

Assessing the role of spectral and intensity cues in spectral ripple detection and discrimination in cochlear-implant users

Elizabeth S. Anderson^{a)}

VM Bloedel Hearing Research Center, CHDD Building, University of Washington, Box 357923, Seattle, Washington 98195

Andrew J. Oxenham

Department of Psychology, University of Minnesota, Elliott Hall, 75 East River Parkway, Minneapolis, Minnesota 55455

Peggy B. Nelson

Department of Speech-Language-Hearing Sciences, University of Minnesota, Shevlin Hall, 164 Pillsbury Drive Southeast, Minneapolis, Minnesota 55455

David A. Nelson

Clinical Psychoacoustics Laboratory, Department of Otolaryngology, University of Minnesota, MMC396, 420 Delaware Street Southeast, Minneapolis, Minnesota 55455

(Received 6 April 2012; revised 5 October 2012; accepted 9 October 2012)

Measures of spectral ripple resolution have become widely used psychophysical tools for assessing spectral resolution in cochlear-implant (CI) listeners. The objective of this study was to compare spectral ripple discrimination and detection in the same group of CI listeners. Ripple detection thresholds were measured over a range of ripple frequencies and were compared to spectral ripple discrimination thresholds previously obtained from the same CI listeners. The data showed that performance on the two measures was correlated, but that individual subjects' thresholds (at a constant spectral modulation depth) for the two tasks were not equivalent. In addition, spectral ripple detection was often found to be possible at higher rates than expected based on the available spectral cues, making it likely that temporal-envelope cues played a role at higher ripple rates. Finally, spectral ripple detection thresholds were compared to previously obtained speech-perception measures. Results confirmed earlier reports of a robust relationship between detection of widely spaced ripples and measures of speech recognition. In contrast, intensity difference limens for broadband noise did not correlate with spectral ripple detection measures, suggesting a dissociation between the ability to detect small changes in intensity across frequency and across time. © 2012 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4763999>]

PACS number(s): 43.66.Ts, 43.66.Fe [EB]

Pages: 3925–3934

I. INTRODUCTION

The ability to discriminate spectral shapes has relevance to the auditory processing of complex acoustic signals such as speech, as many speech cues are contained in the spectral envelope (Peterson and Barney, 1952; Pickett, 1999). One psychophysical task that is believed to measure spectral resolution is spectral ripple discrimination (Supin *et al.*, 1994; Henry *et al.*, 2005; Won *et al.*, 2007; Anderson *et al.*, 2011). For this procedure, the listener is required to discriminate between a spectrally rippled noise stimulus (i.e., broadband noise containing regular variations in amplitude along the frequency axis) and another stimulus in which the positions of spectral peaks and valleys are reversed. Typically, the spectral modulation depth is held constant and the ripple rate is adaptively varied to find the listener's ripple discrimination threshold. Spectral ripple discrimination has been shown to correlate with other measures of spectral resolution in

cochlear-implant (CI) users, such as spatial tuning curves (Anderson *et al.*, 2011).

Spectral ripple detection is a different psychoacoustic measure that also involves the use of rippled noise (Bernstein and Green, 1987; Eddins and Bero, 2007; Litvak *et al.*, 2007). In this task, the listener must distinguish a spectrally rippled noise from an unrippled (spectrally flat) noise. Unlike the spectral ripple reversal task, where the spectral ripple rate is varied and the modulation depth held constant, the spectral ripple detection paradigm typically keeps the ripple rate constant and adaptively varies the modulation depth of the spectral ripples. The spectral ripple detection threshold, or spectral modulation threshold, represents the smallest modulation depth, or spectral contrast, in a rippled noise signal (in dB) that can be discriminated from an unmodulated standard noise stimulus at a given ripple rate (Litvak *et al.*, 2007).

When spectral ripple detection thresholds are measured individually at different spectral modulation rates, the resulting pattern of thresholds as a function of modulation rate is referred to as the spectral modulation transfer function

^{a)}Author to whom correspondence should be addressed. Electronic mail: eander2@u.washington.edu

(SMTF) (Saoji and Eddins, 2007; Saoji *et al.*, 2009). Eddins and Bero (2007) measured the SMTF in normal-hearing listeners for rates between 0.25 and 10 ripples per octave (rpo), and found that the general form of the SMTF in that range is bandpass, with best modulation detection in the region between 2 and 4 rpo, and poorer detection at lower and higher modulation frequencies. Neither carrier bandwidth (1–6 octaves) nor carrier frequency region (200–12 800 Hz) was found to influence spectral ripple detection thresholds, with the exception of carrier bands restricted to very low audio frequencies (e.g., 200–400 Hz), where ripple detection was poorer. Although poorer performance at very low frequencies is qualitatively consistent with expectations based on broader auditory filters at low frequencies (Glasberg and Moore, 1990), the ripple thresholds at these very low frequencies were quantitatively poorer than predicted by excitation patterns based on expected auditory filter bandwidths.

Poor spectral resolution results in a smoothing of the spectral envelope of the internal representation of the acoustic spectrum of a complex stimulus, such as rippled noise or speech (Horst, 1987). In the case of spectral ripple discrimination, a smoothed spectral envelope will result in a lower maximum ripple rate at which ripple phase reversals are detectable. In the case of spectral ripple detection, poorer spectral resolution should lead to poorer (higher) detection thresholds at higher rates, where the smoothing effect of broader filters is most prominent. In general, spectral ripple detection is thought to be limited by poor spectral resolution, as measured by notched-noise estimates of auditory filter bandwidth (e.g., Summers and Leek, 1994). If poorer spectral resolution is thought of as a low-pass filter in the spectral modulation frequency domain (Summers and Leek, 1994; Eddins and Bero, 2007; Saoji and Eddins, 2007), then the effect of poorer spectral resolution (i.e., broader filters) will be to reduce the upper cutoff frequency of the SMTF.

Directly reducing spectral contrasts has been shown to result in decreased speech recognition performance (Bacon and Brandt, 1982; Van Tasell *et al.*, 1987; Baer and Moore, 1994), although the research literature has not consistently shown a significant correlation between measures of spectral resolution and speech perception in normal acoustic hearing (e.g., Dubno and Dirks, 1990; Surprenant and Watson, 2001). In general, speech in quiet remains intelligible even with very poor spectral resolution (e.g., Shannon *et al.*, 1995), although somewhat better spectral resolution is required to understand speech in a background of noise (e.g., Dorman *et al.*, 1998; Friesen *et al.*, 2001).

