A PSYCHOACOUSTIC BASIS FOR SPECTRAL SHARPENING

Peggy B. Nelson, Jeffrey J. DiGiovanni and Robert S. Schlauch

Department of Communication Disorders
University of Minnesota, Twin Cities
nelso477@tc.umn.edu

ABSTRACT

Spectral sharpening of speech has been proposed as a possible method for improving speech understanding by listeners with hearing loss. Animal neural physiological data and cochlear models differ in their predictions as to the potential success of spectral sharpening for improved speech recognition. Data from early implementations of spectral sharpening are also ambiguous. Our recent investigation tested the theoretical viability of spectral sharpening for the detection and discrimination of spectral peaks in broadband noise. Results suggested that spectral sharpening (decrements) surrounding spectral peaks (increments) made those peaks more easily detected and discriminated than were spectral peaks without sharpening. All participants with moderate hearing loss demonstrated benefit from the spectral sharpening for peak detection and discrimination. Implications for speech processing algorithms will be discussed. (Work supported by the NIDCD grant R03 DC 04135).

1. INTRODUCTION

Listeners with cochlear hearing loss have difficulty understanding speech that is not fully explained by the audibility of the signal. As a result, amplification has been an important, but only partially successful tool for the remediation of sensorineural hearing loss.

Speech is the signal of primary importance to most hearing-impaired listeners, and it is characterized by rapidly changing complex formant structure. Figure 1 shows one 7-ms frame of a vowel, illustrating the complex combination of spectral peaks representing the vocal formant frequencies. Spectral peaks labeled C and D in the unamplified stimulus are primary cues to the identity of the vowel. They are separated by a spectral ‘trough’ or ‘valley’ of varying width (in this case approximately 1000 Hz) and depth (shown as d2). A hearing-impaired listener’s thresholds (shown as the black line) cause parts of the unamplified vowel to be inaudible. In this example, the higher-frequency formant peaks in the unamplified stimulus fall at or below the listener’s thresholds. When the vowel is amplified to maximize its audibility, as shown in the upper curve, the peak-to-trough ratio (d1) may be maintained, or may in fact be reduced due to a hearing aid’s amplitude compression.

Many listeners with cochlear hearing losses have reduced spectral resolution abilities, as measured by their auditory filter bandwidths. This reduced frequency selectivity may result in internal ‘smearing’ of the formant peaks in speech. This spectral smearing has been identified as one possible explanation for poor speech recognition performance, particularly in background noise.

One proposed remedy for spectral smearing is the application of spectral sharpening algorithms to speech signals [1]. In spectral sharpening algorithms, the peak-to-trough ratios are exaggerated, or enhanced. Presumably, the enhanced stimulus can compensate at least partially for the presumed broadened auditory filters in the impaired cochlea.

Researchers have disagreed as to whether spectral sharpening should result in improved speech recognition by impaired listeners. Results from specific implementations of sharpening algorithms [1] have produced marginal improvements in listener performance. Models of impaired cochlear function [2] have suggested that impaired auditory filters are sufficiently broad that even a sharpened vowel stimulus passed through the impaired model produces output that suggests no functional difference between the sharpened vowel and an unmodified vowel.

Promising results have been obtained from animal studies of auditory nerve coding of vowels. Miller et al. [3] studied cats with impaired cochleas, and found that contrast enhancement (exaggeration of peak-to-trough ratios) paired with high-frequency amplification successfully restored the auditory nerve’s coding of higher formants when amplification alone did not.

The current investigation was undertaken to test the theoretical viability of spectral sharpening. Listeners with and without hearing loss were tested for their detection and discrimination of spectral peak s in broadband noise.
discrimination of vowel-like spectral peaks in a broadband noise stimulus. Thresholds for detection of spectral peaks and spectral peak frequency difference limens were measured under two general conditions: unsharpened (minimal peak-to-trough ratios) and sharpened (enhanced peak-to-trough ratios in the spectral region immediately surrounding the spectral peak.) We hypothesized that if spectral sharpening improved listeners’ ability to detect and/or discriminate spectral peaks, then spectral sharpening remains a viable goal for hearing aid signal processing.

2. METHODS

The viability of spectral sharpening was evaluated by testing listeners’ detection thresholds and frequency discrimination of spectral peaks in broadband noise.

