ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Psychological and Physiological Acoustics
Session 1pPPb: Psychoacoustics and Perception (Poster Session)

1pPPb12. Investigating the effects of intensity on the bandwidth of peripheral filtering in an amplitude-modulation notch detection task
Matthew L. Richardson, Allison I. Shim and Bruce G. Berg*

*Corresponding author's address: Cognitive Sciences, University of California, Irvine, 2201 SBSG, Irvine, CA 92697-5100, bgberg@uci.edu

The effect of intensity on the effective bandwidth of auditory temporal processing is investigated. Thresholds for detecting sinusoidal amplitude-modulation of a 200-Hz wide band of noise centered at 1000 Hz are measured in the presence of a notched noise masker. The masker consists of two, 200-Hz wide, unmodulated bands of noise placed at frequencies above and below the modulated band. Thresholds for a modulation rate of 10 Hz are estimated for different notch bandwidths ranging from 100 Hz to 2740 Hz. The use of a slow modulation frequency aims to avoid possible central limitations of temporal processing at higher modulation frequencies. Intensity is varied across two conditions, with all three bands of noise presented at either 40 dB SPL or 85 dB SPL. Threshold functions for the two intensity levels are essentially identical. The maximum notch width at which an effect of the masker is observed is approximately 500 Hz. The results are consistent with a hypothesis that the filtering characteristics of temporal processing (e.g. envelope model) and spectral processing (e.g. power spectrum model) are different.

Published by the Acoustical Society of America through the American Institute of Physics
INTRODUCTION

Estimates of effective bandwidths in early stage auditory processing are commonly thought to provide initial spectral and temporal resolution to the auditory system. In traditional theories of spectral resolution, sound information initially passes through a set of narrow bandwidth filters tuned to different characteristics frequencies. This is the basis of the audio-frequency filterbank underlying critical band theory. Furthermore, this process is sometimes assumed to reflect the physical properties of the basilar membrane (Patterson et al. 1992; Dau et al. 1997a,b). The fact that the extent of basilar member displacement increases with sound intensity implies that the bandwidths of peripheral auditory filters also increase with intensity (Moore and Glasberg, 1987). Data reported here are inconsistent with this fundamental premise.

An implicit assumption of current psychoacoustic theory is that the same initial processes that govern spectral resolution also govern the peripheral filtering of temporal information. However, much evidence points to discrepancies between the effective bandwidths for these two kinds of information. Data from a wide variety of discrimination tasks that ostensibly isolate information in the temporal domain produce bandwidth estimates that exceed those of spectral critical bands. (Mathes and Miller, 1947; Viemeister, 1979; Forrest and Green, 1987; Nelson, 1994; Nelson and Schroder, 1995; Strickland and Viemeister, 1997; Berg, 1996, 2007; Oxenham and Dau, 2001; Turner, 2010; Tabuchi, Borucki, and Berg, 2012; Buss, Hall, and Grose, 2013). A valid criticism is that the isolation or control of spectral and temporal information in an experiment is not always a perfect endeavor, so many of these studies cannot stand alone in providing unchallenged evidence for wider bandwidths. One example is the hypothetical use of auditory distortion products in AM/QFM discrimination tasks (Bunnen, 1975). Collectively, however, there is an increasing corpus of data supporting a converging theme that is inconsistent with textbook definitions of critical band theory.

Shim (2012) designed an experiment to circumvent the problems and issues of earlier studies. She applied a version of the tone-in-notched-noise detection procedure traditionally used in spectral critical band estimates (e.g. Patterson, 1976). Instead of a pure tone however, listeners detected sinusoidal amplitude modulation of a narrowband of noise. A modulation rate of 10 Hz was used in order to avoid possible central limitations of temporal processing at higher modulation frequencies (Strickland, 2000). The band of noise was placed between two masking noise bands that were unmodulated (the notches) and thresholds for modulation detection were measured as a function of notch width, or the distance between the noise maskers. The maximum notch width at which the maskers had a discernable effect on thresholds was estimated (i.e. the bandwidth where the threshold function becomes constant for wide notch widths). Four different center frequencies for the modulated noise were tested (0.6, 1, 2, and 4 kHz). All four yielded a ratio of center frequency to maximum notch width, designated $Q_{AM}$, that was consistently close to 2. This suggests broader tuning than estimates of spectral bandwidths which typically have $Q$-values close to 6.