Litvak *et al.* (2007) used the spectral ripple detection paradigm to explore the role of spectral resolution as a possible source of variance in word recognition by CI users. Litvak *et al.* (2007) also introduced varying amounts of spectral smearing to vocoder-processed spectrally rippled noise stimuli by varying the vocoder's carrier filter slopes, in order to match performance of normal-hearing (NH) listeners on the spectral ripple task to that of CI listeners. Their goal was to then compare performance on measures of vowel and consonant identification by CI listeners to performance of NH

listeners with the vocoder parameters that produced the same spectral ripple detection thresholds. Average spectral ripple detection thresholds for ripple rates of 0.25 and 0.5 rpo were found to correlate strongly with vowel recognition across both groups of listeners, with vowel scores for NH listeners decreasing as the filter slopes of the noise-excited vocoder became shallower, simulating poorer spectral resolution. Spectral ripple detection thresholds also correlated with consonant recognition for both groups, but consonant performance decreased less than vowel performance for a given decrease in spectral resolution, perhaps reflecting less dependence of consonant recognition on fine-grained spectral information.

The finding of a relationship between spectral ripple detection for low ripple frequencies and speech recognition is somewhat counterintuitive, given that detection thresholds for higher, not lower, ripple frequencies should be more sensitive to reduced spectral resolution (Eddins and Bero, 2007; Saoji and Eddins, 2007). Saoji *et al.* (2009) investigated this discrepancy by comparing spectral ripple detection thresholds for an extended range of spectral ripple rates (0.25, 0.5, 1.0, and 2.0 rpo) for the same 25 subjects included in the earlier study. They found that the strongest predictors of vowel and consonant identification were detection thresholds at 0.25 and 0.5 rpo, respectively, consistent with the findings of Litvak *et al.* (2007). The authors concluded that the correlation between detection of low spectral modulation frequencies and vowel and consonant identification was not likely related to spectral resolution *per se*, but rather to differences in listeners' ability to compare the amplitudes of widely spaced spectral maxima and minima in the spectral envelope spanning a broad frequency range (cf. Bernstein and Green, 1987; Eddins and Bero, 2007). Reduced spectral resolution may prevent the use of fine spectral detail carried by higher spectral modulation frequencies, requiring CI listeners to rely on features in the broad spectral envelope (i.e., low modulation frequencies).

The aim of the present study was to compare performance on spectral ripple discrimination and detection tasks in the same CI listeners. Spectral ripple detection thresholds were measured over a range of ripple rates for a group of CI subjects and compared to the spectral ripple discrimination thresholds previously obtained from the same subjects (Anderson *et al.*, 2011). It was predicted that if both measures reflect spectral resolution, then the individual results in the two tasks should be related in a lawful manner. In addition, spectral ripple detection thresholds were compared to speech-perception measures from the same subjects, reported in Anderson *et al.* (2011), to test whether the earlier findings showing a relationship between detection of broadly spaced ripples and measures of speech recognition (Litvak *et al.*, 2007; Saoji *et al.*, 2009) could be replicated.

II. EXPERIMENT 1: SPECTRAL RIPPLE DETECTION

A. Subjects

Fifteen CI subjects (5 Clarion I, 5 Clarion II, and 5 Nucleus-22) participated in this study. See Table I for individual subject characteristics.

TABLE I. Individual subject and device characteristics. Gender, age when tested, duration of implant use prior to the study, etiology of deafness, duration of bilateral, severe-to-profound hearing loss prior to implantation, implanted device, and sound processing strategy. The processing strategies include Continuous Interleaved Sampling (CIS), Multiple Pulsatile Stimulation (MPS), HiResolution-Paired (HiRes-P) and Sequential (HiRes-S), and Spectral Peak (SPEAK).

Subject code	M/F	Age (yr)	CI use (yr)	Etiology	Duration of deafness (yr)	Device	Strategy
C03	F	58.8	9.7	Familial Progressive SNHL	27	Clarion I	CIS
C05	M	52.5	10.2	Unknown	<1	Clarion I	CIS
C16	F	54.2	6.7	Progressive SNHL	13	Clarion I	MPS
C18	M	74.0	7.2	Otosclerosis	33	Clarion I	MPS
C23	F	48.1	6.4	Progressive SNHL; Mondini's	27	Clarion I	CIS
D02	F	58.2	6.4	Unknown	1	Clarion II	HiRes-P
D05	F	78.2	6.6	Unknown	3	Clarion II	HiRes-S
D08	F	55.9	5.0	Otosclerosis	13	Clarion II	HiRes-S
D10	F	53.8	5.2	Unknown	8	Clarion II	HiRes-S
D19	F	48.2	3.5	Unknown	7	Clarion II	HiRes-S
N13	M	69.9	17.5	Hereditary; Progressive SNHL	4	Nucleus 22	SPEAK
N14	M	63.5	13.9	Progressive SNHL	1	Nucleus 22	SPEAK
N28	M	68.8	11.8	Meningitis	<1	Nucleus 22	SPEAK
N32	M	40.1	10.3	Maternal Rubella	<1	Nucleus 22	SPEAK
N34	F	62.0	8.4	Mumps; Progressive SNHL	9	Nucleus 22	SPEAK

B. Stimuli

The stimuli were patterned after those used by Litvak *et al.* (2007) and were generated in MATLAB (The Mathworks, Natick, MA), using the AFC software package developed by Stephan Ewert (University of Oldenburg, Germany). Spectral modulations were applied to a broadband (350–5600 Hz) Gaussian noise, producing sinusoidal variations in level (dB) on a log-frequency axis using

$$X(f) = 10^{(D/2)\sin(2\pi(\log_2(f/L))f_s + \theta)/20}, \quad (1)$$

where X is the expected level at frequency f , D is peak-to-valley modulation depth in dB, L is the low cutoff frequency of the noise passband, f_s is the spectral modulation frequency (in rpo), and θ is the starting phase of the ripple function. Noises were generated at seven fixed ripple rates: 0.25, 0.5, 0.75, 1.0, 1.5, 2.0, and 3.0 rpo. The modulation depth was varied adaptively for this task. The signal duration was 400 ms, which included 20 ms raised-cosine onset and offset ramps. The stimuli were played through a Lynx L22 sound card (Lynx Studio Technology, Costa Mesa, CA) at a sampling rate of 48 kHz and presented in a double-walled, sound-attenuating booth through a single loudspeaker (Infinity RS1000, Harman International, Stamford, CT) located 1 m from the subject's seated position at 0° azimuth, at approximately head height. The average (root-mean-square) sound level of the noise was 60 dBA when measured at a point corresponding to the position of the subject's head. The level was roved across intervals within each trial by ± 3 dB, and the starting phase of the spectral modulation was randomized for each trial.

