination. These results suggest that the correlation between SMRT thresholds and speech recognition mainly reflected the relationship between the subjects' AM sensitivity and speech recognition. These findings were consistent with the results reported above that SMRT thresholds were primarily accounted for by subjects' AM detection. The contribution of AM sensitivity to speech recognition has been extensively studied in the past. MDTs measured at a middle electrode or averaged across the entire array have been shown to predict speech recognition in noise and in quiet (Fu, 2002;

Garadat et al., 2012, 2013; Luo et al., 2008). Speech processing maps that use stimulation sites with the best MDTs produced significantly better speech recognition than those using sites with the poorest MDTs, and those using all functioning stimulation sites without considering their AM acuity (Garadat et al., 2013; Zhou & Pfingst, 2012). Furthermore, instead of removing sites with poorer AM sensitivity, Zhou and Pfingst (2014) also showed that MDTs of the poorer sites can be improved by artificially setting threshold levels 5% higher than the true thresholds and the manipulation in turn resulted in improved speech recognition. The present results added further evidence that AM sensitivity is important for speech recognition with a CI. The present results also supported the design of speech processing strategies that emphasize the salience of such temporal cues.

Previous work in our laboratory showed that MDTs (phase duration modulation) measured via direct stimulation depended on the stimulation site's spatial tuning curves (Zhou et al., 2018). Stimulation sites with sharp spatial tuning curves were associated with better MDTs. These results suggest a link between temporal and spectral resolution, at least for local frequency regions. If this relationship extended to the whole electrode array, one would expect that the broad-band acoustic AM thresholds would be correlated with broad-band static ripple thresholds. However, the current results did not support such relationship. Previous research also showed that when stimulation sites estimated to produce broad stimulation (based on low-rate thresholds) were removed, SMRT thresholds and speech recognition both improved compared with using all stimulation sites (Zhou, 2017). There is evidence to show that broad-band spectrotemporal modulation thresholds can be best predicted by a local excitation-pattern model, that is, best local threshold, rather than the average frequency selectivity across the global excitation pattern (Narne et al., 2018). If an otherwise sharply tuned site was interacting with the neighboring sites that produced broad excitation, then deactivating the neighboring sites may result in new best local regions contributing to the broad-band threshold. Based on the current results, it is however more likely that the improvement in SMRT threshold with deactivation was a result of increased salience of AM in the stimulus. Lawler et al. (2017) showed that AM depth in the SMRT stimuli was more salient without spectral smearing than with spectral smearing. This interpretation was consistent with the NH results from Narne et al. (2016), which showed that when AM in the gliding ripples was weakened by spectral smearing, performance in discriminating the gliding ripples worsened and became comparable to that of discriminating the static ripples.

Conclusion

In conclusion, results of this study showed that CI listeners tended to use AM information to discriminate spectrotemporally modulated stimuli used in SMRT. Whether similar reliance on AM cues exists for other spectrotemporally modulated stimuli and tests warrants future research. The underlying mechanism for this preference could be that spectral information is less accessible with a CI. However, similar patterns from the NH listeners, albeit weaker compared with those from CI subjects, suggest that the mechanism of favoring temporal information may be common to both the CI and NH populations.

Acknowledgments

The authors would like to thank our cochlear implant subjects for their dedicated participation and David Landsberger for his input for the interpretation of the data.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by National Institute on Deafness and Other Communication Disorders, part of the National Institutes of Health (Grant R03 DC014771-01A1 and Grant R01 DC017702-01).

ORCID iD

Ning Zhou https://orcid.org/0000-0002-9521-679X

References

Anderson, E. S., Oxenham, A. J., Kreft, H., Nelson, P. B., & Nelson, D. A. (2011). Comparing spectral tuning curves, spectral ripple resolution, and speech perception in cochlear implant users. *The Journal of the Acoustical Society of America*, *130*(1), 364–375. https://doi.org/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. *The Journal of the Acoustical Society of America*, *132*(6), 3925–3934. https://doi.org/10.1121/1.4763999

