

The development of a modified spectral ripple test

Justin M. Aronoff^{a)} and David M. Landsberger

Communication and Neuroscience Division, House Research Institute, 2100 West 3rd Street,
Los Angeles, California 90057
jaronoff@hei.org, dlandsberger@hei.org

Abstract: Poor spectral resolution can be a limiting factor for hearing impaired listeners, particularly for complex listening tasks such as speech understanding in noise. Spectral ripple tests are commonly used to measure spectral resolution, but these tests contain a number of potential confounds that can make interpretation of the results difficult. To measure spectral resolution while avoiding those confounds, a modified spectral ripple test with dynamically changing ripples was created, referred to as the spectral-temporally modulated ripple test (SMRT). This paper describes the SMRT and provides evidence that it is sensitive to changes in spectral resolution.

© 2013 Acoustical Society of America

PACS numbers: 43.66.Fe, 43.66.Yw, 43.66.Ts [DOS]

Date Received: April 6, 2013 Date Accepted: June 28, 2013

1. Introduction

Poor spectral resolution can be a limiting factor for hearing impaired listeners, particularly for complex listening tasks such as speech understanding in noise (Moore, 1996; Friesen *et al.*, 2001). Spectral ripple tests are commonly used to measure cochlear implant users' spectral resolution (e.g., Henry and Turner, 2003; Litvak *et al.*, 2007; Anderson *et al.*, 2011; Jones *et al.*, 2013). With these tests, listeners are asked to discriminate a spectrally rippled stimulus (i.e., a stimulus that is amplitude modulated in the frequency domain) from either a stimulus without a spectral ripple, a phase-reversed spectrally rippled stimulus where the phase of the amplitude modulation is inverted, or a stimulus with a different ripple density (i.e., a different number of ripples per octave). Previous research has indicated that spectral ripple tests can predict speech perception in noise performance (Won *et al.*, 2007) and are correlated with channel interaction in cochlear implant users (Jones *et al.*, 2013). However, as described in Azadpour and McKay (2012), there are a number of potential confounds that can make interpretation of the results difficult. One potential confound is local loudness cues. If a participant can attend to a selected frequency region that contains less than one ripple, there will be an audible loudness difference between the target and reference stimuli in that selected region. Another potential cue occurs at the upper or lower frequency boundaries of the stimulus. If participants are able to attend to either of these regions, the highest or lowest audible frequency will differ between the target and reference stimuli, providing a potential cue. Similarly, if participants can attend to the spectral centroid (i.e., the weighted mean frequency), they may hear an audible shift between the target and reference stimuli. These potential confounds are illustrated in the left half of Fig. 1. Although previous research has validated that the results of a spectral ripple task is related to spectral resolution when using current clinical processors (e.g., Won *et al.*, 2011; Jones *et al.*, 2013), when novel techniques are used that increase the perceptual independence of adjacent electrodes, such as current focusing, there is a substantial risk that the results could be contaminated by these confounds. Given that it is not possible to determine *a priori* the extent to which each novel

^{a)}Also at: University of Southern California, Neurobiology Department, Los Angeles, California 90089.