2.1. Subjects

Participants included four listeners with hearing loss described in Table 1. Four young adult listeners with normal hearing sensitivity also participated in the experiments.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age</th>
<th>HL .5k</th>
<th>HL 1k</th>
<th>HL 2k</th>
<th>HL 4k</th>
</tr>
</thead>
<tbody>
<tr>
<td>HI1</td>
<td>83</td>
<td>10</td>
<td>20</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>HI2</td>
<td>19</td>
<td>55</td>
<td>50</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>HI3</td>
<td>49</td>
<td>35</td>
<td>50</td>
<td>60</td>
<td>65</td>
</tr>
<tr>
<td>HI4</td>
<td>75</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 1: Hearing-impaired subject age (in years) and hearing loss (in dB HL) for .5, 1, 2, and 4 kHz.

2.2. Experiments and stimuli

All listeners participated in four experiments, including estimates of auditory filter bandwidths, nonsense syllable identification, estimates of spectral peak detection (with and without spectral sharpening), and estimates of spectral peak frequency discrimination (with and without sharpening). For all detection and discrimination experiments, subjects were trained to asymptotic performance levels. Four threshold estimates were averaged to obtain each data point reported here.

2.2.1. Auditory Filter Bandwidth Estimates

Thresholds were estimated for listeners’ detection of a 2kHz pure tone in a notched broadband noise using a three-interval forced choice three-down one-up paradigm. Notch widths included 0, 0.2, and 0.4 times the center frequency. Notched noises were presented at 55 and 85 dB HL for normal-hearing listeners and at 85 and 95 dB HL for hearing-impaired listeners. Results were fit with rounded exponential (roex) functions in order to estimate each listener’s auditory filter bandwidths.

2.2.2. Nonsense Syllable Identification

Listeners’ identification of nonsense syllables (NST) [4] was measured at listeners’ most comfortable level using frequency-shaped amplification to match the individual audiograms. Listeners circled the correct responses on a paper response form that contained the closed-set format. No feedback was given.

2.2.3. Spectral Peak Detection

Thresholds were next estimated for listeners’ detection of the presence of a spectral peak at 2 kHz in a broadband noise using a three-interval forced choice three-down one-up paradigm. The spectral peaks were increments of narrowband noise (NBN) centered at 2 kHz with 50Hz bandwidth. The increments were added to the broadband noise (BBN) that had an overall level of 85 dB SPL. The increments were added to the unmodified BBN with no sharpening for some conditions. For other conditions the increments were added with spectral sharpening (spectral decrements) applied to the adjacent frequency regions. For the sharpened conditions, the amount of spectral sharpening was varied in depth (3- and 6-dB) and in width (100 or 200 Hz above and below the increment.) Results from only two conditions will be reported here: the no-sharpening condition, and the 6-dB, 200-Hz sharpened condition. Schematics of the stimuli are shown in Figure 2 above.

2.2.4. Spectral Peak Frequency Discrimination

Listeners’ frequency discrimination for spectral peaks was then measured also using a three-interval forced choice three-down one-up paradigm. The 50-Hz wide NBN increment was added to the BBN, with the standard increment centered at 2 kHz. The center frequency of the signal NBN was higher than the standard, and varied adaptively. The overall level was roved by +/- 2 dB. Spectral peak frequency discrimination was measured with and without spectral sharpening, as above. Results will be reported for the no-sharpening condition and for the 6-dB, 200-Hz wide sharpened condition.

2.3. Instrumentation

Stimuli were generated using MatLab 6.0 interfaced with Tucker-Davis Technologies hardware. All stimuli were generated digitally with 16-bit resolution at a sampling rate of 22.05 kHz. Stimuli were low-pass filtered at 7.5 kHz and presented monaurally through TDH-49P headphones. Subjects were seated in a sound-treated chamber and indicated their responses using a keyboard and monitor.
3. RESULTS

The results for the four experiments are shown in Tables 2 through 4 below. In general, results support the hypothesis that spectral sharpening can improve listeners’ detection and discrimination of spectral peaks in broadband signals.

3.1. Auditory Filter Bandwidths

The estimated filter bandwidths for all listeners are shown in Table 2. The center frequency was 2000 Hz. Hearing-impaired listeners, in general, showed auditory filters that were 2 to 5 times the width of the filters measured in the normal-hearing subjects at 85 dB SPL. Considerable variability among the hearing-impaired subjects is seen, even though the hearing thresholds of the listeners were very similar at 2000 Hz. Auditory filter bandwidths did not seem to be related to age, as the impaired listeners with the narrowest auditory filters (HI2 and HI4) were 19 and 75 years of age, respectively.