Archer-Boyd, A.W, Southwell, R.V., Deeks, J.M., Turner, R. E., & Carlyon, R.P. (2018). Development and validation of a spectro-temporal processing test for cochlear implant listeners. *The Journal of the Acoustical Society of America*, *144*(5), 2983–2997. https://doi.org/10.1121/1.5079636

Aronoff, J. M., & Landsberger, D. M. (2013). The development of a modified spectral ripple test. *The Journal of the Acoustical Society of America*, *134*(2), EL217–EL222. https://doi.org/10.1121/1.4813802

Berenstein, C. K., Mens, L. H., 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 and Hearing*, *29*(2), 250–260. https://doi.org/10.1097/aud.0b013e3181645336

Boothroyd, A., Hanin, L., & Hnath, T. (1985). A sentence test of speech perception: Reliability, set equivalence, and short term learning. *CUNY Academic Works*, 1–17.

Busby, P. A., & Clark, G. M. (2000). Electrode discrimination by early-deafened subjects using the cochlear limited multiple-electrode cochlear implant. *Ear and Hearing*, *21*(4), 291–304. https://doi.org/10.1097/00003446-200008000-00004

De Jong, M. A. M., Briaire, J. J., & Frijns, J. H. M. (2018). Learning effects in psychophysical tests of spectral and temporal resolution. *Ear and Hearing*, *39*(3), 475–481. https://doi.org/10.1097/AUD.0000000000000499

Dorman, M. F., Spahr, A., Gifford, R. H., Cook, S., Zhang, T., Loiselle, L., Yost, W., Cardy, L., Whittingham, J., & Schramm, D. (2012). Current research with cochlear implants at Arizona State University. *Journal of the American Academy of Audiology*, *23*(6), 385–395. https://doi.org/10.3766/jaaa.23.6.2

Drennan, W. R., Anderson, E. S., Won, J. H., & Rubinstein, J. T. (2014). Validation of a clinical assessment of spectral ripple resolution for cochlear implant users. *Ear and Hearing*, *35*(3), e92–e98. https://doi.org/10.1097/AUD.0000000000000009

Drennan, W.R., Jongnion, J.K., Ruffin, C., & Rubinstein, J.T. (2008). Discrimination of Schroeder-phase harmonics complexes by normal-hearing and cochlear-implant listeners. *Journal of the Association for Research in Otolaryngology*, *9*(1), 18–149. https://doi.org/10.1007/s10162-007-0107-6

Drennan, W. R., Won, J. H., Nie, K., Jameyson, E., & Rubinstein, J. T. (2010). Sensitivity of psychophysical measures to signal processor modifications in cochlear implant users. *Hearing Research*, *262*(1–2), 1–8. https://doi.org/10.1016/j.heares.2010.02.003

Finley, C. C., Holden, T. A., Holden, L. K., Whiting, B. R., Chole, R. A., Neely, G. J., Hullar, T. E., & Skinner, M. W. (2008). Role of electrode placement as a contributor to variability in cochlear implant outcomes. *Otology & Neurotology*, *29*(7), 920–928. https://doi.org/10.1097/MAO.0b013e318184f492

Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America*, *110*(2), 1150–1163. https://doi.org/10.1121/1.1381538

Fu, Q. J. (2002). Temporal processing and speech recognition in cochlear implant users. *Neuroreport*, *13*(13), 1635–1639. https://doi.org/10.1097/00001756-200209160-00013

Garadat, S. N., Zwolan, T. A., & Pfingst, B. E. (2012). Across-site patterns of modulation detection: relation to speech recognition. *The Journal of Acoustical Society of America*, *131*(5), 4030–4041. https://doi.org/10.1121/1.3701879

Garadat, S. N., Zwolan, T. A., & Pfingst, B. E. (2013). Using temporal modulation sensitivity to select stimulation sites for processor MAPs in cochlear implant listeners. *Audiology & Neuro-otology*, *18*(4), 247–260. https://doi.org/10.1159/000351302

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. https://doi.org/10.3109/14992027.2013.851800

Gifford, R. H., Noble, J. H., Camarata, S. M., Sunderhaus, L. W., Dwyer, R. T., Dawant, B. M., Dietrich, M. S., Labadie, R. F. (2018). The relationship between spectral modulation detection and speech recognition: Adult versus pediatric cochlear implant recipients. *Trends in Hearing*, *22*, 1–14. https://doi.org/10.1177/2331216518771176