The purpose of the current study is to further investigate the emerging discrepancies between measures of spectral and temporal peripheral filtering. Specifically, we use Shim’s (2012) paradigm to examine the effects of intensity on the effective bandwidths of temporal processing. In order to observe changes in masking that occur as a function of intensity, one center frequency (1000 Hz) is tested at three overall presentation levels—40, 60, and 85 dB SPL. If temporal resolution is limited by “basiliar membrane filtering”—as often thought to be the case for spectral resolution—bandwidth estimates should exhibit a widening effect with increased sound intensity. On the other hand, bandwidth estimates that are invariant with respect to intensity may further suggest that the initial filtering properties for temporal and spectral information are different from one another.

SUBJECTS

Three naïve paid volunteer listeners participated as subjects in this experiment. Two female and one male subject ranged in age from 20 to 49. All subjects were screened for normal hearing and displayed less than 20 dB hearing loss for pure tones ranging from 0.5 to 8 kHz. All procedures are approved by the UCI Institutional Review Board.

STIMULI

Data are collected using a two-interval, forced-choice, adaptive staircase procedure (Levitt, 1971). The signal is a 200-Hz wide, Rayleigh distributed narrowband of noise centered at 1kHz and sinusoidally modulated at a constant modulation frequency of 10Hz. The signal is paired against a non-signal interval with a 200-Hz narrowband
noise without modulation. The masker consists of two, 200-Hz wide, unmodulated bands of noise placed equidistant logarithmically above and below the signal/non signal band, creating the “notch.” Notch width is defined by the distance between the high edge of the low frequency masker and the low edge of the high frequency masker. Signal and noise bands are presented with equal intensities. Each signal/non-signal interval has a 500-ms duration with a 20-ms linear onset and offset ramp. The inter-interval duration is 500-ms.

Digitalized waveforms are played through a two-channel, digital-to-analog converter (E-MU 020 Audio/MIDI interface) with a 44 kHz sampling rate. The signal is then passed through a manual attenuator for intensity calibration, then to a TDT System II headphone buffer where it is split and sent to both channels in set of Sennheesser HD414SL headphones. Listeners are seated in a single-walled, sound attenuating chamber and give their responses on a keyboard. Feedback is provided on a computer monitor after each trial.

**PROCEDURE**

Two overall intensity levels are tested for each subject: 40 and 85dB SPL for one subject and 60 and 85 dB SPL for two subjects. For each intensity level, subjects complete 8 conditions in which notch width is systematically varied around the signal/non-signal bands. Notch bandwidths range from 100 Hz to 2740 Hz. Each notch condition consists of 5 blocks of 70 trials. For each trial, subjects are randomly presented with a signal and no-signal signal interval and instructed to respond for which interval they heard the modulated signal. Modulation depth, $m$, is decreased by 2dB with two consecutive correct responses and increased by 2dB with one incorrect response. Detection thresholds are determined by averaging all signal levels after the first four reversal points. For difficult conditions in which $m$ may exceed a value of 1 and cause over-modulation, $m$ is held at a maximum of 1 until two consecutive correct responses are made.

**RESULTS**

For each subject, averaged thresholds are analyzed as a function of notch bandwidth. Threshold functions derived in this way are used to examine the maximum notch width at which masking of the signal first becomes effective, degrading the detectability of the SAM signal. This point is referred to as the breakpoint. Data for each subject are presented in Figure 1 and estimated breakpoints are indicated by a filled square.

**FIGURE 1:** Threshold functions for three subjects. Subjects JH and VH show thresholds for 60dB and 85dB. Subject MR shows thresholds for 40dB and 85dB. Closed triangles show the average threshold for 5 blocks and error bars indicate the standard error. Closed squares indicate the breakpoint estimates obtained with the regression method.
For each threshold function shown in Fig. 1, breakpoints are obtained by minimizing the least-squared error of fits to the data with two lines using linear regression. The first line consists of the subset of data with two or more points that yields the minimum standard error to a line with a constrained slope of zero. The zero slope constraint has the theoretical implication that detection thresholds are unaffected by maskers at these notch widths. The second line is fit to the remaining data points and the breakpoint is defined as the intersection of the two lines. Estimated breakpoints for each subject are provided in Table 1.