C. Procedure

Subjects wore their own speech processors at typical use settings. The same settings were used as with the previously obtained measures of spectral ripple discrimination in the same listeners (Anderson *et al.*, 2011). A three-interval, three-alternative forced-choice (3I-3AFC)

adaptive procedure was employed. During each trial, subjects heard two intervals of unrippled broadband noise and one interval of rippled noise. The interval containing the rippled noise was selected at random with equal *a priori* probability on each trial. Subjects indicated which interval they judged as sounding different (corresponding to the rippled signal) by selecting the appropriate button on a computer screen offset to the right of the speaker. Correct-answer feedback was provided after each trial. Each test run began with a peak-to-valley ratio for the rippled stimulus of 20 dB; the modulation depth was then varied adaptively in a two-down, one-up adaptive procedure to track the peak-to-valley ratio that could be detected with an accuracy of about 71% correct (Levitt, 1971). The initial step size was 4 dB, changing to 2 dB after the first two reversals, and to 0.5 dB following two more reversals. Termination of the run occurred after a total of ten reversals. Spectral ripple detection threshold was defined for each run as the geometric mean of the peak-to-valley ratios (PVR) at the final six reversal points.

Each subject completed six runs for each of the seven ripple rates, for a total of 42 experimental runs. Thresholds from the first run for each ripple-rate condition were discarded as practice runs, and any thresholds that were more than 3 standard deviations (s.d.) away from the mean of the remaining measurements for that condition were excluded as outliers. Typically, thresholds from five runs for each ripple rate were used to calculate a geometric mean threshold for each subject.

D. Results

Spectral ripple detection thresholds varied widely across subjects and across ripple frequencies. Individual thresholds ranged from about 5 to 45 dB PVR. The pattern of results is in line with that reported by Saoji *et al.* (2009). In general, greater spectral modulation was required for detection as the ripple frequency increased, although some subjects showed non-monotonicities in their SMTFs, as illustrated in Fig. 1. In addition, a range of different shapes of SMTFs can

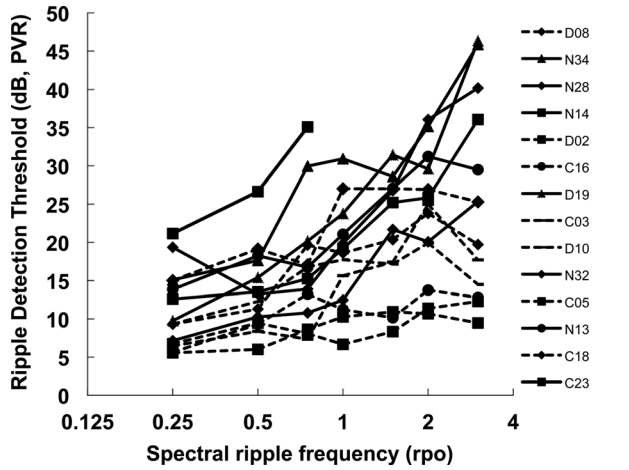


FIG. 1. Spectral modulation transfer functions (SMTFs) for 15 CI subjects. Ripple detection threshold (in dB) for each of seven different spectral ripple frequencies (0.25, 0.5, 0.75, 1.0, 1.5, 2.0, and 3.0 rpo) is depicted.

be observed, similar to the findings of *Saoji et al. (2009)*, with some individual functions steeper than others. As expected from Fig. 1, a repeated-measures analysis of variance (ANOVA) indicated a main effect of ripple frequency [$F(2.29, 29.73) = 26.04, p < 0.001$]. (Here and elsewhere, a Greenhouse–Geisser correction was applied where appropriate to account for lack of sphericity.)

The spectral ripple detection thresholds were compared to spectral ripple discrimination thresholds from *Anderson et al. (2011, Table II)*. The discrimination thresholds were measured with stimuli using a 30 dB spectral PVR; this value represents an approximation of “maximum” depth, close to 95% modulation on a linear modulation scale.

1. Comparing spectral ripple detection and discrimination

If spectral ripple detection and discrimination are both mediated by a common underlying mechanism, then performance in the two tasks should be related. For instance, if it is necessary to detect spectral modulation in order to discriminate changes in its phase, then we might expect that the ripple frequency at which a detection threshold of 30 dB is obtained should correspond to the spectral ripple discrimination threshold, since the modulation depth used in the ripple discrimination task was 30 dB.¹ However, the measured spectral ripple detection thresholds were better (lower) than 30 dB for some subjects even at the highest ripple rate originally tested (3 rpo). Consequently, the spectral ripple detection task was conducted at additional higher ripple frequencies (6, 9, 12, 15, 18, and 21 rpo) for each subject, as needed, to achieve a ripple detection threshold at or above 30 dB. From this extended SMTF, the spectral ripple frequency corresponding to a ripple detection threshold of 30 dB was estimated by a linear interpolation from existing points on the SMTF. Simple linear interpolation, rather than defining best-fit functions for the data points (e.g., *Saoji et al., 2009*), was chosen because of the irregularity and non-monotonicities present in some of the SMTFs. Figure 2 displays the extended SMTF measures for each subject,

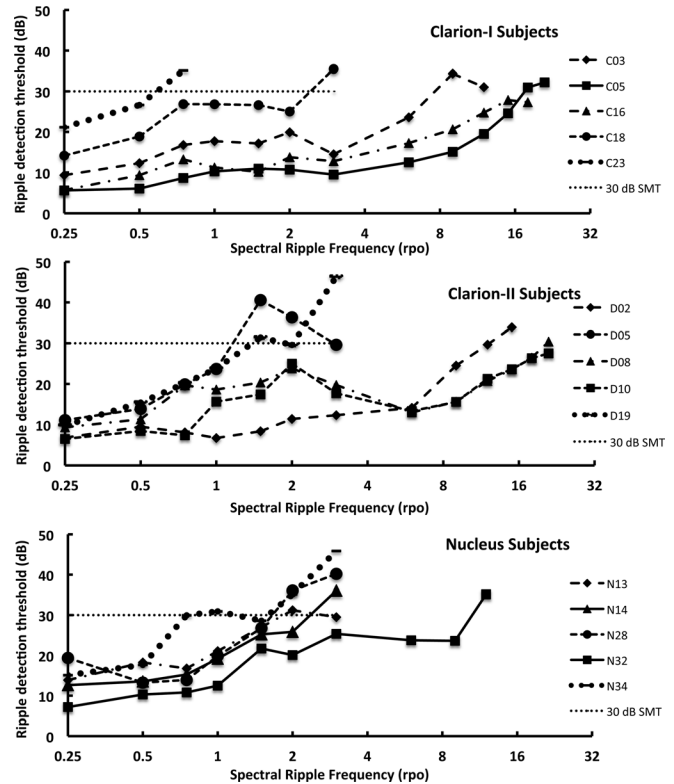


FIG. 2. SMTFs for Clarion-I, Clarion-II, and Nucleus subjects (top, middle, and lower panels, respectively) over an extended range of spectral ripple rates. The dotted horizontal lines correspond to a ripple detection threshold of 30 dB.

grouped by implant device. Note the apparent differences in results as a function of device type. Using a between-subjects ANOVA on the ripple detection thresholds for 3 rpo, a marginally significant effect of device was found [$F(2,13) = 4.313, p = 0.04$], presumably reflecting the generally lower limits found in the Nucleus users; however, the individual differences are so large as to preclude any general statement. The same analysis using the interpolated 30 dB threshold in the detection task showed no significant differences between devices [$F(2,14) = 1.935; p = 0.19$].