<table>
<thead>
<tr>
<th>Subj</th>
<th>Filter BW (Hz) 85 dB SPL</th>
<th>Filter BW (Hz) 95 dB SPL</th>
<th>NST score</th>
</tr>
</thead>
<tbody>
<tr>
<td>HI1</td>
<td>1519</td>
<td>1024</td>
<td>64%</td>
</tr>
<tr>
<td>HI2</td>
<td>469</td>
<td>381</td>
<td>89%</td>
</tr>
<tr>
<td>HI3</td>
<td>1025</td>
<td>444</td>
<td>76%</td>
</tr>
<tr>
<td>HI4</td>
<td>331</td>
<td>306</td>
<td>84%</td>
</tr>
<tr>
<td>NH1</td>
<td>275</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>NH2</td>
<td>232</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>NH3</td>
<td>126</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>NH4</td>
<td>244</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 2: Subject auditory filter bandwidths (in Hz) and Nonsense Syllable Test scores (in percent correct).

3.2. Nonsense syllable testing

The Nonsense Syllable Test scores for the impaired listeners are shown as percent correct in Table 2 above. Some normal-hearing listeners were also tested using the NST, all obtaining scores of 100% correct. Performance varied considerably among the hearing-impaired listeners in this study, and appeared to be generally related to auditory filter bandwidths. The listeners with the widest auditory filters (HI1 and HI3) attained the lowest NST scores; those with the narrowest filters demonstrated NST scores better than 80% correct.

3.3. Spectral peak detection

Results for the spectral peak detection experiment are shown in Table 3 for all listeners. Results are shown as 10 log (dI/I). Hearing-impaired listeners demonstrated elevated spectral peak detection thresholds when compared to normal-hearing listeners. All listeners showed significant improvement in peak detection thresholds for the sharpened stimuli over the unmodified stimuli. Improvement in thresholds ranged from 2 to 3 dB for the impaired listeners, and from 3 to 5 dB for the normal-hearing listeners. Spectral peak detection thresholds do not appear to be strongly related to auditory filter bandwidths; listener HI1 showed the lowest peak detection threshold, but had the widest auditory filter bandwidth. Improvement from the sharpening also did not appear to be related to age or to auditory filter bandwidth. Despite obvious differences in auditory filter bandwidths shown previously, all the impaired listeners improved by 2 – 3 dB when detecting the 6-dB 200-Hz wide spectral decrements that surrounded the target spectral peak.

<table>
<thead>
<tr>
<th>Subj</th>
<th>Peak Detection in BBN, with no sharpening</th>
<th>Peak Detection with sharpening (6-dB, 200-Hz)</th>
<th>Improvement (in dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HI1</td>
<td>8.1</td>
<td>5.6</td>
<td>2.5</td>
</tr>
<tr>
<td>HI2</td>
<td>9.3</td>
<td>6.8</td>
<td>2.5</td>
</tr>
<tr>
<td>HI3</td>
<td>11.6</td>
<td>9.8</td>
<td>1.8</td>
</tr>
<tr>
<td>HI4</td>
<td>12.4</td>
<td>9.2</td>
<td>3.2</td>
</tr>
<tr>
<td>NH1</td>
<td>9.1</td>
<td>5.5</td>
<td>3.6</td>
</tr>
<tr>
<td>NH2</td>
<td>6.0</td>
<td>1.5</td>
<td>4.6</td>
</tr>
<tr>
<td>NH3</td>
<td>8.2</td>
<td>3.2</td>
<td>5.0</td>
</tr>
<tr>
<td>NH4</td>
<td>6.3</td>
<td>1.8</td>
<td>4.5</td>
</tr>
</tbody>
</table>

Table 3: Spectral peak detection threshold results for no-sharpening and 6-dB, 200-Hz sharpening, in dB (10 log (dI/I)). Improvement is shown in the right column.

3.4. Spectral peak frequency discrimination

Results for the spectral peak frequency discrimination experiment are shown in Table 4 for all listeners. Results are shown as just-noticeable-differences in Hz for the no-sharpening condition and for the 6-dB, 200-Hz sharpening. The ratio of the ‘sharpened’ to the ‘no-sharpening’ DLs is shown at the right. As can be seen, all listeners demonstrated significantly better frequency difference limens when spectral sharpening (decrements) of 6-dB depth and 200-Hz width surrounded the spectral peak. The resulting DL for the sharpened spectral peaks was approximately ¾ the original DL for the impaired listeners. For the normal-hearing listeners, the DL for the sharpened condition was approximately 2/3 of their original DL.

