A THEORETICAL STUDY OF INFORMATION TRANSMISSION IN THE AUDITORY SYSTEM USING SIGNAL DETECTION THEORY: FREQUENCY DISCRIMINATION BY NORMAL AND IMPAIRED SYSTEMS

Lisa G. Huettel and Leslie M. Collins

Department of Electrical and Computer Engineering
Duke University, Durham, NC 27708-0291
lisa.huettel@duke.edu, lcollins@ee.duke.edu

ABSTRACT

In this paper, we have investigated the differences between normal and impaired auditory processing for a frequency discrimination task by analyzing the responses of a computational auditory model using signal detection theory. Two detectors, one using all of the information in the signal, the other using only the number of neural responses, were implemented. An evaluation of the performance differences between the two theoretical detectors and experimental data may provide insight into quantifying the type of information present in the auditory system as well as whether the human auditory system uses this information efficiently. Results support previous hypotheses that, for low- and mid-range frequencies, the auditory system is able to use temporal information to perform frequency discrimination [8]. The results also suggest that some temporal information is represented in the neural spike train, even at high frequencies. However, the ability of the auditory system to use this information deteriorates at higher frequencies.

1. INTRODUCTION

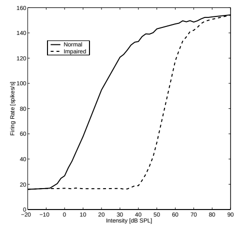
In the auditory literature, signal detection theory (SDT) was originally used to generate theoretical predictions of psychophysical performance based on the stimuli themselves (*e.g.*, [11], [15]). The theoretically optimal detector greatly outperformed the actual human auditory system. This could be attributed to the fact that the theoretical approach, based on the acoustic signals measured outside the ear, essentially ignored the signal transformations that occurred as the acoustic signal propagated through the auditory system. Later, Siebert addressed this issue by generating simulated auditory responses to an acoustic stimulus using a functional auditory model [12, 13]. Using this approach, Siebert was able to study frequency and intensity discrimination in a normal auditory system. More modern computational auditory models (*e.g.*, [1], [10]) may provide more accurate predictions of neural responses to a wider range of acoustic stimuli than the analytical models used previously. Recent work has analyzed several computational models using SDT to predict psychophysical performance on a simultaneous masking task [6] and Heinz *et al.* (1999) have used a different computational model to predict performance for normal-hearing individuals on frequency and intensity discrimination tasks [5]. In this paper, we have developed theoretical predictions of psychophysical performance on a frequency discrimination task by analyzing the signals predicted by the Auditory Image Model (AIM) [10] using SDT. We incorporated a spike generator into the otherwise deterministic model using a Poisson process to represent the internal noise present in the auditory system [7]. We have also investigated differences between normal and impaired auditory processing by modifying AIM as described in Section 2.

Three different approaches were used to generate theoretical predictions of performance. These approaches were selected to provide insight into quantifying the information present in the different neural populations. In one case, the theoretical detector was allowed to process all of the information present in the responses generated by the model. In the second case, temporal information was removed from the signals by only considering rate information before the detector was allowed to process the data. This approach mimics the rate-based processing hypothesis for frequency discrimination. These two cases were chosen so that the differences between the two processors and the experimental data might be used to infer how the auditory system uses temporal information. As both of these approaches are computationally intense, a third approach, the Cramèr-Rao Lower Bound, was also employed.

