Auditory filters and the benefit measured from spectral enhancement

Jeffrey J. DiGiovanni[a] and Padmaja Nair
Auditory Psychophysics and Signal Processing Laboratory, Ohio University, Athens, Ohio 45701

(Received 20 December 2005; revised 15 June 2006; accepted 16 June 2006)

Algorithms designed to improve speech intelligibility for those with sensorineural hearing loss (SNHL) by enhancing peaks in a spectrum have had limited success. Since testing of such algorithms cannot separate the theory of the design from the implementation itself, the contribution of each of these potentially limiting factors is not clear. Therefore, psychophysical paradigms were used to test subjects with either normal hearing or SNHL in detection tasks using well controlled stimuli to predict and assess the limits in performance gain from a spectrally enhancing algorithm. A group of normal-hearing (NH) and hearing-impaired (HI) subjects listened in two experiments: auditory filter measurements and detection of incremented harmonics in a harmonic spectrum. The results show that NH and HI subjects have an improved ability to detect incremented harmonics when there are spectral decrements surrounding the increment. Various decrement widths and depths were compared against subjects' equivalent rectangular bandwidths (ERBs). NH subjects effectively used the available energy cue in their auditory filters. Some HI subjects, while showing significant improvements, underutilized the energy reduction in their auditory filters.
 [DOI: 10.1121/1.2221554]

PACS number(s): 43.66.Ts, 43.66.Fe [DOS] Pages: 1529–1538

I. INTRODUCTION

Speech intelligibility has been identified as one of the foremost problems for hearing-aid users. Although hearing aids provide amplification to compensate for the degree of hearing loss by normalizing loudness and improving thresholds, they do not sufficiently ameliorate the distortion introduced by the damaged cochlea that leads to poor spectral resolution abilities in individuals with sensorineural hearing loss (SNHL) (Plomp, 1978). One of the methods proposed to resolve this issue is spectral enhancement. It refers to the process of selectively amplifying spectral peaks in a speech signal while the spectral valleys or troughs are attenuated or remain unaffected. The goal of this strategy is to effectively improve the signal-noise ratio, leading to better intelligibility.

Numerous researchers have been studying the efficacy of spectral enhancement strategies producing contradictory results (Simpson *et al.*, 1990; Baer *et al.*, 1993; Giguère and Smoorenburg, 1998; Franck *et al.*, 1999; Miller *et al.*, 1999; Lyzenga *et al.*, 2002; DiGiovanni *et al.*, 2005). Miller *et al.* (1999) conducted a physiologic study using a spectral enhancement algorithm called contrast-enhanced frequency shaping (CEFS). They presented both unmodified vowels and vowels processed using CEFS to acoustically traumatized cats and studied their auditory neural responses. Results indicated that phase-locked neural representation of vowel formants was substantially improved with the CEFS algorithm. Giguère and Smoorenburg (1998) used cochlear models to assess spectral enhancement. By simulating the excitation pattern for the vowel /æ/for a normal ear and an ear with a 50% outer hair cell loss, they found no difference in the excitation patterns for the damaged cochlea with and without spectral enhancement. This led them to conclude that spectral enhancement is not a viable solution to the poor spectral resolution caused by broadened auditory filters of the damaged cochlea.

Again, speech perception tests using spectral enhancement algorithms have yielded mixed results. Simpson *et al.* (1990) tested subjects with sensorineural hearing loss on speech-in-noise tests and found that they showed improved results in word and sentence identification tasks. They found an increase of 6.4% in identification of consonant-vowel-consonant words, and an 11.4% increase for Bench-Kowal-Bamford (BKB) sentences using spectral enhancement. Baer *et al.* (1993) changed certain parameters of the same algorithm and carried out further testing; their results however did not show the significant amount of improvement shown by Simpson *et al.* (1990). The only condition in which they were able to demonstrate a significant benefit occurred when spectral enhancement was combined with wide dynamic range compression. In this condition, an 0.8-dB improvement in SNR was noted. A similar study by Franck *et al.* (1999) found that combining spectral enhancement with dynamic range compression reduced the benefit of the spectral enhancement scheme. Although they found an improvement in vowel perception scores in the "spectral enhancement alone" and "spectral enhancement along with single channel compression" conditions, the final consonant perception in noise scores worsened considerably. Additionally, the best speech intelligibility scores were found in the unprocessed condition (Franck *et al.*, 1999). More recently, Lyzenga *et al.* (2002) combined a spectral expansion algorithm with a different algorithm that worked to decrease the effects of upward spread of excitation and observed a 1-dB improvement

[a]Electronic mail: digiovan@ohio.edu

in speech reception threshold (SRT). The range of results for these different studies includes a decrease in intelligibility to a modest improvement.

The incongruity of the results of these studies indicates that the theory of spectral enhancement is inherently flawed or has been inadequately implemented. In a previous study, DiGiovanni *et al.* (2005) attempted to assess the theoretical viability of spectral enhancement. Both normal-hearing and hearing-impaired subjects were required to detect and discriminate a narrow-band peak (the approximate width of a vowel formant) in broadband noise. Frequencies flanking the target narrow-band signal were decremented to improve the spectral contrast and consequently the signal-to-noise ratio. The spectral decrements surrounding the target stimuli were either 100 or 200 Hz wide, and their depth was either 3 or 6 dB (i.e., their intensity was reduced by 3 or 6 dB compared to all other bands including the target band). The test conditions ranged from no decrements to a maximum decrement 6 dB deep and 200 Hz wide. Results showed that spectral decrements significantly improved detection and discrimination of spectral peaks in noise for both normal-hearing and hearing-impaired individuals, leading the authors to the conclusion that spectral enhancement is a viable premise. However, as the stimulus used was broadband noise, conclusions regarding the effect of spectral enhancement on speech were limited. The authors further noted that the amount of benefit shown by hearing-impaired subjects, although significant, was less than the benefit demonstrated by normal-hearing subjects. They posited that the broadened auditory filters of the damaged cochlea in individuals with sensorineural hearing loss limited the amount of measured improvement, though auditory filters were not measured in that study.