Figure 3 illustrates the relationship between the interpolated spectral ripple detection threshold and spectral ripple discrimination thresholds obtained from *Anderson et al. (2011)*. Linear regression analysis revealed a significant correlation between the two variables ($r^2 = 0.42, p = 0.01$). Despite the significant correlation, the correspondence is surprisingly poor. For instance, listeners with very similar ripple discrimination thresholds (between 2 and 3 rpo) had 30 dB detection thresholds at rates ranging from about 3 to more than 17 rpo. For most (but not all) listeners, the spectral ripple rate at which a 30 dB PVR was detected was substantially higher than the corresponding ripple discrimination threshold, and the function relating the ripple rate corresponding to a detection threshold of 30 dB to the ripple discrimination threshold had a slope of approximately 5. This pattern of results is difficult to explain in terms of the detectability of the spectral contrasts within each stimulus.¹ Potential reasons for this apparent discrepancy are addressed in Sec. III E.

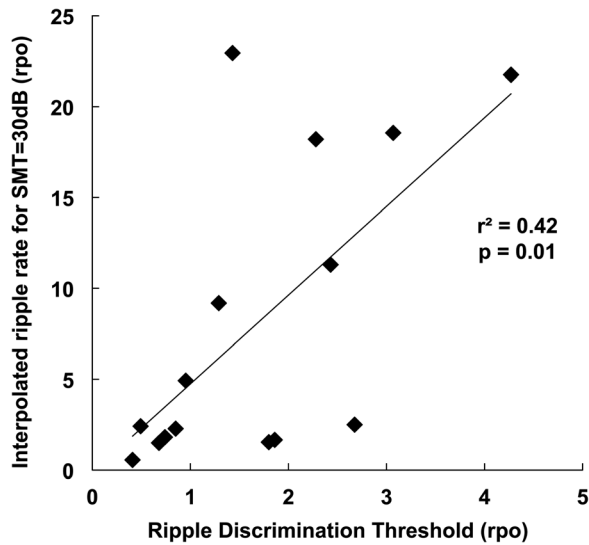


FIG. 3. Interpolated ripple frequency (in rpo) for ripple detection threshold, or spectral modulation threshold (SMT) = 30 dB, as a function of spectral ripple discrimination threshold (rpo).

2. Spectral ripple detection and speech recognition

Spectral ripple detection thresholds were examined in relation to sentence and vowel recognition performance [see Anderson *et al.* (2011) for additional details on speech recognition testing]. Two lists of IEEE sentences (IEEE, 1969) recorded by one male and one female talker were administered for each experimental condition. Each list contained ten sentences with five key words, for a total of 100 key words per condition. Subjects orally repeated the words they heard, following presentation. Hillenbrand vowels (Hillenbrand *et al.*, 1995) included 11 vowels in an h/V/d context spoken by six male talkers. Subjects identified the token they heard by selecting the corresponding item from the full set of possible items on a computer screen. Stationary Gaussian background noise was spectrally shaped to match the long-term average spectrum of each type of speech material (sentences and vowels). Speech and noise were mixed to produce the appropriate speech-to-noise ratios (SNRs), and acoustic verification was performed with a sound level meter. Speech materials were presented in the sound field at 65 dBA, in quiet and at fixed SNRs (0, 5, 10, 15, and 20 dB). Performance in quiet was quantified in rationalized arcsine units, or RAU scores (Studebaker, 1985), representing correct key word recognition for sentence materials and for vowel identification. Performance in noise is reported as the SNR corresponding to 50% correct recognition, derived from logistic functions fitted to the speech psychometric functions (percent correct as a function of SNR).

To analyze the relationship between speech recognition and spectral ripple detection, three summary measures of ripple detection were derived. The first was the average threshold at ripple rates of 0.25 and 0.5 rpo, representing performance at low ripple rates. The second was the average threshold at ripple rates of 2 and 3 rpo, representing performance at higher ripple rates. The third measure was the

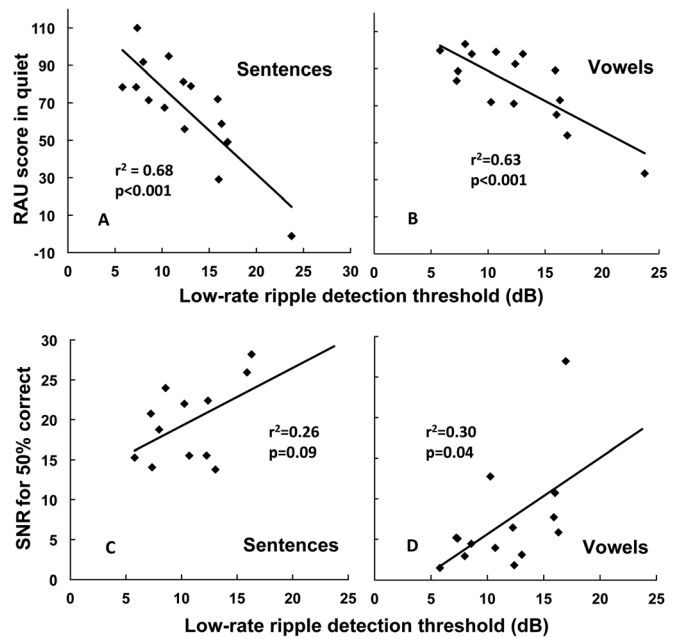


FIG. 4. RAU scores for sentence recognition (A) and vowel recognition (B) as a function of average SMT for 0.25 and 0.5 rpo. (C), (D) SNR_{50%} for sentences and vowels, respectively, as a function of the average of ripple detection thresholds at 0.25 and 0.5 rpo.

interpolated ripple rate at which the threshold PVR was 30 dB, representing the upper limit of spectral ripple detection.

The average of spectral ripple detection thresholds at 0.25 and 0.5 rpo showed a strong relationship with sentence recognition in quiet ($r^2 = 0.68$, $p < 0.001$), as shown in Fig. 4(A), and vowel recognition in quiet ($r^2 = 0.63$, $p < 0.001$), shown in Fig. 4(B). Data from all 15 subjects are included.