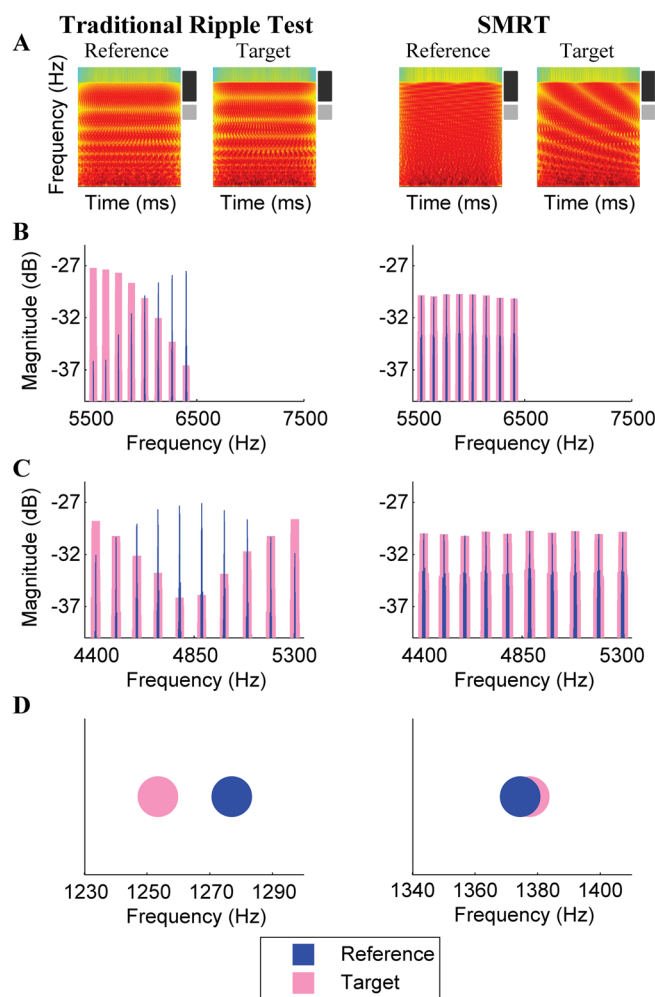


Fig. 1. (Color online) Comparison of the traditional ripple stimuli and the stimuli for the SMRT. Both the traditional and SMRT stimuli were generated using Eq. (1). (A) Spectrograms of reference and target stimuli for the traditional ripple task and the SMRT. (B) Spectrums of the high frequency edge, indicated by the dark rectangles next to the spectrograms in (A), for the two tasks. These spectrums indicate that, unlike with the traditional ripple task, there is no shift in the high frequency edge between the reference and target stimuli for the SMRT. (C) Spectrum of a portion of the frequency region, indicated by the light rectangles next to the spectrograms in (A). These spectrums indicate that local loudness cues are present in the traditional ripple task but absent in the SMRT. (D) Spectral centroid for the reference and target stimuli. These plots illustrate how the centroids for the target and masker differ in the traditional ripple task versus in the SMRT.

technique increases the impact of these confounds, it is necessary to either re-validate the traditional spectral ripple task with each new technique using a method similar to Jones *et al.* (2013) or to create a modified spectral ripple task that eliminates those potential confounds from the stimuli. Although it may very well be true that traditional spectral ripple tests measure spectral resolution in all situations, to avoid re-validating the test for use with each new technique, we propose a small modification to the standard ripple stimuli that eliminates this potential issue.

This paper describes a modified version of the spectral ripple test, referred to as the spectral-temporally modulated ripple test (SMRT), that was designed to eliminate the potential confounds in the traditional spectral ripple stimuli by using a spectral ripple with a modulation phase that drifts with time [see Fig. 1(a)]. As a result, all

frequency regions receive all loudness levels over the duration of the stimulus, thus avoiding local loudness cues and edge effects. Similarly, since the spectral centroid constantly shifts throughout both the target and the reference stimuli, this avoids confounding spectral centroid cues. The right half of Fig. 1 illustrates that the potential confounds in the standard ripple stimuli are absent in the SMRT stimuli. In this paper we present the details of the SMRT and an experiment to demonstrate that the results from this test are related to spectral resolution.

2. Spectral-temporally modulated ripple test

The SMRT consists of an adaptive procedure whereby the ripple density of the target stimulus is modified until the listener cannot distinguish between the reference and the target stimuli. The stimuli and procedures are described in detail below.

2.1 Stimuli

Each stimulus for the SMRT is 500 ms with 100 ms onset and offset linear ramps and is generated with a 44.1 kHz sampling rate. The stimuli are generated using a non-harmonic tone complex with 202 equal amplitude pure-tone frequency components, spaced every 1/33.333 of an octave from 100 to 6400 Hz. The amplitudes of the pure tones are modulated by a sine wave based on the following equation:

$$S(t) = \sum_{i=1}^{202} P(i) \times \left(\left| D \times \sin \left[\frac{i \times \text{RD} \times \pi}{33.333} + (\text{RR} \times \pi \times t) + \varphi \right] \right| + D \right), \quad (1)$$

where S is the SMRT stimulus, P is the amplitude of the pure tone with index i (100 Hz for $i = 1$, 102.1 Hz for $i = 2$, etc.), t is time, RD is the ripple density defined by the number of ripples per octave (RPO), φ determines the phase of the ripple at the onset of the stimulus, RR is the ripple repetition rate, indicating the number of times the ripple pattern repeats each second, and D scales the modulation depth of each ripple. The orthogonal effects of manipulating RD and RR are shown in Fig. 2. For the SMRT, only RD and φ are varied across stimuli to test spectral resolution, but it is possible to use this equation to create stimuli that test temporal resolution by varying RR instead of RD.

