Evaluating Human Neural Tuning Curves From a Mechanical Model of the Cochlea by Relating Them to Psychophysical Masking Data

by

Peter A. Z. Garbes

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degrees of

Master of Science

and

Bachelor of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1994

Author .....
Department of Electrical Engineering and Computer Science
May 12, 1994

Certified by .....
Dennis M. Freeman
Thesis Supervisor (Academic)

Certified by .....
Jont B. Allen
Company Supervisor (AT&T Bell Laboratories)

Accepted by .....
Frederic R. Morgenthaler
Chairman, Departmental Committee on Graduate Students

Eng.

Evaluating Human Neural Tuning Curves From a Mechanical Model of the Cochlea by Relating Them to Psychophysical Masking Data

by

Peter A. Z. Garbes

Submitted to the Department of Electrical Engineering and Computer Science
on May 12, 1994, in partial fulfillment of the
requirements for the degrees of
Master of Science
and
Bachelor of Science

Abstract

Human neural threshold tuning curves are estimated by scaling the parameters of Allen's [1] resonant tectorial membrane model of the cat cochlea. A way to evaluate the derived tuning curves using psychophysical data is developed, based on a psychophysical detection model which relates the physiological tuning curves to psychophysical masking data. A detection criterion, defined by a relationship among the bandwidth of the frequency tuning curves, expressed as an equivalent rectangular bandwidth (ERB), the width of the excitation patterns, expressed as an equivalent rectangular spread (ERS), and the psychophysical critical ratio, is explored and verified using cat data. The detection criterion is then used to test the derived human curves by making predictions of psychophysical masking and comparing the predictions to experimental data. The detection model may also provide a deeper understanding of the frequency resolving properties of the cochlea.

Thesis Supervisor: Dennis M. Freeman
Title: Research Scientist, Research Laboratory of Electronics

Thesis Supervisor: Jont B. Allen
Title: Distinguished Member of Technical Staff, AT&T Bell Laboratories

Acknowledgments

First, I would like to thank my mentor, Jont Allen, for everything he has taught me and for giving me the great opportunity to do my thesis reserach with him at Bell Laboratories; my officemate at Bell Labs, Mark Sydorenko, for the encouragement when things got rough and for helping make my stay in New Jersey an enjoyable one: and my MIT supervisor, Denny Freeman, for pushing me through the writing stages and helping mold this thesis into a readable document.

I would also like to thank my friend Mike Ismert, who went through the thesis process last year, for reading through some of the preliminary drafts of my thesis and reassuring me that I would finish, especially when I was frustrated with all the work.

Finally, this work is dedicated to my parents, who have been behind me every step of the way, and whose love and support have helped me get through 5 years of MIT.

Contents

Chapter 1

Introduction

1.1 Background

Modeling the cochlea has been a goal of hearing research dating before Wegel and Lane's transmission line model in 1924 [2]. A model of the human cochlea would be useful in attempts to include auditory models in commercial products such as hearing aids and speech recognizers. In addition, it would be a valuable tool in further human auditory research. The biggest roadblock to modeling the human cochlea is the lack of available physiological data on live human cochleas due to the impracticality of performing the experiments on living humans. Thus, most cochlear models which have been quantitatively developed are tested on animal data. This thesis explores the problem of extrapolating results obtained for cat to human. Then, the problem of evaluating those results using psychophysical masking data is explored.

1.2 Overview

Four types of data are studied in this thesis (Fig. 1-1): psychophysical and physiological for the cat and psychophysical and physiological for the human. A relation (represented by the top horizontal arrow) between cat and human psychophysics was shown by Allen [3] and was the inspiration for this thesis. The idea is that the critical ratio, a measure of psychophysical masking, is approximately the same in both species