The goal of the present study was to examine the effects of increasing spectral contrast on detectability of raised harmonics in a broadband harmonic spectrum. A second goal was to explore the relationship of auditory filter bandwidths and the improvements in increment detection when spectral enhancements are added. In the first experiment, auditory filters were measured for all subjects. In the second experiment, both normal-hearing and hearing-impaired subjects were required to detect a narrow-band peak (approximately the width of a vowel formant) in a harmonic spectrum with a fundamental frequency of 50 Hz and harmonics up to and including 5000 Hz. Spectral decrements of varying width and intensity relative to the overall spectrum level were inserted into the bands flanking the peak.

II. EXPERIMENT 1: AUDITORY FILTERS WIDTHS

Individuals with sensorineural hearing loss experience greater difficulty in speech discrimination tasks and this has, in part, been attributed to their wider auditory filters. DiGiovanni *et al.* (2005) found that, although hearing-impaired subjects showed a significant improvement with spectral enhancement, they demonstrated less benefit as compared to normal-hearing subjects. In the present study, the auditory filter bandwidth of hearing-impaired subjects was measured at 2 kHz, since this frequency lies well inside the region where second formants mostly occur. Next, listeners' benefit of spectral enhancement was related to their auditory filter widths to assess the hypothesis that wider auditory filters lead to poorer benefit from spectral enhancement. Quantitative relationships between a subject's improvement and their auditory filter bandwidth will be made.

TABLE I. Audiometric thresholds and ages of subjects in the HI group.

		Audiometric thresholds (kHz)							
Subject	Age (years)	0.25	0.5	1.0	1.5	2.0	3.0	4.0	8.0
I1	66	20	25	40	45	50	45	50	65
I2	75	15	20	35	⋯	40	⋯	50	60
I3	78	55	45	45	⋯	40	⋯	55	80
I4	67	55	60	55	45	40	35	35	65

A. Methods

1. Subjects

Two groups of subjects participated in this study: four individuals with normal hearing (NH group) and four individuals with sensorineural hearing loss (HI group). Informed consent was obtained from all subjects and they were paid for participation.

The NH group had audiometric thresholds of 15 dB HL or better from 0.25 to 8 kHz and a negative report of auditory pathology. The HI group had mild to severe sensorineural hearing loss, with thresholds ranging from 35 to 50 dB HL from 1.5 to 3.0 kHz. Audiograms for this second group were measured at octave intervals, and interoctave intervals when appropriate, from 0.25 to 8.0 kHz. Audiometric data for the HI group are shown in Table I. Also, to be included in the experiment, participants were required to have a relatively constant threshold microstructure around the test frequency, 2.0 kHz. For frequencies between 1.5 and 2.5 kHz, thresholds were measured at 25-Hz intervals to ensure that hearing sensitivity did not vary more than 10 dB in the region around 2.0 kHz. Thresholds measured between 1.5 and 2.5 kHz were obtained using a computer-implemented version of Békésy discrete-frequency audiometry. Participants pressed a button on a computer keyboard to signal "louder" or "softer," which raised or lowered the signal level in 1.5-dB steps. Participants were instructed to press the louder button if the signal was inaudible and the softer button if it was audible. The stimulus for the Békésy audiometry task was a 250-ms tone (with 20-ms raised cosine ramps). Thresholds for each frequency were based on the mean of the last 10 of 12 reversals in stimulus level direction.

2. Stimuli

The stimuli were generated digitally at a sampling rate of 24.414 kHz using a computer (Dell Dimension, DIM 4550). The computer was equipped with a signal processor (Tucker-Davis Technologies, RP2.1) 24-bit digital-to-analog converter and, after filtering and attenuation (Tucker-Davis Technologies, PA-5), sounds were presented to one ear of a listener through an earphone (Telephonics, TDH-39P). Bands of noise and 2-kHz pure tones were generated using the aforementioned hardware. The stimuli were 500 ms in duration with a 500-ms interstimulus interval. The pure tone had a raised cosine ramp of 10 ms. The noise was low-pass filtered at 8.0 kHz and was generated using a Gaussian distribution. The noise was fixed in level at 46 dB/Hz, which is equivalent to an overall level of 85 dB SPL. The pure tone varied in level depending on the subject's response.

3. Procedure

A three-interval forced-choice procedure was used to measure the subjects' auditory filter bandwidths. A two-up one-down rule was used to track 70.7% on the psychometric function (Levitt, 1971). Each block consisted of ten reversals and the threshold was calculated from the final eight reversals. The step size was 4 dB for the first two reversals and decreased to 2 dB for the remainder of the block. In any given trial, two of the intervals contained notched noise (standard intervals) while one of the intervals contained both the noise and a 2-kHz tone centered in the spectral gap of the noise (target interval). f_{lo} and f_{hi} defined the cutoff frequencies for the high-pass and low-pass conditions, respectively. The subjects' task was to determine the interval containing the tone. Correct answer feedback was given following each response for a given trial. The subject was seated in a sound attenuating booth. Each subject underwent a minimum of two hours of practice before data collection began. Three blocks were collected for a given point, or notched-noise condition, for a total of 15 blocks per subject. The three cutoff frequencies that were collected were derived from g values of 0.0, 0.2, and 0.4, resulting in five points per filter (0,0; 0.2, 0.2; 0.4, 0.4; 0.2, 0.4; 0.4, 0.2) (Glasberg and Moore, 1990) where

$$g = \frac{|f - f_0|}{f_0} = \frac{\Delta f}{f}. \qquad (1)$$

For instance, for $g=0.20$ and $f_0=2.0$ kHz, $f_{lo}=1.6$ kHz and $f_{hi}=2.4$ kHz.

B. Results

The auditory filters were modeled from the collected data using Patterson *et al.*'s (1982) formula for curve fitting [Roex(p,r)]:

$$W(g) = (1-r)(1-pg)e^{-pg} + r, \qquad (2)$$

where p determines the passband width and the rate of fall of the skirts of the filter and r is the dynamic range limiter.

Auditory filter bandwidths were derived by fitting the five points of data into the two-parameter Roex(p,r) model. Table II shows the equivalent rectangular bandwidths (ERBs) of the modeled data along with the subject age and group. The average ERBs were 276 and 473 Hz for the NH and HI subjects, respectively.