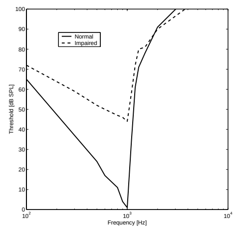
2. IMPAIRING THE MODEL

A flat (constant as a function of frequency), 40-dB hearing loss was induced by shifting the thresholds of the normal model by 40 dB. To verify the impaired model, comparisons were made between rate-intensity curves, neural tuning curves, and response area curves generated by the normal and impaired models. These comparisons are presented in Figure 1. The rate-intensity curve (Figure 1a), which plots the firing rate as a function of signal intensity, shows that the threshold, or point at which the firing rate exceeds the spontaneous rate, for the impaired model is shifted by approximately 40 dB. The steeper slope for the impaired model also illustrates the phenomenon of loudness recruitment [9]. In addition to indicating a 40-dB threshold shift for the impaired model, the neural tuning curves (presented in Figure 1b for a fiber with a best frequency (BF) at 1045 Hz responding to a 1000-Hz tone by plotting the threshold as a function of frequency) depict a broadening of the filter in the impaired case. This feature has implications for the design of hearing aids in that it suggests that more than simple amplification might be necessary to restore the normal functioning of an impaired ear. Finally, the response area curve (Figure 1c), which plots the firing rate as a function of filter center frequency, indicates how an elevated threshold affects the responses of fibers tuned to frequencies other than the stimulus frequency. Specifically, more fibers respond to a stimulus at a given

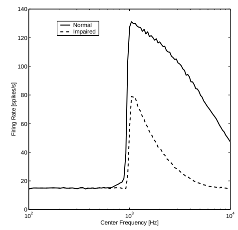
Support for this work was provided by the Whitaker Foundation and the Army Research Office.



(a) Rate-Intensity curves



(b) Neural tuning curves



(c) Response area curves

Fig. 1. Comparison between normal (solid lines) and impaired (dashed lines) models: (a) Rate-Intensity curves for a fiber with CF=1045 Hz in response to a 1000-Hz tone; (b) Neural tuning curves for a fiber with CF=1045 Hz in response to a 1000-Hz tone; (c) Response area curves for a 1000-Hz tone at 55 dBSPL.

level in the normal system than in the impaired system. The implications are that, since more fibers respond, more information may be available in the normal system on which to base a decision. Based on these three comparisons, it appears that, to at least a first order, the desired impairment has been induced and, consequently, that it is reasonable to make comparisons between the normal and impaired systems on a frequency discrimination task.

3. METHODS USED TO CALCULATE Δf_{JND}

Theoretical predictions of the difference limen for frequency (DLF) were generated three ways. Two detectors, one utilizing the entire neural response predicted by the computational model and another restricted to using the number of spikes only, were derived using SDT. The first detector, referred to as the optimal detector, was applied to the spike trains generated by a non-stationary Poisson process driven by the neural firing rate predicted by the auditory model and was implemented to provide a bound on performance. For a single fiber located in the i^{th} filter, the detector is:

$$\sum_{n=1}^{N} \frac{\ln r_i(ST_n, f)}{\ln r_i(ST_n, f+\Delta f)} \underset{H_0}{\overset{H_1}{\gtrless}} \beta, \qquad (1)$$

where $r_i(ST_n, f)$ is the time-varying firing rate in response to a tone with frequency f sampled at ST_n, the time of the n^{th} spike.

The second detector, based on the total number of spikes on each fiber, was implemented to study the theoretical contribution of temporal information on frequency discrimination. Temporal information was removed from the signal by integrating over the neural firing rate and using this *average* rate to drive a stationary Poisson process. Then, the total number of spikes was computed, disregarding arrival times. The form of this detector was simply:

$$N \underset{H_0}{\overset{H_1}{\gtrless}} \beta, \qquad (2)$$

where N is the total number of spikes observed in a given fiber. The difference between the count-based results and the optimal detector result may provide insight into how much potentially useful temporal information is present in the auditory system. For both detectors, it was assumed that the filters were independent. Therefore, the overall detectability, d', was equal to the square root of the sum of the squares of the detectability in each individual filter. An adaptive procedure was used to determine Δf_{JND}, defined as the difference in frequency required to obtain $d' = 1$.

The third method used to calculate the DLFs was to compute the Cramèr-Rao Lower Bound [2]:

$$\Delta f_{JND} = \left\{ \sum_i \int \frac{1}{r_i(t,f)} \left[\frac{\partial r_i(t,f)}{\partial f} \right]^2 dt \right\}^{-(1/2)}. \qquad (3)$$

This approach is only optimal (i.e., equivalent to the optimum detector) when an efficient estimator exists. Therefore, if such an estimator does not exist, the CRLB may underestimate the bound.