The relationships between spectral ripple detection and sentence and vowel recognition in noise are depicted in Figs. 4(C) and 4(D), respectively. Plotted are the SNRs corresponding to 50% correct performance as a function of ripple detection threshold. Three subjects (C18, C23, N28) were excluded from sentence analysis and one (C23) from vowel analysis because their asymptotic performance in quiet was below 50%. The SNR required for 50% word recognition in sentences for the remaining 12 subjects showed trends, but the relationships failed to reach significance ($r^2 = 0.26$, $p = 0.09$); the corresponding SNR for vowel recognition in noise for 14 subjects showed a borderline significant correlation with ripple detection ($r^2 = 0.30$, $p = 0.04$). The poorer correlations found for speech in noise may be due, at least in part, to the smaller number of subjects in the noise conditions.

Table II displays results from all of the regression analyses for each of the three summary ripple detection measures, with each of the speech measures obtained, uncorrected for multiple comparisons. The average threshold at ripple rates of 0.25 and 0.5 rpo shows consistently strong correlations for sentence and vowel recognition in quiet, whereas correlations with speech measures at higher ripple rates are less robust, in agreement with other studies (Litvak *et al.*, 2007; Saoji *et al.*, 2009).

TABLE II. Correlations and corresponding p values (with no post-hoc Bonferroni corrections for multiple comparisons) for three summary measures of ripple detection (the average threshold at ripple rates of 0.25 and 0.5 rpo, the average threshold at ripple rates of 2 and 3 rpo, and the interpolated ripple rate at which the threshold peak-to-valley ratio was 30 dB), and speech recognition measures.

Speech measure	Ripple detection measure		
	Avg. (0.25, 0.5 rpo)	Avg. (2.0, 3.0 rpo)	SMT _{30 dB}
Sentence recognition, Q (RAU)	$r^2 = 0.68$ $p < 0.001$	$r^2 = 0.37$ $p = 0.02$	$r^2 = 0.33$ $p = 0.02$
Vowel identification, Q (RAU)	$r^2 = 0.63$ $p < 0.001$	$r^2 = 0.36$ $p = 0.02$	$r^2 = 0.16$ $p = 0.15$
SNR _{50%} , sentences	$r^2 = 0.26$ $p = 0.09$	$r^2 = 0.14$ $p = 0.24$	$r^2 = 0.09$ $p = 0.35$
SNR _{50%} , vowels	$r^2 = 0.30$ $p = 0.04$	$r^2 = 0.09$ $p = 0.31$	$r^2 = 0.04$ $p = 0.51$

E. Discussion

1. Comparing spectral ripple detection and discrimination as measures of spectral resolution

One aim of this study was to test whether spectral ripple detection and discrimination were both measures of the same underlying mechanisms involving spectral resolution. Although the two measures were correlated, the relationship was not clear-cut and did not conform to expectations based on predictions from simple models of detection theory. The results fail to support the suggestion of [Saoji et al. \(2009\)](#) that thresholds from the spectral ripple discrimination task represent points on the high-modulation-frequency end of the SMTF. In fact, spectral ripple detection for a modulation depth of 30 dB generally fell at a higher ripple rate than the corresponding ripple discrimination thresholds.

The better performance for the detection task might reflect differences in listening strategies used by individuals for the two tasks. For instance, a strategy for the spectral ripple detection task might be to compare each interval to an internal template of the standard stimulus, unmodulated noise, which will remain constant across all trials (e.g., [Sabin et al., 2012](#)). On the other hand, for the discrimination task a listener would not have a constant internal template, because of the randomized starting phase used in all tasks, so that a direct comparison of all three intervals with each other is necessary for successful performance. The demands on working memory may therefore be greater than for the detection task.

However, the extended-frequency ripple detection results raise a further question of the validity of the ripple detection technique as a measure of spectral resolution. Some CI users were able to detect spectral ripples as closely spaced as 20 rpo, well beyond their spectral ripple discrimination thresholds and predicted capabilities based on the bandwidths of the CI analysis filters. The filters of the individual subjects were never narrower than 1/4 octave on average, meaning that spectral ripple rates greater than 4 rpo should not have been reliably detected. Figure 5 demonstrates an example, where a non-monotonicity occurs in the function, with a dip

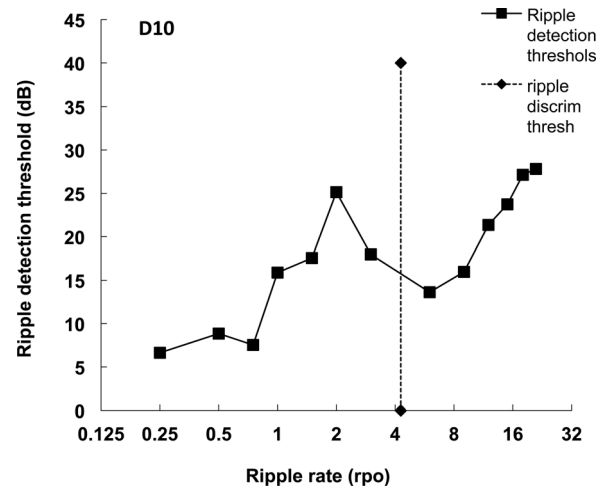


FIG. 5. Individual SMTF (ripple detection threshold as a function of spectral ripple frequency) for subject D10. The vertical line indicates the spectral ripple discrimination threshold for this listener.

representing better performance (smaller modulation depth at threshold) at higher ripple densities. The non-monotonicity may indicate a transition from one cue (e.g., spectral contrast) at low ripple rates to another cue at higher rates.

To further investigate factors underlying spectral ripple detection, three NH subjects (ages 34–52 with audiometric thresholds no greater than 20 dB HL at octave frequencies between 250 and 8000 Hz) were tested on an identical spectral ripple detection procedure in the sound field, except that an extended range of spectral ripple rates was used: 0.25, 0.5, 0.75, 1.0, 1.5, 2.0, 3.0, 6.0, 9.0, 12.0, 15.0, 21.0, 30.0, 45.0, and 60.0 rpo. Note that according to estimates of human auditory filter bandwidths (e.g., [Glasberg and Moore, 1990](#); [Oxenham and Shera, 2003](#)), filters are typically no narrower than one-sixth to one-tenth of an octave (depending on the center frequency and the method of measurement). Therefore, the detection of spectral ripples should be limited to ripple rates below about 10 rpo. (This expected limit may explain why earlier studies in NH listeners have typically been limited to rates below about 10 rpo.) The results from the three NH listeners are shown in Fig. 6. Measurable thresholds were obtained for ripple rates well beyond the

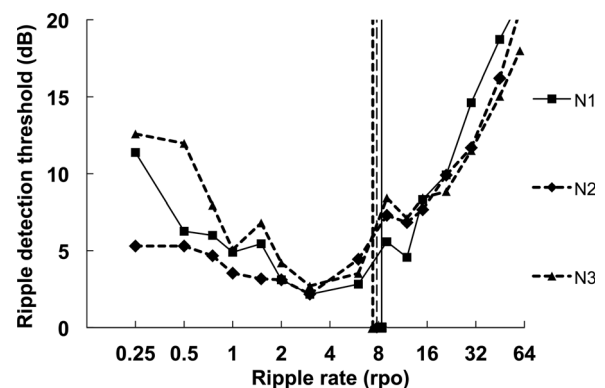


FIG. 6. SMTFs for three normal-hearing subjects. Ripple detection thresholds (in dB) for each of 15 different spectral ripple frequencies (0.25–60 rpo) is displayed. Spectral ripple discrimination thresholds are indicated by vertical lines.

expected theoretical limit, and all three subjects were able to detect ripples out to the highest ripple rate tested of 60 rpo. Although not as prominent as in Fig. 5, there is a hint of a non-monotonicity in the function, with a local peak just below 10 rpo, which is found in all three NH listeners.