C. Discussion

Auditory filter bandwidths measured in this experiment largely agree with those established in literature. Dubno and Dirks (1989) measured filters in nine normal-hearing subjects at a spectrum level of 40 dB/Hz. At 2.0 kHz, their subjects had ERBs ranging from 220 to 320 Hz. The range for our NH subjects was 222–300 Hz. They also measured auditory filters in hearing-impaired subjects. Their hearing-impaired subjects had ERBs ranging from 400 to 1640 Hz, whereas the ERBs for the HI subjects in this study ranged from 356 to 648 Hz. Glasberg and Moore (1986) also measured auditory filter bandwidths in normal and hearing-impaired ears. Their ranges were in agreement with the current data as well as with that of Dubno and Dirks' (1989).

The difference between the performance of the NH and HI groups can be attributed to a loss in frequency resolution. Broadened auditory filters in hearing-impaired subjects is a common finding (Glasberg and Moore, 1986; Dubno and Dirks, 1989). The current data also reveal that the HI subjects performed worse than the NH subjects. Overall, our HI subjects' had an average ERB 1.7 times that of the NH subjects' ERBs.

TABLE II. ERBs and ages for HI and NH subjects.

Subject	Age (years)	ERB (Hz)
I1	66	647.6
I2	75	449.1
I3	78	355.7
I4	67	440.6
HI Average		473.3
N1	21	222.4
N2	21	300.2
N3	24	296.7
N4	24	283.0
NH average		275.5

III. EXPERIMENT 2: INCREMENT DETECTION THRESHOLDS

DiGiovanni *et al.* (2005) demonstrated that spectral enhancement showed promise as a viable method to improve speech intelligibility in hearing-impaired individuals. The stimulus used in their study was broadband noise with a 100-Hz-wide narrow-band peak at 2 kHz. Spectral decrements were inserted adjacent to this peak in order to provide spectral contrast and subsequently improve the signal-to-noise ratio. Although the study demonstrated significant improvement for both normal-hearing and hearing-impaired subjects, it did not indicate whether the results could be generalized to speech stimuli as well. In this study, subjects listened in detection tasks using a broadband harmonic spectrum. Test conditions included various depths and widths of spectral decrements to assess their relationship with improved performance on spectral enhancement tasks, if any. Furthermore, a second set of conditions was tested with stimuli that differed in one respect; a three-harmonic, fixed 10-dB increment was added to the standard and target intervals. These conditions were included to test the possibility that by introducing a first formantlike fixed increment, upward spread of masking may reduce the benefit of enhancing the spectrum.

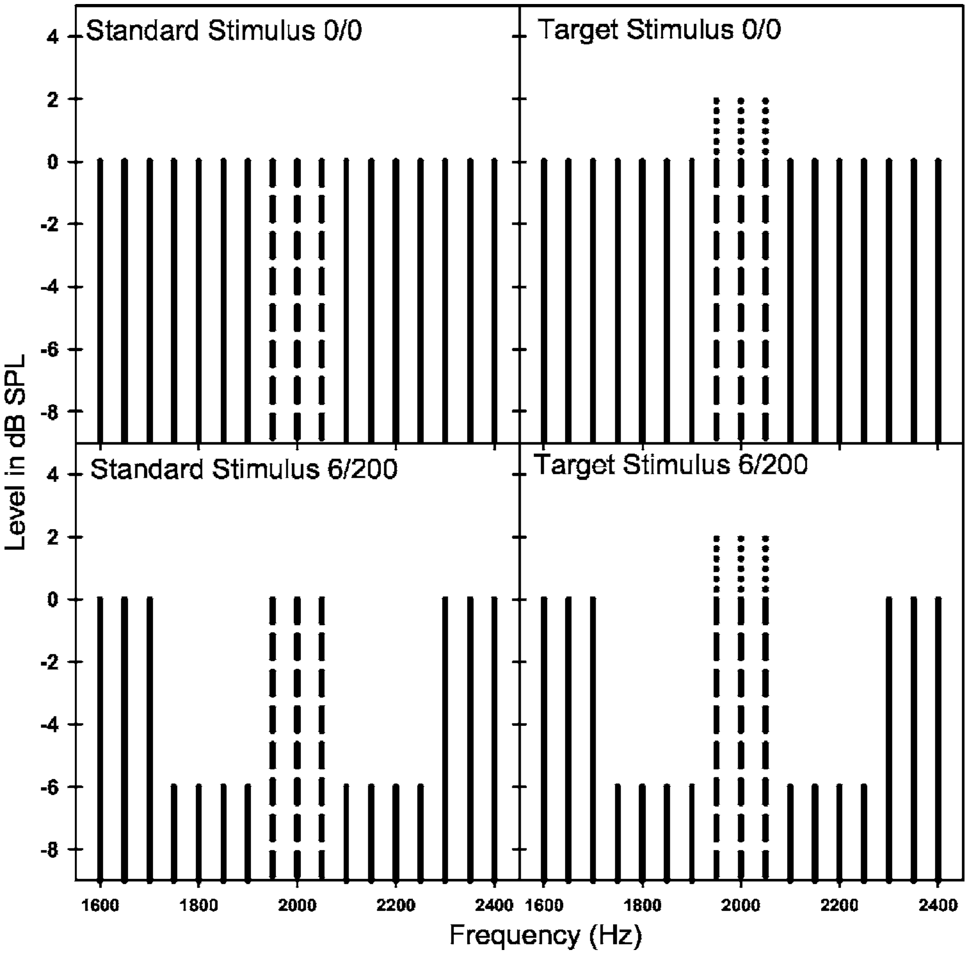

FIG. 1. Schematic representations of the stimuli used in experiment 2. Shown in the two left panels are the standard stimuli for the 0/0 (upper panel) and the 6/200 (lower panel) conditions. Shown in the two right panels are the target stimuli for the 0/0 (upper panel) and 6/200 (lower panel) conditions. In this example the target stimuli are shown with an increment of 2 dB relative to the spectrum level of the harmonic spectrum.

A. Methods

1. Subjects

Subjects used were the same as in experiment 1.

2. Stimuli

Two sets of conditions were run, one with the standard stimulus and another with a stimulus that had a first formant-like increment in its frequency spectrum, called the F1 stimulus. While the standard stimulus for this experiment was a harmonic spectrum with $F_0=50$ Hz (band limited at 5000 Hz), the F1 stimulus had an additional 100-Hz-wide, 10-dB increment inserted at the first formant region (i.e., 950, 1000, and 1050 Hz). The stimuli were 500 ms in duration with a 500-ms interstimulus interval. The target stimulus for both sets of conditions had a 100-Hz-wide (three harmonics), narrow-band increment centered at 2 kHz. Spectral decrements of varying widths were inserted adjacent to the 2 -kHz increment. The convention used to refer to these conditions will be the depth (dB)/width (Hz). For instance, a 6 -dB-deep, 200-Hz-wide decrement will be referred to as 6/200. The control condition with no decrements will be referred to as 0/0. F1 stimuli are denoted similarly, but with the addition of F1 in the designation (e.g., 6/200F1). Figure 1 shows a schematic of part of the spectrum for the 0/0 and 6/200 conditions for both the standard and the target stimuli.