### Table 1: Breakpoint results from regression analysis.

<table>
<thead>
<tr>
<th>Intensity (dB SPL)</th>
<th>Breakpoint (dB SPL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40/60</td>
<td>511.15</td>
</tr>
<tr>
<td>85</td>
<td>436.7</td>
</tr>
</tbody>
</table>

As seen in Fig 1 and Table 1, all subjects yield threshold functions and breakpoints that are essentially identical with no apparent effect of the stimulus intensity. The maximum notch width at which an effect of the masker is observed is approximately 500Hz. This result is further analyzed by converting breakpoint estimates into a dimensionless value, $Q_{AM}$, which is calculated by dividing the center frequency (1kHz) by the estimated breakpoint. $Q$-values provide an estimate of the sharpness of tuning regarding theoretical filtering characteristics of spectral and temporal information, with higher $Q$-values indicate sharper tuning. Here, a $Q_{AM}$ of approximately 2 is obtained for all intensities. This is shown in Figure 2.

**Figure 2**: $Q_{AM}$ for each individual are indicated by open symbols. Average $Q_{AM}$ across all subjects for 60dB and 85dB are indicated by filled shapes. At all intensities, $Q$ is approximately equal to 2.

**Discussion**

Results from this experiment are consistent with past studies that have demonstrated broad filtering for initial temporal processes. This is highlighted by the low $Q_{AM}$ obtained in this study. Auditory spectral filters on the other hand—long defined by their “critical bandwidths”—are characterized by sharply tuned filters with higher $Q$-values. This is important because sharply tuned filterbanks are often incorporated into models of peripheral processing for both spectral and temporal information. Such models include the gammatone filterbanks of Patterson et al. (1992) and Dau et al. (1997a,b). In these models, the gammatone filterbank is an initial stage of narrowly tuned
bandpass filters consistent with critical band theory. One assumption about this stage is the process known as “basilar membrane filtering,” in which spectral and temporal information is initially parsed along the basilar member through multiple narrow channels.

This study questions assumptions related to basilar membrane filtering in two ways, first in the evidence it provides for broadly tuned filters, and secondly in the lack of intensity effects. It is well known that both the displacement of the basilar membrane and the tuning of primary afferents increase with increasing intensity. Critical bandwidths obtained with tone-in-noise detection tasks also increase with intensity (Moore and Glasberg, 1987). The lack of an intensity effect for this study suggests that something other than basilar membrane filtering is being measured. Similar results have been reported by Wojtczak (2011), where the effects of carrier level on tuning in a modulation masking experiment were tested using noise and tonal carriers. The results for the noise carrier showed no effects of carrier level on thresholds or estimated bandwidths of modulation filters.

In light of these findings, altering the bandwidths of auditory filters as a function of intensity might create an unstable system with respect to variables such as a change in distance. One hypothesis is that “auditory filters” associated with the temporal system are constructed through neural integration at a level beyond the auditory nerve, possibly at the cochlear nucleus. The bandwidth of filters for the temporal system might depend on neural micro-circuitry rather than the extent of basilar membrane displacement. As noted by Berg (2013), it may be useful to consider peripheral integration rather than peripheral filtering with regards to the temporal processing system.

Combining the results of this experiment with previous studies of temporal resolution, it is reasonable to question the basis of a singular auditory process governing both temporal and frequency information at the peripheral level. By using the same notched-noise paradigm used to determine the characteristics of filters in critical band theory, it has been shown that the model of narrow bandwidth filters originating in the basilar membrane is too limiting to sufficiently transmit temporal modulation information. In response, Shim (2012) proposes that the wider bandwidths observed in temporal tasks be recognized as a “temporal critical band,” complimenting the traditional “spectral critical band. This opens the possibility of a more dynamic system that can adjust its bandwidth to the stimulus context (Strickland and Viemeister, 1997). Alternatively, Berg (2004) proposed discrete filtering systems for spectral and temporal processes that reflect this distinction. What is certain, is that the auditory filter bandwidths derived in temporal tasks require further investigation and characterization to account for a more accurate understanding of early stage auditory processing.

REFERENCES