Henry, B. A., & Turner, C. W. (2003). The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *The Journal of Acoustical Society of America*, *113*(5), 2861–2873. https://doi.org/10.1121/1.1561900

Holden, L. K., Firszt, J. B., Reeder, R. M., Uchanski, R. M., Dwyer, N. Y., & Holden, T. A. (2016). Factors affecting outcomes in cochlear implant recipients implanted with a perimodiolar electrode array located in Scala Tympani. *Otology & Neurotology*, *37*(10), 1662–1668. https://doi.org/10.1097/MAO.0000000000001241

Kawano, A., Seldon, H. L., Clark, G. M., Ramsden, R. T., & Raine, C. H. (1998). Intracochlear factors contributing to psychophysical percepts following cochlear implantation. *Acta Otolaryngology*, *118*(3), 313–326. https://doi.org/10.1080/00016489850183386

Landsberger, D.M., Dwyer, R.T., Stupak, N., Gifford, RH. (2019). Validating a quick spectral modulation detection task. *Ear and Hearing*, *40*(6), 1478–1480. https://doi.org/10.1097/AUD.0000000000000713

Lawler, M., Yu, J., Aronoff, J. M. (2017). Comparison of the spectral-temporally modulated ripple test with the Arizona Biomedical Institute Sentence Test in cochlear implant users. *Ear and Hearing*, *38*(6), 760–766. https://doi.org/10.1097/AUD.0000000000000496

Luo, X., Fu, Q. J., Wei, C. G., Cao, K. L. (2008). Speech recognition and temporal amplitude modulation processing by Mandarin-speaking cochlear implant users. *Ear and Hearing*, *29*(6), 957–970. https://doi.org/10.1097/AUD.0b013e3181888f61

McKay, C. M., Azadpour, M., & Akhoun, I. (2009). *In search of frequency resolution* [Conference session]. Conference on Implantable Auditory Protheses, Tahoe City, CA, United States. Abstract book page 54.

McKay, C. M., Remine, M. D., & McDermott, H. J. (2001). Loudness summation for pulsatile electrical stimulation of the cochlea: Effects of rate, electrode separation, level, and mode of stimulation. *The Journal of the Acoustical Society of America*, *110*(3 Pt 1), 1514–1524. https://doi.org/10.1121/1.1394222

Mehraei, G., Gallun, F. J., Leek, M. R., & Bernstein, J. G. W. (2014). Spectrotemporal modulation sensitivity for hearing-impaired listeners: Dependence on carrier center frequency and the relationship to speech intelligibility. *The Journal of the Acoustical Society of America*, *136*(301), 301–316. https://doi.org/10.1121/1.4881918

Nadol, J. B., Jr., Shiao, J. Y., Burgess, B. J., Ketten, D. R., Eddington, D. K., Gantz, B.J., Kos, I., Montandon, P., Coker, N. J., Roland, J. T., Jr., Shallop, J. K. (2011). Histopathy of cochlear implants in humans. *Annals of Otology, Rhinology & Laryngology*, *110*(9), 883–891. https://doi.org/10.1177/000348940111000914

Narne, V. K., Prabhu, P. P., Van Dun, B., & Moore, B. C. J. (2018). Ripple glide direction discrimination and its relationship to frequency selectivity estimated using notched noise. *Acta Acustica united with Acustica*, *104*(6),1063–1074. https://doi.org/10.3813/AAA.919272

Narne, V. K., Sharma, M., Van Dun, B., Bansal, S., Prabhu, L., & Moore, B. C. (2016). Effects of spectral smearing on performance of the spectral ripple and spectro-temporal ripple tests. *Journal of Acoustical Society of America*, *140*(6), 4298–4306. https://doi.org/10.1121/1.4971419

Nelson, D. A., Kreft, H. A., Anderson, E. S., & Donaldson, G. S. (2011). Spatial tuning curves from apical, middle, and basal electrodes in cochlear implant users. *The Journal of Acoustical Society of America*, *129*(6), 3916–3933. https://doi.org/10.1121/1.3583503