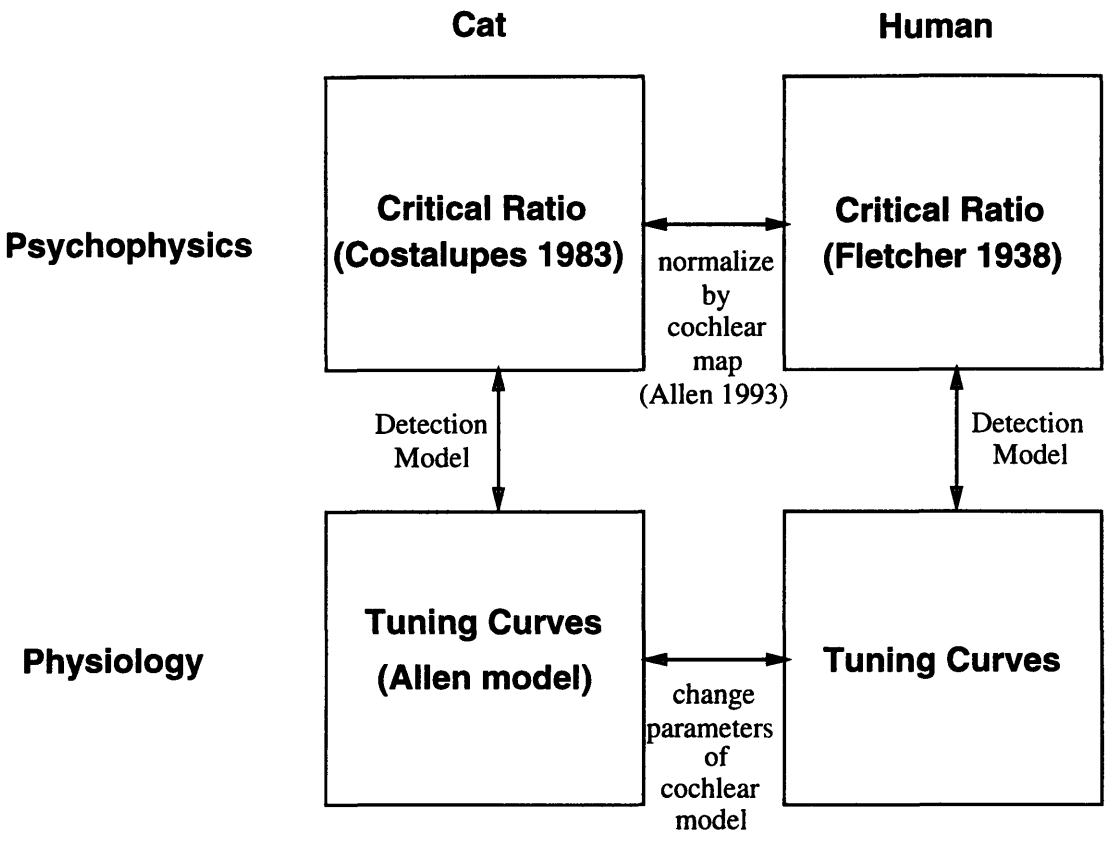

Figure 1-1: Diagram of relationships among data analyzed in this work.

when the data are normalized to the length of the basilar membrane (Fig. 1-2).

Transforming the cat cochlear model into a human cochlear model (represented by the bottom horizontal arrow in Fig. 1-1 is the subject of Chapter 2. Developing the relation between the physiology and psychophysics (represented by the vertical arrows in Fig. 1-1 is the subject of Chapter 3.

1.2.1 Deriving the Human Model

Chapter 2 deals with the derivation of the human cochlear model. A human cochlear model will be developed based on Allen's resonant tectorial membrane model [1]. The scope of this portion of the work rests on the assumption that the cat auditory system and the human auditory system work similarly. Thus, to the extent that the resonant TM model accurately simulates the cat's neural threshold tuning curves, then by appropriately scaling the parameters, the corresponding human model can