It seems highly unlikely that listeners are able to resolve spectral peaks presented at 60 rpo; instead, some other cue must be available. It is possible that the interactions between adjacent peaks in the spectrally rippled noise provide usable temporal cues. Specifically, closely spaced spectral peaks might produce regular temporal fluctuations or beats in the temporal envelope, providing a salient cue for discriminating the rippled signal from the flat-spectrum noise standard. Viewed in this light, the non-monotonicity illustrated in Fig. 5, and to a lesser extent in Fig. 6, may reflect two regions in which different cues are being used to perform the task: At low ripple rates, subjects are detecting the spectral ripples via a spectral-contrast mechanism because the temporal cues, based on peaks that are spaced far apart, are too weak because of the high rate of fluctuation, based on the difference in frequency of two adjacent peaks, and because of the attenuating effects of the auditory filters. At higher ripple rates, the spectral cues become weaker, but the temporal cues become more salient, as the denser spectral peaks produce slower and more detectable temporal fluctuations. To our knowledge, there exist no published reports addressing this possibility systematically, but work using iterated rippled noise points to the fact that temporal periodicity in noise waveforms can be observed as spectral peaks, and vice versa (e.g., Yost, 1996; Yost *et al.*, 1998; Wiegube and Patterson, 1999).

In summary, spectral ripple detection at ripple rates higher than 2–3 rpo in CI users (or 8–10 rpo in NH listeners) may be influenced by temporal-envelope cues, and therefore do not solely reflect spectral resolution. This problem may not apply to spectral ripple discrimination, because the temporal-envelope fluctuations would be present in both the target and reference stimuli and therefore may not provide a reliable cue to perform the task.

Returning to spectral ripple detection at very low ripple rates, it is unclear the extent to which the results reflect spectral resolution, as opposed to spectral profile discrimination (e.g., Bernstein and Green, 1988), or intensity resolution: At low ripple rates, if the spectral ripples of the stimulus are much broader than the bandwidths of the effective analysis filters (either in the CI processor or in terms of current spread within the cochlea), then changes in the effective filter bandwidths are unlikely to affect spectral ripple detection thresholds. Instead, thresholds are likely to be determined primarily by the subject's ability to detect across-frequency (or across-channel) differences in intensity. If so, then it is possible that this ability is related to intensity difference limens—the just-noticeable difference in intensity of a sound. This hypothesis is tested in experiment 2.

2. Relationship between ripple detection and speech perception

Three measures of spectral ripple detection were compared to speech recognition performance: (1) Detection

threshold for broadly spaced spectral ripples (the average of 0.25 and 0.5 rpo); (2) detection threshold for more closely spaced ripples (the average of 2.0 and 3.0 rpo); (3) the interpolated ripple density from each subject's SMTF corresponding to a 30 dB peak-to-valley modulation depth. It might be predicted that the ability to detect more closely spaced ripples would be most predictive of speech perception, given that this measure is most likely to reflect spectral resolution, rather than intensity resolution (low rates) or temporal processing (interpolated 30 dB threshold). However, the measure of ripple detection that correlated most strongly with speech recognition measures in quiet was for the low ripple rates of 0.25 and 0.5 rpo, corroborating the results reported by Litvak *et al.* (2007) and Saoji *et al.* (2009). Correlations were not as robust for any of the speech measures in noise, in part perhaps because of the smaller number of subjects included in the comparison.

In general, the relationship between detection of low spectral ripple rates and speech perception is consistent with a study by Liu and Eddins (2008), in which they measured vowel identification by NH listeners for vowel stimuli that were progressively high-pass filtered in the spectral modulation domain. Their data suggested that the spectral cues most important for vowel identification are represented by spectral modulation frequencies below 2 cycles/octave.

If spectral resolution strongly influences speech perception, then these results are somewhat counterintuitive, given that detection thresholds for higher ripple frequencies, not lower ones, should be more sensitive to reduced spectral resolution and therefore might be expected to correlate more strongly with speech recognition. Saoji *et al.* (2009) suggested that the correlation between detection of low spectral modulation frequencies and vowel and consonant identification might not be related to spectral resolution *per se*, but rather to differences in listeners' ability to compare widely spaced spectral maxima and minima in the spectral envelope spanning a broad frequency range, i.e., profile analysis (Green *et al.*, 1984; Bernstein and Green, 1988). In other words, adequate speech perception may not require access to high spectral modulation frequencies, so long as the intensity contrasts at lower spectral modulation frequencies are well resolved. Another possibility is that subjects were basing their judgments on small changes in overall loudness, rather than spectral contrasts. Although this possibility is rendered less likely by our use of level roving and randomized starting modulation phase, the ability of subjects to make use of any overall loudness cue would again depend on their ability to resolve small differences in intensity. This possibility is explored further in experiment 2.

III. EXPERIMENT 2: INTENSITY DISCRIMINATION

In this experiment, intensity difference limens were measured using broadband, spectrally flat noise stimuli, similar in bandwidth to the stimuli from experiment 1. The rationale was based on the possibility that performance in the spectral ripple detection task, particularly at low ripple rates, was governed primarily by intensity resolution, rather than spectral resolution. This was mostly likely to be in the

form of spectral profile analysis, with across-frequency (or across-channel) comparisons of intensity, but might also be in the form of comparisons of the overall perceived intensity (loudness) of the spectrally modulated and unmodulated stimuli. In either case, we might expect to see a relationship between spectral ripple detection thresholds at low ripple rates and intensity difference limens.

A. Subjects

Fourteen of the original 15 CI subjects (4 Clarion I, 5 Clarion II, and 5 Nucleus-22) participated in this study. Subject C23 was no longer available to participate.