2.2 Procedure

SMRT consists of a three-interval, forced choice task. Two of the intervals contain a reference stimulus with 20 RPO. The target stimulus initially has 0.5 RPO and is modified using a 1-up/1-down adaptive procedure with a step size of 0.2 RPO. φ is randomly selected separately for each target and reference stimulus from one of four values: 0, $\pi/2$, π , and $3\pi/2$. The test is completed after ten reversals. Thresholds are calculated based on the average of the last six reversals. D is set to 20 and RR is set to 5 Hz. The stimuli are presented at 65 dB(A) from a speaker located in front of the listener at ear level at a distance of 1 m from the head. Software to conduct the SMRT is available free of charge at <http://smrt.tigerspeech.com>.

3. Verifying that the SMRT is sensitive to changes in spectral resolution

A key characteristic of any test of spectral resolution is that it is sensitive to changes in spectral resolution. To verify that the SMRT is sensitive to changes in spectral resolution, normal hearing (NH) participants were tested with the SMRT while the number of available spectral channels was systematically manipulated using a vocoder.

3.1 Methods

Eight naive NH listeners with pure tone thresholds ≤ 25 dB hearing level from 0.25 to 8 kHz participated in this experiment. The stimuli consisted of those described in Sec. 2.1 as well as vocoded versions of the stimuli. The stimuli were vocoded by first high-pass filtering the stimuli at 1200 Hz with a 6 dB per octave roll-off to add

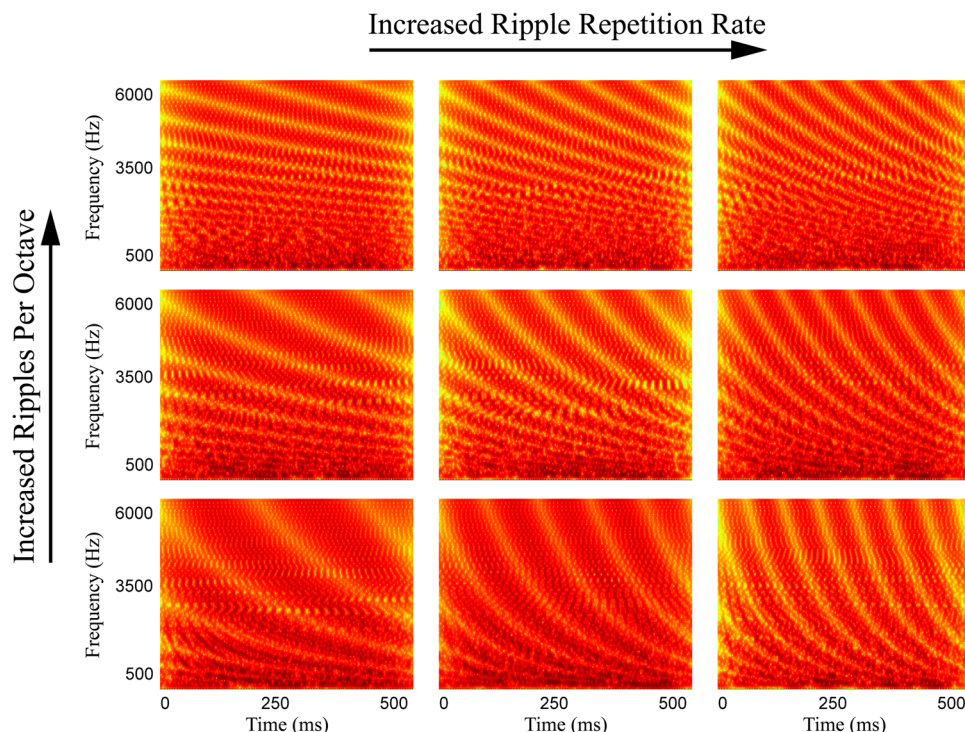


Fig. 2. (Color online) Spectrograms of stimuli generated using Eq. (1) showing the orthogonal effects of manipulating the ripple repetition rate and the number of ripples per octave. The first column represents the ripple repetition rate used for the SMRT.