3. Procedure

A three-alternative forced-choice task was used to measure the subjects' ability to detect the narrow-band increment in a harmonic series. A two-down one-up paradigm tracking 70.7% on the psychometric function was used (Levitt, 1971). The step size was 3 dB for the first two reversals and 1 dB for the remaining reversals. A block of trials ended after the completion of ten response reversals. The first two reversals were discarded and threshold was the mean of the final eight reversals. The standard harmonic spectrum was presented in any two of three intervals in a given trial, and in the other interval, the harmonic spectrum with the 2-kHz increment was presented. The interval order was determined from a uniform, random distribution. The subjects were instructed to pick the interval that was different from the other two and respond using the computer keyboard. Feedback was given on a trial-by-trial basis.

Several conditions were collected: 0/0, 6/200, 6/400, 9/200, 9/400, 6/400F1, and 9/400F1 by varying the width and depth of the spectral decrements surrounding the 2 -kHz increment. The order of conditions was randomized within subjects. A minimum of three blocks per condition were measured to calculate thresholds. More blocks were run if the standard deviation exceeded 3 dB until the standard deviation was below 3 dB. The number of blocks never exceeded five.

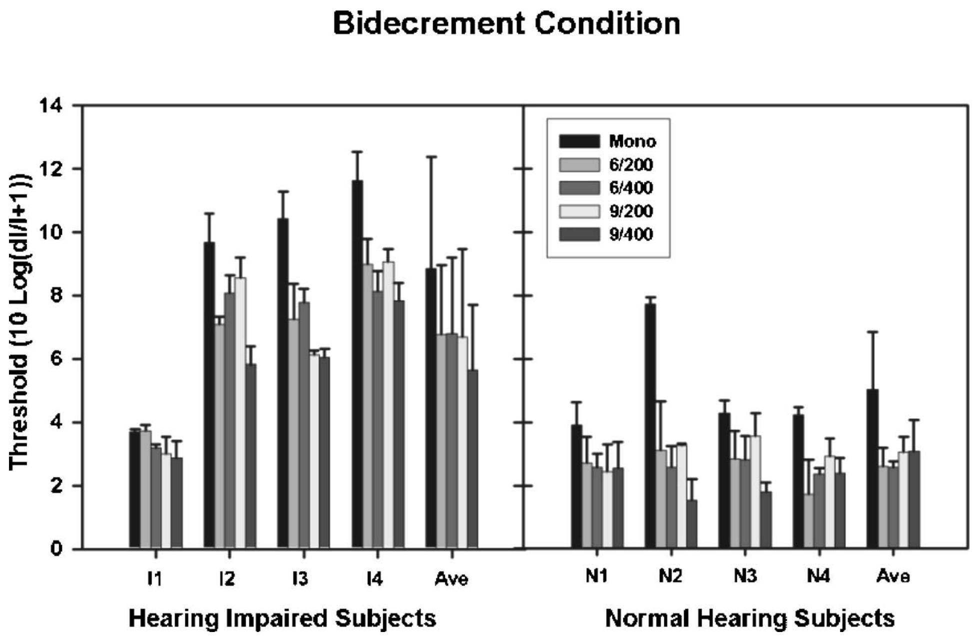

FIG. 2. The increment detection results of the NH and HI groups for the bidecrement conditions in experiment 2 are shown. The data are presented after they were converted into ΔL (i.e., the difference between the centerband spectrum level of the signal and standard intervals). Data for the HI subjects are shown in the left panel and data for the NH subjects are shown in the right panel. The increment thresholds in dB for each subject as well as the group average are shown with corresponding standard deviation bars. The 0/0 condition is referred to as Mono.

B. Results

Increment detection thresholds for both subject groups are shown in Figs. 2 and 3. Narrow-band target thresholds were converted from $10 \log(\Delta I/I)$ to $10 \log(\Delta I/I+1)$ in order to demonstrate the relative difference at threshold between the signal increment and the harmonic spectrum. The spectral enhancement conditions are thus expressed in terms of decibel increments above the spectrum level of the harmonic series. The spectral enhancement conditions ranged from no enhancement (0/0) to a maximum enhancement in the 9/400 condition. Additionally, F1 conditions (9/400F1 and 6/400F1), which included a spectral increment of 10 dB at 950, 1000, and 1050 Hz, were tested. Benefit as it is understood in this study is the difference between threshold in a given enhancement condition relative to the 0/0 condition.

A repeated measures analysis of variance was performed on the entire listener pool to measure the effects of spectral enhancement in the standard and the F1 conditions. A significant group interaction was found on the standard condition $[F(1,6)=6.42,\ p<0.05]$, indicating that, overall, normal-hearing subjects had better thresholds than the hearing-impaired group. Additionally, significant effects of decrement depth $[F(2,5)=10.66,\ p<0.05]$ and width $[F(2,5)=10.51,\ p<0.05]$ were found, indicating that greater the depth and width of the decrements, the better the detection thresholds for both groups. No significant depth by group $[F(2,5)=1.45,\ p=0.318]$ or width by group $[F(2,5)=0.43,\ p=0.67]$ interactions were found, suggesting that there was no differential effect of decrement depth or width on either of the two groups. The three-way interaction (depth $\times$ width $\times$ group) was not significant $[F(4,3)=0.48,\ p=0.76]$. Overall, the analysis along with the data shown in Fig. 4 suggests that both groups showed greatest benefit in the condition with the greatest spectral decrement (9/400).