4. DISCRIMINATION TASK

Frequency discrimination based on the model was determined as a function of both stimulus frequency and sensation level. First, the DLF of a stimulus presented at 55 dBSPL was measured at the test frequencies of 250, 500, 1000, 2000, 4000, and 8000 Hz. Next, the DLF was measured at 1000 Hz for stimuli presented at sensation levels ranging from 10 to 70 dB. In all simulations, the stimuli were 200 ms in duration with a 20 ms rise/fall time. An adaptive procedure was used to determine the theoretical DLF (see Section 3). Simulation results are compared to human data from experiments with comparable stimulus parameters [3, 8, 14, 16].

5. SIMULATION RESULTS

In the first simulation, frequency discrimination was determined as a function of frequency for a stimulus presented at a fixed level of 55 dBSPL. The normal-hearing data used for comparison was collected by Moore (1973) for stimuli 200-ms long presented at 60 dBSPL [8]. The geometric mean of the data (in terms of the Weber fraction, $\Delta f/f$) for three subjects is plotted with a solid line in Figure 2. The sharp increase in the Weber fraction at higher frequencies suggested to Moore that a temporal mechanism was more efficient at frequencies below 5 kHz and a place mechanism was more efficient above this frequency [8]. A similar trend can be observed in the impaired-hearing data collected by Simon and Yund (1993) with the sharp change occurring closer to 2 kHz (dashed line in Figure 2). They measured frequency discrimination for 34 hearing-impaired subjects using stimuli 320 ms in duration presented at 80 dBSPL or 10 dBSL, whichever was larger [14].

As seen in Figure 2, none of the theoretical detectors accurately predict the absolute level of human performance. The larger-than-experimental values of $\Delta f/f$ for the count-based detector may be attributed to the fact that the model only simulates 100 nerve fibers, whereas there are approximately 30000 fibers in a

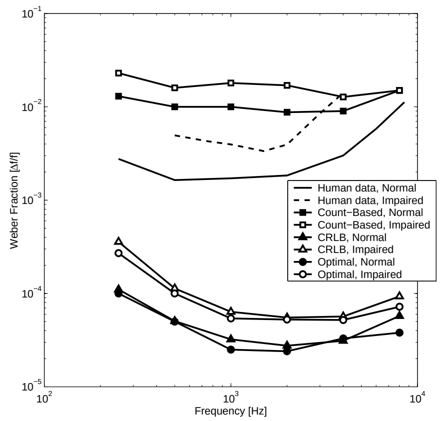


Fig. 2. Frequency discrimination as a function of frequency.

f [Hz]	$\Delta f_I / \Delta f_N$			
	Expt.	**Optimal**	**CRLB**	**Count-Based**
500	2.1	2.0	2.2	1.6
1000	1.6	2.2	2.0	1.8
2000	2.5	2.2	2.0	1.9
4000	—	1.6	1.8	1.4
8000	—	1.9	1.6	1.0

Table 1. The experimental and theoretical ratios of the DLF in impaired and normal ears as a function of frequency.

healthy human auditory system [4]. Theoretically, each fiber contains additional information; therefore, if more fibers were simulated, theoretical performance would improve by a factor on the order of the square root of the factor of increase in number of fibers (*i.e.*, approximately $\sqrt{300}$). This would result in theoretical DLFs less than experimental data, as in the case of the optimal and CRLB detectors (whose predictions would also decrease if the number of fibers modeled was increased). When theoretical predictions of $\Delta f / f$ are smaller than those measured experimentally, it suggests that the theoretical detectors have more information to operate on or are more efficient than the human auditory system. In the first case, this implies that the model does not accurately simulate the information loss in the auditory system. One solution is to modify the model. The other alternative is that the information exists in the signals encoded in the peripheral auditory system, but the auditory system does not use this information efficiently. In this case, the auditory system does not behave optimally and the theoretical detector can be modified to reflect this.

Clearly, we are interested in predicting not just absolute levels, but also trends in the data. The count-based predictions are fairly flat across frequency, whereas the experimental data show a marked increase in the Weber fraction above 2000 Hz. Since a strategy based solely on spike counts is approximately independent of frequency (vis-à-vis the count-based detector), it can be inferred that a different mechanism mediates the performance of the auditory system above 2000 Hz causing a sharp increase in the Weber fraction. This supports Moore's interpretation of experimental results suggesting a temporal/place mechanism tradeoff [8].