B. Stimuli

Broadband (350–5600 Hz) Gaussian noise stimuli were generated using MATLAB software. The stimulus duration was 400 ms, including 20 ms raised-cosine onset and offset ramps. The intensity of the target stimulus was increased by adding a synchronous, uncorrelated but statistically identical noise (the “signal”) to one of the standard stimuli. The stimuli were presented in a double-walled, sound-attenuating booth through a single loudspeaker (Infinity RS1000) located 1 m from the subject’s seated position at approximately head height. The average sound level of the standard noise was set to 60 dBA when measured at the location corresponding to the subject’s head.

C. Procedure

Subjects wore their own speech processors at typical use settings, as in experiment 1. A 3I-3AFC adaptive procedure was used. The overall level of the stimuli was roved by ± 3 dB across trials with uniform distribution. During each trial, subjects heard two intervals of the standard noise and one interval of a noise that was higher in level than the other two. The presentation order of the intervals within each trial was randomized, so that all intervals had the same *a priori* probability of containing the signal on each trial. Subjects indicated which interval they judged as louder by selecting the appropriate virtual button on a computer screen. Correct-answer feedback was provided after each trial. The level of the signal was varied adaptively in a two-down, one-up psychometric procedure to track the 70.7% point on the psychometric function. The initial signal level was 10 dB higher than the standard stimulus, leading to an overall level difference (ΔL) between the target interval (i.e., the interval containing the signal) and reference intervals of about 10.4 dB. Initially, the level of the signal was varied in steps of 4 dB. After four reversals in the tracking procedure, the step size was reduced to 2 dB for the last six reversals, and threshold was defined for each run as the arithmetic mean of the level difference between the signal and reference noise at the final six reversal points.

Each subject completed six runs of the intensity discrimination experiment. Thresholds from the first run were eliminated as “practice” runs, and any thresholds that were more than 3 s.d. away from the mean of the remaining measurements for that condition were excluded; in general, thresholds

from five runs were used to calculate an arithmetic mean threshold for each subject. These threshold values, denoting the signal-to-standard ratio, in dB, were used for all statistical analyses. However, for the purposes of display, these ratios were converted to ΔL values, i.e., the level difference between the standard interval and the target interval in dB.

D. Results

Intensity discrimination thresholds varied widely across subjects. Individual level difference limens (ΔL) ranged from about 1.0 to 6.6 dB. Figure 7 displays ΔL for each subject.

Intensity difference limens were compared to spectral ripple detection thresholds, using regression analysis; intensity discrimination did not correlate with spectral ripple detection at any modulation frequency, or with spectral ripple discrimination thresholds (from Anderson *et al.*, 2011); all regression analyses resulted in $p > 0.05$.

E. Discussion

If spectral ripple detection at very low ripple rates by CI users depends on the ability to discriminate differences in intensity across frequency, then one might intuitively predict a correlation between low-rate ripple detection (as measured in experiment 1) and intensity discrimination of gated noise across time (as measured in this experiment). The fact that none was found suggests that different mechanisms mediate intensity discrimination over time (as in the intensity discrimination experiment) and intensity discrimination across frequency (as in spectral profile analysis). In fact, an earlier study by Green and Mason (1985), comparing spectral profile analysis and intensity discrimination (using a simple Weber fraction experiment with single pure tone signals), also showed no relationship between performance on the two tasks, in line with our findings.

If detection of ripples at higher ripple densities were mediated by overall loudness or intensity cues, then we might expect a correlation between performance on spectral ripple detection for high ripple rates and intensity discrimination tasks. The fact that none was found may be taken as an indication that, as expected, overall loudness cues did not play a major role in determining spectral ripple detection thresholds.

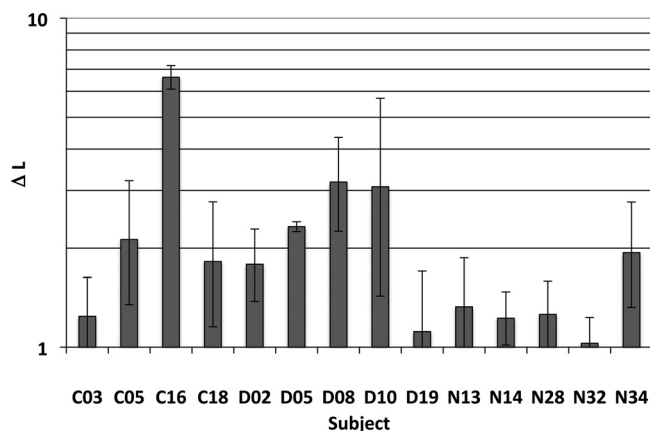


FIG. 7. Intensity difference limens (ΔL) for each subject. Error bars represent 1 s.d.

IV. SUMMARY

In order to better understand the relationship between spectral resolution and speech perception, this study examined the relations between performance on spectral ripple discrimination, spectral ripple detection, and intensity discrimination. In particular, the correspondence between spectral ripple detection measures at higher ripple rates and ripple discrimination thresholds was examined to test the hypothesis that they both reflect underlying spectral resolution. Although the measures were correlated, they were not related by any obvious mechanistic relationship, suggesting that the two tasks may be tapping into different auditory processes. In particular, the fact that spectral ripple detection was often possible at much higher ripple rates than predicted by the underlying spectral resolution in both CI and NH listeners suggests that subjects were able to use cues related to regularities in the temporal structure, unrelated to spectral resolution.

Strong relationships between spectral ripple detection and sentence and vowel recognition were found, in agreement with other reports (e.g., Litvak *et al.*, 2007; Saoji *et al.*, 2009). The ripple detection measures that correlated best with speech perception were those at very low spectral rates, which seem less likely to test spectral resolution *per se*. Rather, the spectral ripple detection measures may reflect across-channel capacities related to spectral intensity profile analysis. However, in line with earlier studies in NH listeners, no correlation was found between across-channel intensity comparisons (experiment 1) and across-time intensity comparisons (experiment 2), suggesting that the mechanisms underlying the two tasks are not identical.

In summary, the relationship between spectral ripple detection and spectral ripple discrimination cannot be easily explained in terms of a single underlying factor, such as peripheral spectral resolution. Spectral ripple detection at high ripple rates seems to be influenced by other factors, such as potential temporal-envelope cues, making such high-rate measures unsuitable as measures of spectral resolution. Nevertheless, the results support the earlier finding (Litvak *et al.*, 2007; Saoji *et al.*, 2009) of high correlations between ripple detection at low ripple rates and speech perception, suggesting that low-rate spectral ripple detection remains a viable diagnostic tool for assessing speech-relevant auditory capabilities in CI users.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health (Grant No. R01 DC 006699 and R01 DC 008306) and by the Lions International Hearing Foundation. We thank Heather Kreft for her assistance on this project, Christophe Micheyl for his helpful insights, and all of our research subjects for their participation.