O'Connell, B. P., Hunter, J. B., & Wanna, G. B. (2016). The importance of electrode location in cochlear implantation. *Laryngoscope Investigative Otolaryngology*, *1*(6), 169–174. https://doi.org/10.1002/lio2.42

O'Neill, E.R., Kreft, H.A., & Oxenham, A. J. (2019). Speech perception with spectrally non-overlapping maskers as measure of spectral resolution in cochlear implant users. *Journal of the Association for Research in Otolaryngology*, *20*(2), 151–167. https://doi.org/10.1007/s10162-018-00702-2

Saoji, A. A., Litvak, L., Spahr, A. J., & Eddins, D. A. (2009). Spectral modulation detection and vowel and consonant identifications in cochlear implant listeners. *The Journal of the Acoustical Society of America*, *126*(3), 955–958. https://doi.org/10.1121/1.3179670

Souza, P., Gallun, F., & Wright, R. (2020). Contributions to speech-cue weighting in older adults with impaired hearing. *Journal of Speech, Language, and Hearing Research*, *63*(1), 334–344. https://doi.org/10.1044/2019_JSLHR-19-00176

Supin, A. Y., Popov, V. V., Milekhina, O.N., & Tarakanov, M. B. (1998). Ripple density resolution for various rippled-noise patterns. *The Journal of the Acoustical Society of America*, *103*(4), 2042–2050. https://doi.org/10.1121/1.421351

Winn, M.B., & O'Brien, G. (2019, May 31). Flaws in spectral ripple stimuli for listeners with cochlear implants. *PsyArXiv* Preprint. https://psyarxiv.com/cbwgh/

Won, J. H., Drennan, W. R., Kang, R. S., & Rubinstein, J. T. (2010). Psychoacoustic abilities associated with music perception to cochlear implant users. *Ear and Hearing*, *31*(6), 796–805. https://doi.org/10.1097/AUD.0b013e3181e8b7bd

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. https://doi.org/10.1007/s10162-007-0085-8

Won, J. H., Jones, G. L., Drennan, W. R., Jameyson, E. M., & Rubinstein, J. T. (2011). Evidence of across-channel processing for spectral-ripple discrimination in cochlear implant listeners. *The Journal of the Acoustical Society of America*, *130*(4), 2088–2097. https://doi.org/10.1121/1.3624820

Won J. H., Moon, I. J., Jin, S., Park, H., Woo, J., Cho, Y. S., Chung, W. H., & Hong, S. H. (2015). Spectrotemporal modulation detection and speech perception by cochlear implant users. *PLoS One*, *10*(10), 1–24. https://doi.org/10.1371/journal.pone.0140920

Zheng, Y., Escabi, M., & Litovsky, R. Y. (2017). Spectro-temporal cues enhance modulation sensitivity in Cochlear Implant users. *Hearing Research*, *351*, 45–54. https://doi.org/10.1016/j.heares.2017.05.009

Zhou, N. (2017). Deactivating stimulation sites based on low-rate thresholds improves spectral ripple and speech reception thresholds in cochlear implant users. *The Journal of the Acoustical Society of America*, *141*(3), EL243–EL248. https://doi.org/10.1121/1.4977235

Zhou, N., Cadmus, M., Dong, L., & Matthews, J. (2018). Temporal modulation detection depends on sharpness of spatial tuning. *Journal of the Association for Research in Otolaryngology*, *19*(3), 317–330. https://doi.org/10.1007/s10162-018-0663-y

Zhou, N., & Pfingst, B. E. (2012). Psychophysically based site selection coupled with dichotic stimulation improves speech recognition in noise with bilateral cochlear implants. *The Journal of Acoustical Society of America*, *132*(2), 994–1008. https://doi.org/10.1121/1.4730907

Zhou, N., & Pfingst, B. E. (2014). Effects of site-specific level adjustment on speech recognition with cochlear implants. *Ear and Hearing*, *35*(1), 30–40. https://doi.org/10.1097/AUD.0b013e31829d15cc

Zwolan, T. A., Collins, L. M., & Wakefield, G. H. (1997). Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects. *The Journal of the Acoustical Society of America*, *102*(5), 3673–3685. https://doi.org/10.1121/1.420401