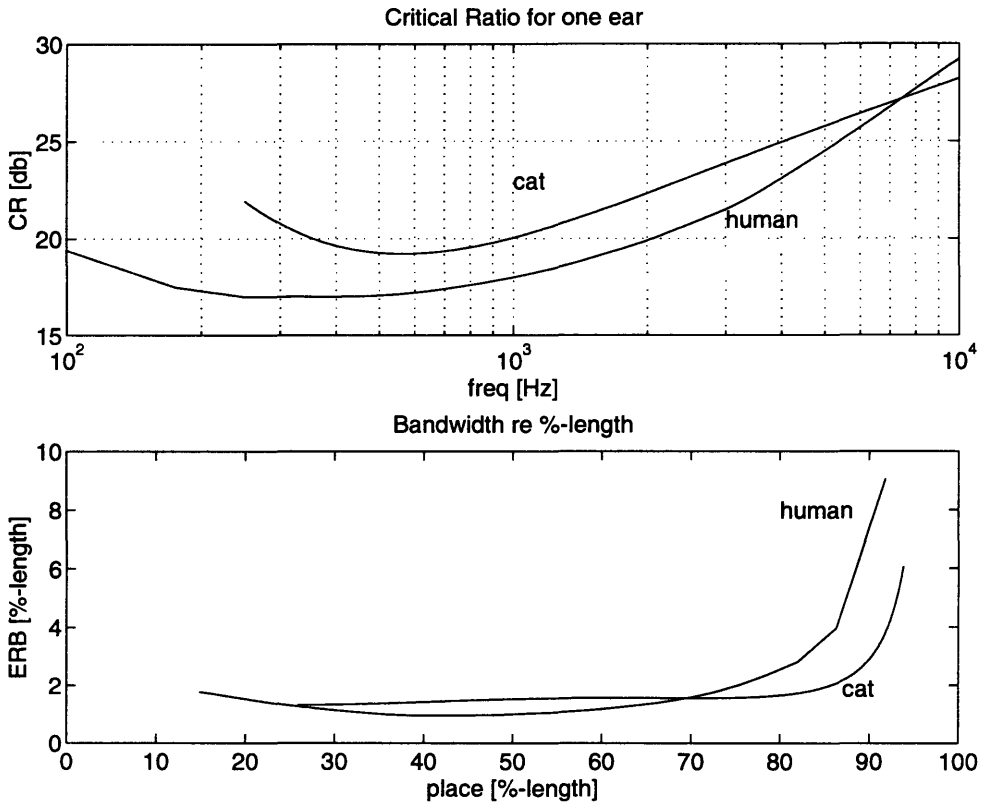

Figure 1-2: Comparison of cat and human critical ratios. The upper panel shows critical ratio data from Fletcher [4] (for human) and Costalupes [5] (for cat). The lower panel shows same data, normalized to percent length along the BM. Adapted from [3].

be developed.

1.2.2 Evaluating the Human Tuning Curves

Chapter 3 deals with evaluating the human tuning curves. The first step is to develop a relation between tuning curves and performance in a psychophysical task. The task is detection of tones in noise and the performance metric is critical ratio, i.e. the signal-to-noise ratio at which a tone in wide band noise is just detectable. The proposed relation between physiology and psychophysics can then be tested directly for cats because both the physiology and psychophysics have been measured. It can then be used to predict human psychophysical function from the derived human tuning curves. To the extent that the predictions match measured human psychophysical masking, the validity of the method for estimating human physiological tuning curves is supported.

Chapter 2

Deriving Human Tuning Curves

In this chapter, the shapes of human neural tuning curves are estimated from a physical model of the cochlea. The model is based on Allen's resonant tectorial membrane model [1], a passive model of the cochlea which includes a "second filter". This second filter is modeled as a resonance in the tectorial membrane that transduces a relatively broad basilar membrane response into a sharper neural response.

2.1 The Resonant Tectorial Membrane Model

The cochlea is assumed to be divided into two scalae and filled with an incompressible fluid. Sound waves enter the ear and travel through the ear canal to vibrate the tympanic membrane. The tympanic membrane then vibrates the ossicles, which act as an impedance matcher to transfer sound from the tympanic membrane to the smaller oval window without a large loss in energy. The last of the three ossicles is the stapes, whose motion displaces the oval window, which in turn displaces fluid in the cochlea. The fluid displacement causes a "traveling wave" on the basilar membrane. The movement of the basilar membrane causes hair cells to shear against the tectorial membrane and this shearing is the stimulus which sends impulses down the auditory nerve.

Although controversial, some measured neural responses have been shown to be more sharply frequency selective than the basilar membrane response [6, 7]. Allen's

model accounts for this by introducing a "second filter" or "transduction filter" as a resonance of the tectorial membrane. This second filter introduces a zero into the system at a frequency below the characteristic frequency, and thus sharpens the broader response of the basilar membrane alone. The transduction filter also accounts for the π phase shift found in the phase data of Kim *et al.* (see [1] for further information).