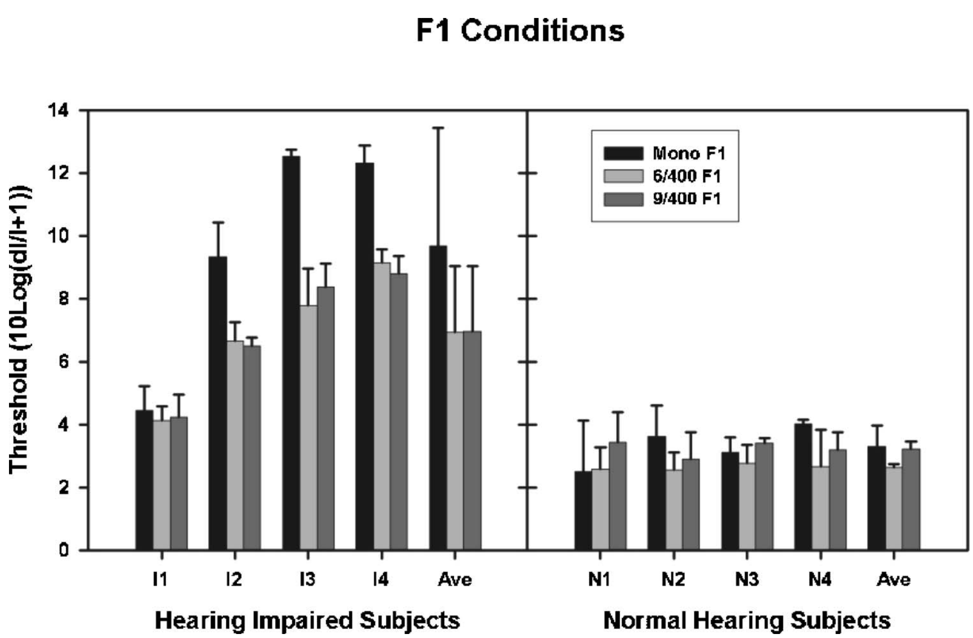

FIG. 3. As in Fig. 2, but now showing results for the F1 stimuli.

Improvement in Bidecrement Conditions

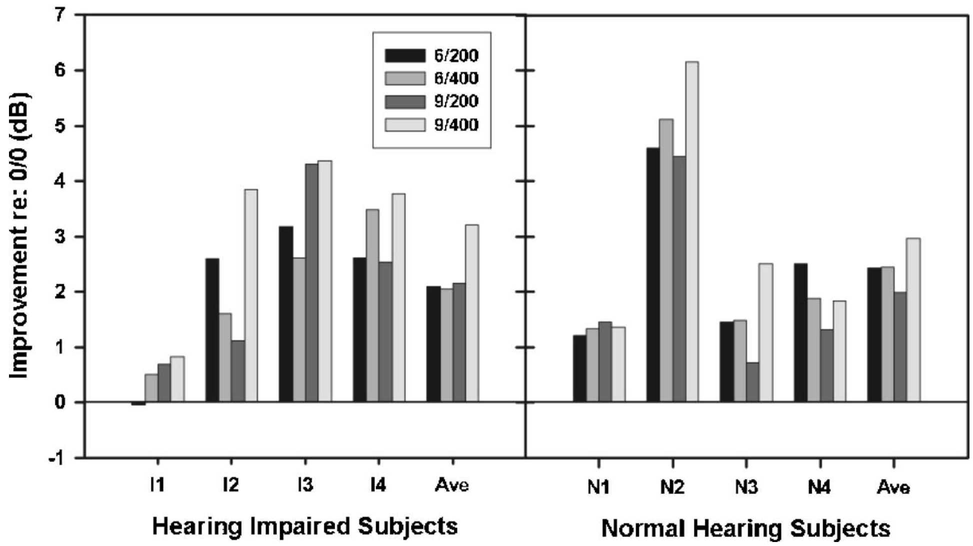

FIG. 4. The benefit measured for each decrement condition in experiment 2 is shown relative to the 0/0 condition. The comparisons were made from the ΔL converted data. The left panel shows individual benefit along with group average for the HI group and the right panel shows the same data for the NH group.

For the F1 conditions, a significant group interaction was found [$F(1,6)=14.25$, $p<0.05$], indicating that the normal-hearing subjects had better thresholds than the hearing-impaired subjects. Also, significant effects of decrement depth [$F(2,5)=13.46$, $p<0.05$] and width [$F(1,6)=9.99$, $p<0.05$] were found, indicating that wider and deeper decrements led to better detection thresholds in the F1 conditions as well. The significant depth by group interaction [$F(2,5)=6.1$, $p<0.05$] suggests that there was a differential effect of decrement depth on each group. Furthermore, a significant interaction between width and depth was found [$F(2,5)=13.46$, $p<0.05$], suggesting that the relative importance of the decrement width varied depending on the depth of the decrement. The three-way interaction (depth × width × group) was also significant [$F(2,5)=6.1$, $p<0.05$].

These analyses show that spectral enhancement demonstrated significant benefit for both normal-hearing and hearing-impaired listeners. The F1 condition, which approximates a steady-state vowel more closely than the standard stimulus, suggests that spectral enhancement could be viable for real-life situations. The greatest amount of improvement was seen for the most extreme 9/400 spectral enhancement condition, and the least improvement was seen for the 6/200 condition.

C. Discussion

The findings from experiment 2 agree with the findings from DiGiovanni *et al.* (2005). In their study, one of the experiments was similar to the standard stimulus in this study, except that they used noise rather than a harmonic complex. Both studies showed that the HI subjects performed worse than the NH. While they suggested that there may be a link between the measured improvement and auditory filter bandwidths, auditory filters were not measured and therefore this relationship could not be directly evaluated. From Figs. 4 and 5 it can be seen that the current HI subjects showed similar improvement to the NH subjects for the standard conditions and greater improvement for the F1 conditions, whereas the NH group in DiGiovanni *et al.* (2005) showed more improvement than the HI subjects. Overall, the findings of their study and the current study are in agreement. It is clear that both the NH and HI subjects in this study demonstrated benefit in detection of a set of harmonics when the spectrum was enhanced.

While both the NH and HI subjects generally demonstrated benefit in the F1 conditions, the NH subjects did not show nearly as much benefit as the HI subjects, as shown in Fig. 5. On average, the HI subjects realized a 2.72-dB benefit in the F1 conditions while the NH subjects only showed a 0.38-dB improvement. This finding is contrary to our prediction. We predicted that subject performance, and consequently their improvement, would be related to the analysis band that was used, namely their auditory filter bandwidth in equivalent rectangular bandwidth (ERB) units. Our NH subjects had consistently narrower auditory filters than our HI subjects, which led to our prediction that NH subjects would perform better on this task as well as show greater improvement.