¹Different expectations for the relationship between spectral ripple discrimination and detection can be derived based on different assumptions with respect to the decision rule. For instance, if the maximal spectral contrast in each interval (in dB) is compared within each trial, then the difference

between two out-of-phase rippled stimuli will be twice that of the difference between a rippled and a flat stimulus. If anything, therefore, one might expect limits in the detection task to be reached at a lower ripple rate than in the discrimination task. As shown in Fig. 2, the opposite was found, further suggesting different detection cues in the two tasks, at least at high ripple rates and large modulation depths.

- Anderson, E. S., Nelson, D. A., Kreft, H., Nelson, P. B., and Oxenham, A. J. (2011). "Comparing spatial tuning curves, spectral ripple resolution, and speech perception in cochlear implant users." *J. Acoust. Soc. Am.* **130**, 364–375.
- Bacon, S. P., and Brandt, J. F. (1982). "Auditory processing of vowels by normal-hearing and hearing-impaired listeners." *J. Speech Hear. Res.* **25**, 339–347.
- Baer, T., and Moore, B. C. J. (1994). "Spectral enhancement to compensate for reduced frequency selectivity." *J. Acoust. Soc. Am.* **95**, 2992.
- Bernstein, L. R., and Green, D. M. (1987). "Detection of simple and complex changes of spectral shape." *J. Acoust. Soc. Am.* **82**, 1587–1592.
- Bernstein, L. R., and Green, D. M. (1988). "Detection of changes in spectral shape: Uniform vs. non-uniform background spectra." *Hear. Res.* **34**, 157–165.
- Dorman, M., Loizou, P., Fitzke, J., and Tu, Z. (1998). "The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6–20 channels." *J. Acoust. Soc. Am.* **104**, 3583–3585.
- Dubno, J. R., and Dirks, D. D. (1990). "Associations among frequency and temporal resolution and consonant recognition for hearing-impaired listeners." *Acta Oto-Laryngol., Suppl.* **469**, 23–29.
- Eddins, D. A., and Bero, E. M. (2007). "Spectral modulation detection as a function of modulation frequency, carrier bandwidth, and carrier frequency region." *J. Acoust. Soc. Am.* **121**, 363–372.
- Friesen, L. M., Shannon, R. V., Baskent, D., and Wang, X. (2001). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants." *J. Acoust. Soc. Am.* **110**, 1150–1163.
- Glasberg, B. R., and Moore, B. C. J. (1990). "Derivation of auditory filter shapes from notched-noise data." *Hear. Res.* **47**, 103–138.
- Green, D. M., and Mason, C. R. (1985). "Auditory profile analysis: Frequency, phase, and Weber's law." *J. Acoust. Soc. Am.* **77**, 1155–1161.
- Green, D. M., Mason, C. R., and Kidd, G., Jr. (1984). "Profile analysis: Critical bands and duration." *J. Acoust. Soc. Am.* **75**, 1163–1167.
- Henry, B. A., Turner, C. W., and Behrens, A. (2005). "Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners." *J. Acoust. Soc. Am.* **118**, 1111–1121.
- Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (1995). "Acoustic characteristics of American English vowels." *J. Acoust. Soc. Am.* **97**, 3099–3111.
- Horst, J. W. (1987). "Frequency discrimination of complex signals, frequency selectivity, and speech perception in hearing-impaired subjects." *J. Acoust. Soc. Am.* **82**, 874–885.
- IEEE (1969). "IEEE recommended practice for speech quality measurements." *IEEE Trans. Audio Electroacoust.* **AU-17**, 225–246.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics." *J. Acoust. Soc. Am.* **49**, 467–477.
- Litvak, L. M., Spahr, A. J., Saoji, A. A., and Fridman, G. Y. (2007). "Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners." *J. Acoust. Soc. Am.* **122**, 982–991.
- Liu, C., and Eddins, D. A. (2008). "Effects of spectral modulation filtering on vowel identification." *J. Acoust. Soc. Am.* **124**, 1704–1715.
- Oxenham, A. J., and Shera, C. S. (2003). "Estimates of human cochlear tuning at low levels using forward and simultaneous masking." *J. Assoc. Res. Otolaryngol.* **4**, 541–554.
- Peterson, G. E., and Barney, H. L. (1952). "Control methods used in a study of the vowels." *J. Acoust. Soc. Am.* **24**, 175–184.
- Pickett, B. (1999). *The Acoustics of Speech Communication* (Allyn & Bacon, Needham Heights, MA).
- Sabin, A. T., Eddins, D. A., and Wright, B. A. (2012). "Perceptual learning evidence for tuning to spectrotemporal modulation in the human auditory system." *J. Neurosci.* **32**, 6542–6549.

- Saoji, A. A., and Eddins, D. A. (2007). "Spectral modulation masking patterns reveal tuning to spectral envelope frequency," *J. Acoust. Soc. Am.* **122**, 1004–1013.
- Saoji, A. A., Litvak, L., Spahr, A. J., and Eddins, D. A. (2009). "Spectral modulation detection and vowel and consonant identifications in cochlear implant listeners," *J. Acoust. Soc. Am.* **126**, 955–958.
- Shannon, R. V., Zeng, F., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.
- Studebaker, G. A. (1985). "A 'rationalized' arcsine transform," *J. Speech Hear. Res.* **28**, 455–462.
- Summers, V., and Leek, M. L. (1994). "The internal representation of spectral contrast in hearing-impaired listeners," *J. Acoust. Soc. Am.* **95**, 3518–3528.
- Supin, A., Popov, V. V., Milekhina, O. N., and Tarakanov, M. B. (1994). "Frequency resolving power measured by rippled noise," *Hear. Res.* **78**, 31–40.
- Surprenant, A. M., and Watson, C. S. (2001). "Individual differences in the processing of speech and nonspeech sounds by normal-hearing listeners," *J. Acoust. Soc. Am.* **110**, 2085–2095.
- Van Tasell, D. J., Soli, S. D., Kirby, V. M., and Widin, G. P. (1987). "Speech waveform envelope cues for consonant recognition," *J. Acoust. Soc. Am.* **82**, 1152–1161.
- Wiegrebe, L., and Patterson, R. D. (1999). "The role of envelope modulation in spectrally unresolved iterated rippled noise," *Hear. Res.* **132**, 94–108.
- Won, J. H., Drennan, W. R., and Rubinstein, J. T. (2007). "Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users," *J. Assoc. Res. Otolaryngol.* **8**, 384–392.
- Yost, W. A. (1996). "Pitch of iterated rippled noise," *J. Acoust. Soc. Am.* **100**, 511–518.
- Yost, W. A., Patterson, R., and Sheft, S. (1998). "The role of the envelope in processing iterated rippled noise," *J. Acoust. Soc. Am.* **104**, 2349–2361.