Finally, the cochlea is known to have non-linear properties. To account for this, Allen proposes that the stiffness of the basilar membrane varies with signal level for several reasons [1, 8]. In this thesis, however, we will deal with a linear model to avoid the difficulties brought on by nonlinearity. Further work incorporating the nonlinear properties of the cochlea is an important area for future investigation.

2.1.1 Parameters

The following is a list of parameters which are used in the computer simulation [9], divided into three groups.

Geometric Parameters

- L, length of the cochlea
- h, height of the organ of Corti
- w_{par}, width of the cochlear partition
- w_{bm}, width of the basilar membrane
- m_0, specific mass of the organ of Corti
- m_t, specific mass of the tectorial membrane

Tuning Parameters

- ω_{cf}, the cochlear map
- ω_z, spectral zero frequency of transduction filter

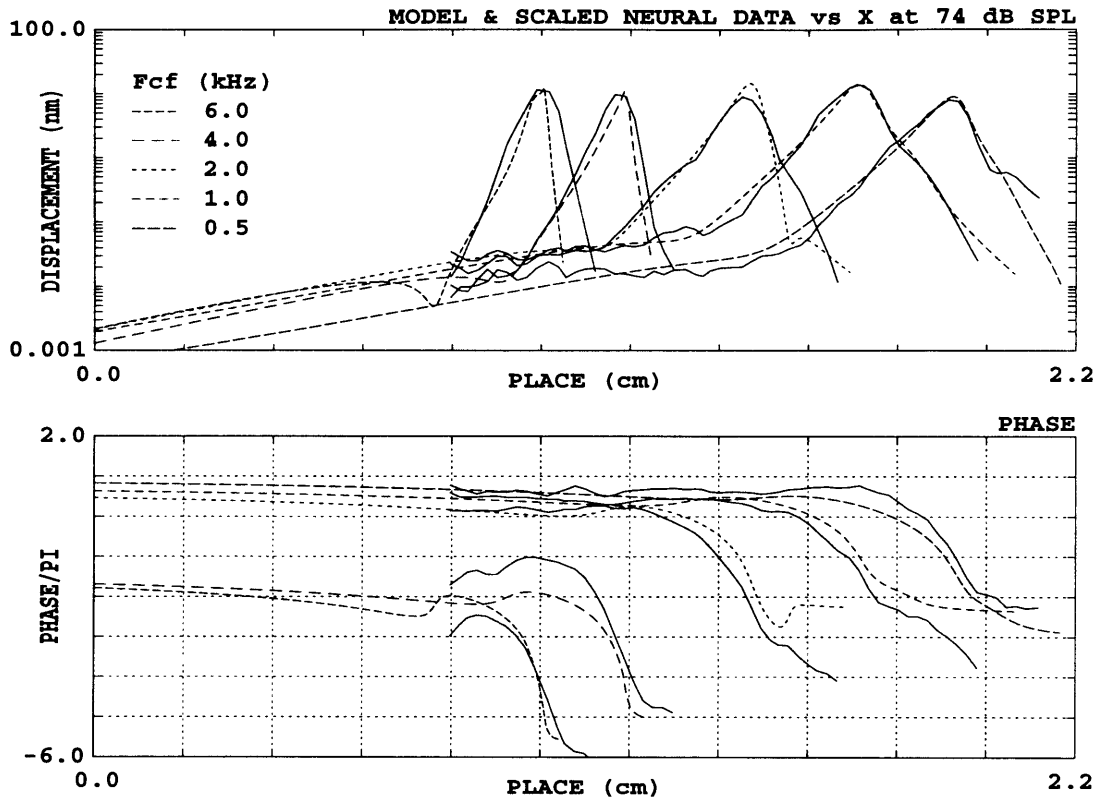

Figure 2-1: Cat cochlear model output. From Allen [8]. The dashed lines are the model data and the solid lines are scaled neural data.