Despite an average improvement of almost 3 dB for both the F1 conditions, the amount of improvement varied among the HI subjects. It is difficult to predict if this variability will carry over to a spectral-enhancement algorithm used in a real-world environment. Another analysis was performed to better understand the effect of F1 on performance. The original analysis compared the enhanced conditions to the 0/0 condition separately for experiments 1 and 2. However, a comparison to the 0/0 data is not required. To understand the impact of F1, the three conditions (0/0, 6/400, and 9/400) can be directly compared between the standard and F1 data from experiment 2. To do this, the 0/0 threshold is subtracted from the 0/0F1 threshold. This is repeated for the 6/400(F1) and the 9/400(F1) conditions. When this analysis is done, most of the findings are as expected. That is, the 6/400 and 9/400 conditions have the same or worse thresholds for both groups. This is also true for the 0/0 condition for the HI subjects. However, the 0/0 threshold for the NH

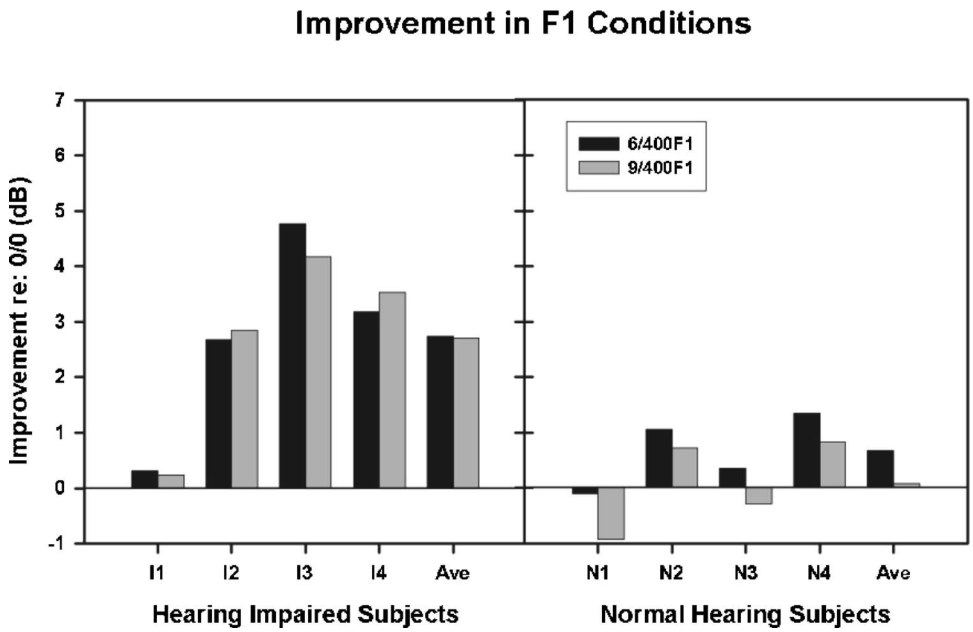

FIG. 5. As in Fig. 4, but now showing results for the F1 stimuli.

group was better for all subjects in the presence of F1. The source of this finding is uncertain for two reasons: (1) all subjects were trained for at least 4 h prior to data collection and (2) the presentation order was randomized. All other findings, however, were as expected. Therefore, the finding that HI subjects show more improvement in the F1 conditions should be made with caution.

IV. GENERAL DISCUSSION

The effective power within an auditory filter may be the significant factor in determining performance in experiment 2 (Bilger, 1978; Summers and Leek, 1994). Similar to Summers and Leek's (1994) ripple detection results, our HI subjects performed worse than the NH subjects. Despite this, the HI subjects showed improvement in the spectrally enhanced conditions. Summers and Leek (1994) posited that broadened auditory filters resulted in a flattened internal spectrum leading to the reduced performance. In the current study, auditory filter data show that the HI subjects have broader auditory filters than the NH subjects. This might explain the poorer overall performance of the HI group. If the overall power within the auditory filter is the determining factor for performance, then the relative signal-to-noise ratio or, rather, the increment-to-spectrum ratio would decrease with increasing bandwidth. Furthermore, this would explain the improvement measured for both groups when spectral decrements are inserted. Removing energy from the background spectrum results in an increase in the relative power of the increment to the spectrum. From this, it follows that little or no benefit would be realized for extremely wide filter bandwidths. As the filter widens, the effect of removing energy from part of the spectrum becomes less significant and, ultimately, will have a negligible impact on performance. However, since the subject groups were not age matched, differences in performance between groups may be partially a result of age differences rather than auditory-filter bandwidths.

To better understand the amount of improvement related to auditory filter bandwidths, two quantitative models were developed. Multiple cues have been well described to account for detection of increments in broadband signals (Formby *et al.*, 1994). Two of these cues are relevant to this analysis: overall-energy and spectral-profile cues. The overall-energy cue simply assumes that performance improvements can be predicted by the energy difference between the standard stimulus and target stimulus within a band. The spectral-profile cue assumes that the listener utilizes changes in the spectral profile of between the standard and target stimulus without the contribution of the increment. (Formby *et al.*, 1994; Heinz and Formby, 1999). While neither of these cues incorporate profile analysis *per se*, a profilelike analysis could have been used by the listener to detect the increment (Green, 1988). Using these cues, two models were developed: (1) the overall-energy model and (2) the spectral-profile model. These models were developed to predict the maximum possible benefit subjects could realize, depending on the possible listening cues.

The overall-energy model assumed that the main cue for detection was the comparison of overall energy in a given bandwidth (i.e., the auditory filter bandwidth) between the standard and comparison stimuli. To calculate benefit, the standard stimulus used the 0/0 condition. The overall level was calculated for 0/0 and each enhancement condition within each subject's auditory filter bandwidth. To do this, the number of harmonics within the auditory filter was calculated. Next, the overall energy for these harmonics was calculated considering the individual levels of the increment harmonics (without the pedestal), the decremented harmonics in the enhancement condition, and the fixed-level harmonics in the event the auditory filter bandwidth was wider than the decrement width plus the increment width. In this manner, as the auditory filter bandwidth exceeded the width of the increment plus decrement, less benefit would be predicted. The benefit for a particular condition was the difference in energy between these two stimuli.