<table>
<thead>
<tr>
<th>Subj</th>
<th>Frequency DL1 in BBN, with no sharpening</th>
<th>Frequency DL2 with sharpening (6-dB, 200-Hz)</th>
<th>FDL2/ FDL1</th>
</tr>
</thead>
<tbody>
<tr>
<td>HI1</td>
<td>130.7</td>
<td>94.1</td>
<td>.72</td>
</tr>
<tr>
<td>HI2</td>
<td>68.2</td>
<td>52.3</td>
<td>.77</td>
</tr>
<tr>
<td>HI3</td>
<td>179.7</td>
<td>132.2</td>
<td>.74</td>
</tr>
<tr>
<td>HI4</td>
<td>229.1</td>
<td>178.8</td>
<td>.77</td>
</tr>
<tr>
<td>NH1</td>
<td>86.9</td>
<td>49.5</td>
<td>.57</td>
</tr>
<tr>
<td>NH2</td>
<td>35.3</td>
<td>23.6</td>
<td>.67</td>
</tr>
<tr>
<td>NH3</td>
<td>54.2</td>
<td>29.3</td>
<td>.54</td>
</tr>
<tr>
<td>NH4</td>
<td>42.3</td>
<td>30.1</td>
<td>.71</td>
</tr>
</tbody>
</table>

Table 4: Spectral peak frequency discrimination results shown in difference limens (in Hz). The ratio of the DLs is shown in the right column.
3. DISCUSSION AND CONCLUSIONS

Results from this study indicate that spectral sharpening can improve listeners’ detection and discrimination of spectral peaks in broadband signals. The results are in agreement with the physiological findings of Miller et al. [3] and have direct implications for speech processing algorithms in hearing aids.

First, all listeners, with and without hearing loss, demonstrated benefit from the spectral sharpening. Normal-hearing listeners demonstrated greater benefit than did impaired listeners. This was not an expected finding, and caused us to question whether only those listeners with near-normal auditory filters could benefit from spectral sharpening. However, all four impaired listeners studied showed benefit from the sharpening, regardless of their auditory filter bandwidths. The presence of sharpening around spectral peaks improved impaired listeners’ detection thresholds by 2-3 dB and reduced frequency difference limens to ¾ of their original values.

Apparently the amount of benefit obtained from sharpening may be somewhat independent of the auditory filter bandwidths of the listeners. Although auditory filter bandwidths were generally related to NST scores, the benefit from sharpening was not. This seems contrary to the model-based theories of Giguere and Smoorenberg [2]. Apparently a simple modeling of the signal-to-noise ratio in each critical band does not fully predict the auditory system response to spectral peaks.

In addition, the benefit obtained from sharpening did not appear to be related to subject age. One might hypothesize that elderly impaired listeners would perform differently when compared to younger impaired listeners. In this investigation, that hypothetical difference was not observed. The two elderly listeners (HI1 and HI4) were vastly different in their auditory filter bandwidths. HI1 demonstrated very wide auditory filters, but good thresholds for detecting the spectral peaks and small difference limens for frequency. In contrast, HI4 demonstrated narrower auditory filters, poorer thresholds for detecting spectral peaks, and poorer frequency difference limens. Both demonstrated significant (and similar) benefit from the sharpening algorithm for improved detection and discrimination. The two younger subjects in this investigation demonstrated differences in performance of the same magnitude.

Based on these results, one would predict that listeners with moderate hearing impairments, regardless of auditory filter bandwidths, can realize significant benefit from spectral sharpening. Previous pilot results of a real-time speech sharpening algorithm [5] suggest that improved speech recognition in quiet and in noise can be attained. Improving peak-to-trough ratios in running speech could make formant peak frequencies easier to identify and to discriminate. In addition, under some conditions spectral sharpening can have the effect of reducing background noise and improving signal-to-noise ratios. If one presumes that the spectral peaks arise from the vocal signal, and lower-level noise has the effect of filling the troughs, then expanding peak-to-trough ratios can improve signal-to-noise ratio.

Results from these basic psychoacoustic detection and discrimination experiments demonstrate that spectral sharpening is a viable signal-processing strategy that can result in improved listener performance. Additional factors affecting implementation will also be discussed.

5. REFERENCES