Examination of the theoretical predictions obtained using the optimal and CRLB detectors reveal that these two approaches yield similar results. However, unlike experimental results, the optimal and CRLB detectors show a large decrease in the Weber fraction between 250 and 1000 Hz and only a slight increase above 2000 Hz. This behavior is more like human performance than the behavior of the count-based detector, but the differences at higher frequencies suggest once again that either the theoretical detectors utilize temporal information more efficiently than the auditory system or the signals generated by the model contain more information than the signals actually present in the auditory system. One possible solution, not pursued in this work, is to tweak the parameters of the model, such as the slope of the filters, until more accurate predictions are obtained.

Table 1 lists the ratio of the impaired DLFs and the normal DLFs ($\Delta f_I / \Delta f_N$) for experimental data as well as the three theoretical cases. Experimental data collected by Turner (1987) for four normal and four impaired subjects are used to calculate the experimental ratio in Table 1. The test signals were 205-ms in duration and presented at 80 dBSPL [16]. All three theoretical detectors yield ratios on the same order of magnitude as the experimental data. However, none of them reflect the V-shaped trend evident in the experimental data. One must exercise caution when comparing theoretical and experimental results since the experimental subjects did not have exactly the same impairment as was simulated in the model. It can be concluded that the simulated 40-dB flat hearing loss affects the frequency discrimination performance of the optimal and CRLB detectors approximately equally across the mid-range frequencies, with a slight decrease above 2000 Hz. However, in the case of the count-based detector, it appears that the ratio converges to unity at 8000 Hz. These results suggest that a more accurate model of hearing impairment is needed.

The results of the second simulation, frequency discrimination at 1000 Hz as a function of sensation level, are presented in Figure 3. Due to limitations in the model, impaired simulations were limited to sensation levels 40 dB and smaller. Human data for a similar task ($f = 1000$ Hz, 300-ms duration) from Freyman and Nelson (1991) for normal subjects (solid line) and impaired subjects (dashed line) are also shown [3]. In these data, there is a decrease in the DLF as the sensation level increases from 10 to 20 dBSL and then a more gradual decrease over the remaining range of levels. A similar trend is observed in the theoretical predictions obtained using the optimal and CRLB detectors, although generally the slope is greater than that of the experimental data. In contrast, the predictions generated using the count-based detector display the opposite trend at high levels with the DLF *increasing* with increasing level. This increase may be attributed to the spread in response across fibers in the model as the stimulus intensity increases. The count-based detector, which discriminates between the two frequencies based on the total number of spikes on each fiber, has increasing difficulty in distinguishing between frequencies at higher levels because the firing rates of more and more fibers are saturated resulting in the same number of spikes in each fiber. Thus, contrary to our intuition, the amount of useful information available to the count-based detector actually decreases as the stimulus level increases above a certain level. This suggests that the auditory system is able to use temporal information, like the optimal and CRLB detectors. Thus, it appears that, although

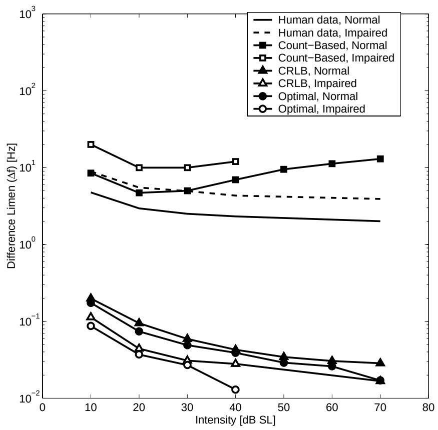


Fig. 3. Frequency discrimination as a function of sensation level.

none of the theoretical detectors accurately predict the absolute level of human performance (for reasons discussed previously), the optimal or CRLB detector provides the more accurate prediction of the trend in human performance as a function of level.

A different conclusion must be drawn, however, when comparing normal and impaired results. Table 2 lists the ratio of DLFs between impaired and normal systems ($\Delta f_I/\Delta f_N$) as a function of sensation level for the experimental data and the theoretical detectors. It appears that the count-based detector represents the relation between normal and impaired hearing better than the optimal or CRLB detectors which actually indicate that the impaired model performs better than the normal model ($\Delta f_I/\Delta f_N < 1$). It can be concluded that the impaired model is not an accurate representation of hearing impairment and requires further modification.