The spectral-profile model assumed listeners are using the spectral differences within a band as the cue. The model

TABLE III. Comparison of benefit predicted by overall-energy model, spectral profile model, and the measured benefit for each subject.

	I1 (648 Hz)			I2 (449 Hz)			I3 (356 Hz)			I4 (441 Hz)		
	Energy (dB)	Profile (dB)	Measured (dB)	Energy (dB)	Profile (dB)	Measured (dB)	Energy (dB)	Profile (dB)	Measured (dB)	Energy (dB)	Profile (dB)	Measured (dB)
6/200	3.5	6.0	−0.04	3.2	4.4	2.59	2.9	6.0	3.17	3.5	6.0	2.62
6/400	3.5	6.0	0.51	4.2	6.0	1.60	2.9	6.0	2.62	3.5	6.0	3.48
9/200	4.9	9.0	0.69	4.4	6.3	1.11	4.0	9.0	4.30	4.9	9.0	2.54
9/400	4.9	9.0	0.82	6.0	9.0	3.85	4.0	9.0	4.37	4.9	9.0	3.77
	N1 (222 Hz)			N2 (300 Hz)			N3 (297 Hz)			N4 (283 Hz)		
6/200	1.9	6.0	1.21	2.9	6.0	4.60	1.9	6.0	1.46	1.9	6.0	2.50
6/400	1.9	6.0	1.33	2.9	6.0	5.12	1.9	6.0	1.48	1.9	6.0	1.88
9/200	2.6	9.0	1.46	4.0	9.0	4.44	2.6	9.0	0.72	2.6	9.0	1.32
9/400	2.6	9.0	1.37	4.0	9.0	6.16	2.6	9.0	2.50	2.6	9.0	1.83

assumes that listeners are not incorporating the level of the increment with the pedestal as part of the cue. Therefore, the energy within a given bandwidth was calculated by comparing the spectral differences between the standard and target stimuli of the magnitude spectrum within the auditory filter but outside the increment. Therefore, the number of harmonics within each subject's auditory filter was calculated. The three increment harmonics were subtracted. The remaining harmonics were then considered. For the standard stimulus, the overall level within the auditory filter minus the level of the increment was calculated by adding the levels of the remaining harmonics. For the target stimulus, the overall level within the auditory filter minus the level of the increment level was calculated by adding the level of the fixed-level and decremented harmonics, the number of which depended on the enhancement condition and the auditory filter bandwidth. The predicted improvement was calculated as the difference between the level of this band for the 0/0 condition and an enhancement condition.

The maximum achievable benefit was calculated using the subjects' measured auditory filter bandwidths for both models. Using the subjects' own bandwidths allows for the most direct test of the model. The model predictions along with the measured benefit are shown in Table III. Several qualitative observations can be made. First, for both models, greater depth always predicted more benefit. Increasing width, however, required an analysis band that exceeded the subjects' ERB in order to predict greater benefit. This impacted the HI subjects more due to their wider ERBs. Subject data shown in Fig. 4 indicate that, on average, greater depths and widths contributed to improved performance. Second, both models predict less benefit for broader ERBs. However, subject data did not consistently agree with this trend. HI subjects provided a greater range of ERBs, but their performance did not monotonically vary with their ERB. It is likely that a greater pool of subjects, especially HI subjects having broader ERBs, would smooth out this trend. Third, both models overpredicted the benefit. It may be that subjects do not make ideal use of the cues. The spectral-profile model consistently overpredicted the data while the overall-energy model predicted numbers within the proximity of the data.

To quantify the predictive value of each of the models, an error analysis was performed. The error analysis involved a calculation of the root-mean-square (rms) error, in dB, between the predicted and measured benefit. The results of this analysis are shown in Table IV. The overall-energy model has a modest 1.84-dB rms error for all subjects and all conditions combined. The spectral-profile model, however, has almost three times the rms error than the overall-energy model. Furthermore, the spectral-profile model overpredicts the measured benefit for every condition and every subject. The error analysis for the overall-energy model was calculated for the NH and HI group means separately to assess if the model was more accurate for one of the groups. The rms error was 1.19 and 2.32 dB for the NH and HI groups, respectively. This reveals that the overall-energy model is a better predictor of benefit for the NH listeners than the HI listeners. Therefore, subjects likely are using the overall energy in a particular analysis band as a partial or total cue, that determines their performance in these listening tasks. Moreover, the model predicts the NH performance within 1.19 dB rms error while the HI error is approximately double. It may be that the HI listeners in this study could not utilize the maximal benefit. However, it appears that the HI subjects had a greater variability than the NH subjects for this model. From Table III, subject I3 shows excellent agreement between predicted and measured improvement. I4 shows a lesser degree of agreement, and I1 and I2 showed the least agreement. Given this variability for HI subjects, it is possible that the auditory filter data may not be sufficient to describe the performance benefit for all HI subjects.

This study measured auditory filters and increment detection thresholds for NH and HI subjects to elaborate on a suggested relationship between detection thresholds and auditory filters as discussed in DiGiovanni *et al.* (2005). DiGiovanni *et al.* (2005) measured increment detection thresholds

TABLE IV. Error analysis for predicted and measured benefit.

	rms error (dB)	
	Energy	Profile
All subjects	1.8	5.3
NH group	1.2	5.5
HI group	2.3	5.2

as well as frequency discrimination for a narrow-band increment in a broadband noise. They found that subjects improved similarly in both tasks, thus both tasks were providing the same type of information. The current study limited the measurement to increment detection thresholds largely due to this fact as well as that the usage of a harmonic spectrum did not lend itself to a frequency discrimination task.

Despite the fact that at high levels an increment may have a fairly broad excitation, literature has consistently shown that the important aspect of neural coding an increment (or formant) in a broadband spectrum (e.g., a vowel spectrum) is through synchrony and not neural excitation, and even at high presentation levels that cause a broadband excitation, the frequency specificity of synchrony was maintained (Sachs and Young, 1980; Horst *et al.*, 1985). In this manner, if a spectral peak can be "reintroduced" into the signal, it follows that regardless of the bandwidth of excitation, synchrony will increase local to the place of the spectral peak in a similar way that formants are coded at higher levels. Therefore, an algorithm that can partially restore the spectral representation of speech will likely improve speech intelligibility.