pre-emphasis. Next, one, four, eight, or 16 bandpass filters were used for each ear to create different numbers of spectral channels. In all cases, the bandpass filters covered the same frequency range of 200 Hz to 7 kHz and used fourth order Butterworth filters with forward filtering. These filters were designed to sample frequency ranges that were equally spaced along the cochlea based on the equation by Greenwood (1990). The envelope of each band was extracted by half-wave rectification followed by low pass filtering at 160 Hz using a fourth order Butterworth filter. The envelopes for each channel were then convolved with white noise and the white noise was filtered using the same fourth order Butterworth filter used to sample each spectral channel. Finally, the output of all channels was summed. One potential confound with this technique is that increasing the number of channels while covering the same frequency range results in narrower spectral channels, and thus less spectral smearing. As such, it is possible that improved performance with increasing numbers of channels would simply reflect the reduced spectral smearing within each channel rather than the number of available spectral channels. To rule out that possibility, an additional vocoded simulation was created for use by a subset of participants whereby sixteen bandpass filters were used but only the output from the eighth bandpass filter was preserved, creating a narrow-band one channel vocoded condition. To distinguish the 2 one channel vocoded conditions, this condition will be referred to as the narrow one channel condition and the stimuli which used one bandpass filter that covered the entire frequency range (200 Hz to 7 kHz) will be referred to as the wide one channel condition.

Testing procedures followed those described in Sec. 2.2. Participants were tested in blocks consisting of tests for the wide one channel, four channel, eight channel, 16 channel, and unprocessed (not vocoded) conditions, presented in a random

order. Participants completed five blocks. A subset of four participants also completed three blocks containing the narrow one channel condition.

3.2 Results and discussion

Robust statistical techniques and measures were adopted to minimize the potential effects of any outliers or non-normality in the data [for more detail, see the Appendix in Aronoff *et al.*, (2011)]. These techniques included trimmed means, which are a cross between a mean and a median, and bootstrap analyses, which avoid the assumption of normality by conducting tests on distributions based on the original data set rather than on normal distributions that may not accurately reflect the data. Performance across the five conditions completed by all participants (wide one channel, four channel, eight channel, 16 channel, and unprocessed) was compared using a percentile-t bootstrap repeated measures analysis of variance with 20% trimmed means. The results indicated that there was a significant effect of condition ($p < 0.01$; see Fig. 3). Percentile-t bootstrap pairwise comparisons were conducted to determine whether increasing the number of vocoded channels yielded higher (better) SMRT thresholds. Familywise type I error was controlled using Rom's method (Rom, 1990). The results indicated that, in all cases, increasing the number of channels yielded a significant improvement in SMRT thresholds ($p < 0.0001$ for all comparisons; see Fig. 3). To determine the relationship between the number of vocoded channels and thresholds, the thresholds for the four, eight, and 16 channel conditions were analyzed with a mixed effect regression. Thresholds for the 1 channel condition were excluded to minimize contamination by floor effects. The fit was significantly better than chance ($p < 0.05$). The slope of the regression line indicated that thresholds improved by 0.25 ripples per octave for each additional spectral channel.