L [dBSL]	$\Delta f_I/\Delta f_N$			
	Expt.	**Optimal**	**CRLB**	**Count-Based**
10	1.9	0.5	0.6	2.4
20	1.9	0.5	0.5	2.1
30	1.9	0.6	0.5	2.0
40	1.9	0.3	0.7	1.7

Table 2. The experimental and theoretical ratios of the DLF in impaired and normal ears as a function of sensation level.

6. SUMMARY

Two different detectors were derived using signal detection theory to analyze information transmission in the auditory system. The difference in the detectors highlights the contribution temporal information may make in a frequency discrimination task. The results support the hypothesis that the human auditory system uses temporal information to perform this task but that a second mechanism dominates at higher frequencies. The combination of SDT and computational models has be shown to do a fairly good job at predicting trends in performance for most normal and some impaired-hearing cases. However, the comparison between theoretical and experimental predictions for normal and impaired cases suggests that the model used to simulate impairments needs to be further modified to improve the predictive power of this approach.

7. REFERENCES

[1] Carney, L.H. (**1993**). "A model for the responses of low-frequency auditory-nerve fibers in cats," J. Acoust. Soc. Am. **93**, 401-417.

[2] Cramèr, H. (**1951**). *Mathematical Methods of Statistics* (Princeton University Press, Princeton, NJ).

[3] Freyman, R. L. and Nelson, D. A. (**1991**). "Frequency discrimination as a function of signal frequency and level in normal-hearing and hearing-impaired listeners," J. Speech Hear. Res. **34**, 1371-1386.

[4] Harrison, J. M. and Howe, M. E. (**1974**). "Anatomy of the afferent auditory nervous system of mammals," in *Handbook of Sensory Physiology*, eds. W. D. Keidel and W. D. Neff, (Springer, Berlin).

[5] Heinz, M. G., Carney, L. H., and Colburn, H. S. (**1999**). "Performance limits for frequency and intensity discrimination based on a computational auditory-nerve model," Abstracts of the 22rd Midwinter Meeting of the Assoc. for Research in Otolaryngology, pp. 212.

[6] Huettel, L. G. and Collins, L. M. (**1999**). "Using computational auditory models to predict simultaneous masking data: Model comparison," IEEE Trans. Biomed. Eng. **46**, 1432-1440.

[7] Huettel, L. G. and Collins, L. M. (**2000**). "The theoretical effects of neural uncertainty on acoustic signal detectability: Single vs. multiple ANFs," Abstracts of the 23rd Midwinter Meeting of the Assoc. for Research in Otolaryngology, pp. 4.

[8] Moore, B. C. J. (**1973**). "Frequency difference limens for short-duration tones," J. Acoust. Soc. Am. **54**, 610-619.

[9] Moore, B. C. J. (**1989**). *An Introduction to the Psychology of Hearing* (Academic Press, San Diego, CA), pp 76-78.

[10] Patterson, R. D., Allerhand, M. H., and Giguère, C. (**1995**). "Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform," J. Acoust. Soc. Am. **98**, 1890-1894.

[11] Peterson, W. W., Birdsall, T. G., and Fox, W. C. (**1954**). "The theory of signal detectability," Trans. IRE **PGIT-4**, 171-212.

[12] Siebert, W. M. (**1968**). "Stimulus Transformations in the Peripheral Auditory System" in *Recognizing Patterns*, ed. Kolers and Eden, (MIT Press, Cambridge, MA).

[13] Siebert, W. M. (**1970**). "Frequency discrimination in the auditory system: Place or periodicity mechanisms?," IEEE Proc. **58**, 723-730.

[14] Simon, H.J. and Yund, E. W. (**1993**). "Frequency discrimination in listeners with sensorineural hearing loss," Ear and Hearing **14**, 190-201.

[15] Tanner, W. P., Jr., and Birdsall, T. G. (**1958**). "Definitions of d' and η as psychophysical measures," J. Acoust. Soc. Am. **32**, 1140-1147.

[16] Turner, C. W. (**1987**). "Effects of noise and hearing loss upon frequency discrimination," Audiology **26**, 133-140.