V. CONCLUSIONS

Based on psychophysical, physiological, and clinical studies, it appears that modest improvements can be achieved by enhancing peaks in a spectrum by either selectively amplifying the peak or attenuating the energy adjacent to a peak. Based on the current data and those from literature (as cited), several inferences can be made: (1) deeper and wider decrements are better, at least to 400 Hz wide and 9 dB deep; (2) auditory filter bandwidths are related to improvement in enhancement conditions; (3) F1 should not be amplified significantly (Miller *et al.*, 1999); (4) amplification should be focused on F2 and F3 frequency regions (Miller *et al.*, 1999); (5) the effects of combined processing (e.g., spectral enhancement and compression) are not well understood (Franck *et al.*, 1999); and (5) since it is well known that consonant perception is a critical factor for speech intelligibility, any spectral enhancement algorithm needs to account for consonant perception.

These experiments suggest that further research and the development of an algorithm for spectral enhancement remain worthy ventures. Historical arguments suggest that broadened auditory filters eliminate the possible benefit from spectral enhancement. However, animal studies have shown that spectral enhancement is effective in restoring the formant representation at the neural level (Miller *et al.*, 1999). DiGiovanni *et al.* (2005) showed promising results by testing normal and hearing-impaired listeners in two psychophysical studies. To further these findings using stimuli that are more speechlike, two psychophysical experiments were performed to test the NH and HI benefit from spectral enhancement. These experiments included (1) auditory filter measurements and (2) narrow-band signal detection with and without adjacent spectral decrements as well as with and without an added 10-dB fixed increment to simulate the effect of a lower formant (i.e., the F1 condition). Even though the HI subjects performed worse overall, they showed more improvement than the NH subjects for both experiments. Therefore, we conclude that the NH and HI listeners generally benefit from spectral enhancements when discriminating a narrow-band signal in a broadband harmonic spectrum. Furthermore, HI subjects received more benefit from the enhancements, contrary to the finding of DiGiovanni *et al.* (2005). Finally, the analysis of auditory filter bandwidths showed that the NH listeners effectively use all the energy in their ERB, while some of the HI subjects underperform the predictions made by the overall-energy model, suggesting that, for these subjects, the reduced performance is not fully accounted for by their wider ERBs. It is possible that this is a result of the physiological limitation imposed by SNHL. In sum, spectral enhancement remains a viable goal to make modest improvements in the local signal-to-noise reduction for the detection of spectral peaks.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their insightful comments on an earlier version of the manuscript. This work was supported by an Ohio University Scholarly Activity Award.

Baer, T., Moore, B. C. J., and Gatehouse, S. (**1993**). "Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: effects on intelligibility, quality, and response times," J. Rehabil. Res. Dev. **30**, 49–72.

Bilger, R. C. (**1978**). "A revised critical band hypothesis," in *Hearing and Davis: Essays Honoring Hallowell Davis*, edited by S. K. Hirsh, D. H. Eldridge, I. J. Hirsh, and S. R. Silverman (Washington U.P., St. Louis, MO), pp. 191–198.

DiGiovanni, J. J., Nelson, P. B., and Schlauch, R. S. (**2005**). "A psychoacoustic evaluation of spectral enhancement," J. Speech Lang. Hear. Res. **48**, 1–15.

Dubno, J. R., and Dirks, D. D. (**1989**). "Auditory filter characteristics and consonant recognition for hearing-impaired listeners," J. Acoust. Soc. Am. **38**, 1666–1680.

Formby, G., Heinz, M. G., Luna, C. E., and Shaheen, M. K. (**1994**). "Masked detection thresholds and temporal integration for noise band signals," J. Acoust. Soc. Am. **96**, 102–114.

Franck, B. A., Sidonne, C., van Kreveld-Bos, G. M., Dreschler, W. A., and Verschuure H. (**1999**). "Evaluation of spectral enhancement in hearing aids, combined with phonemic compression," J. Acoust. Soc. Am. **106**, 1452–1464.

Giguère, C., and Smoorenburg, G. F. (**1998**). "Computational modeling of outer hair cell damage: Implications for hearing aid signal processing," in *Psychophysics, Physiology and Models of Hearing*, edited by T. Dau, B. Kollmeier, and V. Hohmann (World Scientific, Singapore), pp. 155–165.

Glasberg, B. R., and Moore, B. C. J. (**1986**). "Auditory filter shapes in subjects with unilateral and bilateral cochlear implants," J. Acoust. Soc. Am. **79**, 1020–1033.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Deviation of auditory filter shapes from notch-noise data," Hear. Res. **47**, 103–138.

Green, D. M., *Profile Analysis* (Oxford U. P., New York, **1988**), p. 93.

Heinz, M. G., and Formby, C. (**1999**). "Detection of time and bandlimited increments and decrements in a random-level noise," J. Acoust. Soc. Am. **106**, 313–326.

Horst, J. W., Javel, E., and Farley, G. R. (**1985**). "Extraction and enhancement of spectral structure by the cochlea," J. Acoust. Soc. Am. **78**, 1898–1901.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Lyzenga, J., Festen, J. M., and Houtgast, T. (**2002**). "A speech enhancement scheme incorporating spectral expansion evaluated with simulated loss of frequency selectivity," J. Acoust. Soc. Am. **112**, 1145–1157.

Miller, R. L., Calhoun, B. M., and Young, E. D. (**1999**). "Contrast enhancement improves the representation of /eh/-like vowels in the hearing-impaired auditory nerve," J. Acoust. Soc. Am. **106**, 2693–2708.

Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (**1982**). "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold," J. Acoust. Soc. Am. **72**, 1788–1803.

Plomp, R. (**1978**). "Auditory handicap of hearing impairment and the limited benefit of hearing aids," J. Acoust. Soc. Am. **63**, 533–549.

Sachs, M. B., and Young, E. D. (**1980**). "Effects of nonlinearities on speech encoding in the auditory nerve," J. Acoust. Soc. Am. **68**, 858–875.

Simpson, M., Moore, B. C. J., and Glasberg, B. R. (**1990**). "Spectral enhancement to improve the intelligibility of speech in noise for hearing-impaired listeners," Acta Oto-Laryngol. **469**, 101–107.

Summers, V, and Leek, M. R. (**1994**). "The internal representation of spectral contrast in hearing-impaired listeners," J. Acoust. Soc. Am. **95**, 3518–3528.