To determine if the improved performance with the 16 channel condition simply reflected reduced spectral smearing resulting from narrower spectral channels, performance for the narrow one channel condition was compared to that for the 16 channel condition. Both conditions contain minimal spectral smearing by using narrow bandpass filters. Because of the small number of data points (only four participants were tested with the narrow one channel condition), traditional paired t-tests were used instead of bootstrap paired comparisons, which require larger data sets. Rom's correction was used to control for familywise type I error. One subject was unable to distinguish the 20 RPO reference stimuli from a 0.1 RPO stimulus (the lowest ripple density used) for two trials

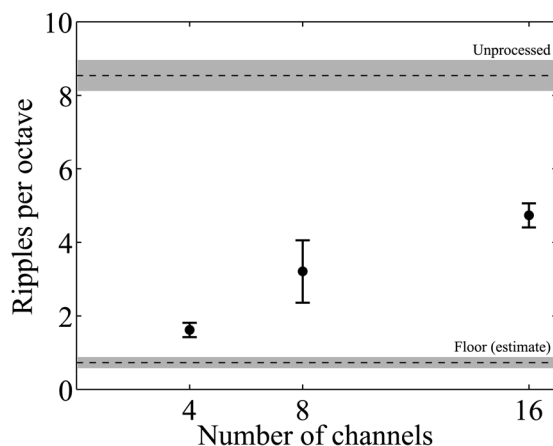


Fig. 3. Results indicating that SMRT thresholds are sensitive to changes in spectral resolution. Circles and dashed lines indicate trimmed means. Error bars and gray areas indicate ± 1 Winsorized standard error. Floor reflects performance with the wide 1 channel vocoder, at which point performance is based on temporal rather than spectral resolution. Unprocessed data indicates the ceiling for the vocoded version of the SMRT.

for the narrow one channel condition and their threshold from the remaining trial was used for this analysis. For the other subjects, the average of all trials was used. The results suggested that the narrower bandpass filters used in the 16 channel condition could not fully explain the performance in that condition. Mean threshold for the narrow one channel condition (1.7 RPO) was significantly worse than that for the 16 channel condition (5.5 RPO for the subset of four listeners; $p < 0.05$), indicating that the improved performance with increasing numbers of spectral channels does not simply reflect decreased spectral smearing.

4. Conclusions

Spectral ripple tests have a number of strengths including being able to be used for quick acute testing of normal hearing and hearing impaired listeners. However, the static nature of the ripples also creates a number of potential confounds making it difficult to interpret the results. SMRT eliminates these confounds by using a dynamically changing ripple. The results from the validation study indicate that SMRT is sensitive to changes in spectral resolution.

Acknowledgments

We thank our participants for their time and effort. We also thank Akiko Amano and Mark Robert for providing engineering expertise and Leo Litvak for providing methodological suggestions for the perceptual experiment. This work was supported by a grant from the National Organization for Hearing Research, and NIDCD Grants Nos. T32DC009975, R01-DC12152, R01-DC-001526, R01-DC004993, and R03-DC-010064.

References and links

- 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.
- Aronoff, J. M., Freed, D. J., Fisher, L., Pal, I., and Soli, S. D. (2011). "The effect of different cochlear implant microphones on acoustic hearing individuals' binaural benefits for speech perception in noise," *Ear Hear.* **32**, 468–484.
- Azadpour, M., and McKay, C. M. (2012). "A psychophysical method for measuring spatial resolution in cochlear implants," *J. Assoc. Res. Otolaryngol.* **13**, 145–157.
- 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.
- Greenwood, D. D. (1990). "A cochlear frequency-position function for several species—29 years later," *J. Acoust. Soc. Am.* **87**, 2592–2605.
- Henry, B. A., and Turner, C. W. (2003). "The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners," *J. Acoust. Soc. Am.* **113**, 2861–2873.
- Jones, G. L., Ho Won, J., Drennan, W. R., and Rubinstein, J. T. (2013). "Relationship between channel interaction and spectral-ripple discrimination in cochlear implant users," *J. Acoust. Soc. Am.* **133**, 425–433.
- 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.
- Moore, B. C. (1996). "Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids," *Ear Hear.* **17**, 133–161.
- Rom, D. M. (1990). "A sequentially rejective test procedure based on a modified Bonferroni inequality," *Biometrika* **77**, 663–666.
- 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.
- Won, J. H., Jones, G. L., Drennan, W. R., Jameyson, E. M., and Rubinstein, J. T. (2011). "Evidence of across-channel processing for spectral-ripple discrimination in cochlear implant listeners," *J. Acoust. Soc. Am.* **130**, 2088–2097.