- ω_p, spectral pole frequency of transduction filter

Damping and Gain Parameters

- ζ_z, damping ratio of transduction filter
- r_{bm}, basilar membrane resistance per unit area
- G, shear gain

The cat parameters were chosen by Allen [3] as a best fit to current physiological data available. A plot of excitation patterns calculated for the cat taken from Allen [8] is shown in Fig. 2-1. The computer simulation of the model suffers from a known artifact at low frequencies (below 300 Hz in the cat) due to the boundary condition at the helicotrema. In addition, undersampling of the data points in either frequency or space can also introduce small numerical errors. For these reasons, the data from

the model is best used in the frequency range from 700 to 10^4 Hz, where the model is close to neural measurements.

2.2 Scaling the Parameters

The model parameters are divided into three groups, according to the method used to modify them. First, the geometric parameters were scaled by the ratio of the lengths of the cat and human cochleas. Next, the cat tuning parameters were scaled by replacing the cat cochlear maps with human cochlear maps in order to rescale the frequency mapping of the model. Finally, the damping and gain parameters were left unchanged because there was no obvious reason to alter them.

2.3 Geometric Scaling

Assuming that the human and cat cochleas are essentially scale models of each other, the geometric parameters can be scaled by the ratio of the lengths of the cochleas.[1] The assumed length of the cat cochlea is 2.1 cm, and the assumed length of the human cochlea is 3.5 cm, so all length parameters were increased by a factor of 3.5/2.1, which is approximately a scale factor of 1.59.

Organ of Corti height and cochlear partition width

The height of the organ of Corti and the width of the cochlear partition are the simplest parameters to scale. In the cat, h has a value of 0.1 cm, so the human value is scaled up to 0.159 cm. Similarly, the cochlear partition width has a value of 0.16 cm in the cat, so the corresponding scaled value is 0.256 cm.

[1] Some of the parameters in this section have been measured in humans. An alternative approach would have been to use measured data where it exists.

Masses

The next items to consider are the mass parameters. These are defined in the model as specific mass, which is mass per area. These parameters were calculated by multiplying the height of the structure by the density of the structure [10], so they scale in only one dimension (height). For the cat, m_0 is given as 0.04 g/cm^2 and m_t is 0.02 g/cm^2. Thus for the human these values are 0.064 g/cm^2 and 0.032 g/cm^2, respectively.

Width of the Basilar Membrane

The width of the basilar membrane varies with position along the long axis of the cochlear partition. In this model, it is assumed that the width varies exponentially, and is described by the function

$$w_{bm}(x) = A \exp \frac{\alpha x}{L} \quad \text{(cm)}. \tag{2.1}$$

For the cat, A is equal to 0.011, and α is 1.19. For the human, the exponent was left unchanged so that the variation with normalized position would remain the same, and the coefficient A was changed to 0.0176 to account for the geometric scaling.

2.4 Frequency Scaling

Some other model elements are estimated from tuning parameters of the cochlea. The frequency spectrum of the model is then scaled in a manner that is similar to the previous geometric scaling. The model has three frequency parameters: the cochlear map, which is the function mapping the frequency to place of the maxima of the tuning curves; and the maps of the spectral zero and spectral pole of the transduction filter.

2.4.1 The Cochlear Map

The Cat Cochlear Map

The cochlear map for the cat was determined by Liberman [11] by measuring auditory nerve fibers' threshold tuning curves, and then labeling them with horseradish peroxidase (HRP). The HRP stains the fibers and is transported to the synapse at the hair cell. In this way, the exact location on the basilar membrane which is connected to an auditory nerve fiber can be seen on surface preparations. From this data, a frequency-to-place function was determined. If f_{CF} represents the characteristic frequency in Hz and x is the place on the basilar membrane expressed in terms of percent length from the stapes, Liberman's formula is

$$f_{CF} = 456(10^{2.1(1-\frac{x}{100})} - 0.8). \tag{2.2}$$

The Human Cochlear Map

One of the earliest attempts to characterize the cochlear map was by Wegel and Lane in 1924 [2] using just noticeable differences (JNDs) in frequency. Fletcher extended Wegel and Lane's frequency JNDs, and proposed using masking to determine a cochlear map function in 1938 [12]. Greenwood used the critical bandwidth concept to calculate the cochlear map in 1961 [13], and further revised his calculations of the human cochlear map in 1990 [14]. Greenwood's equation for the human cochlea, with f_{CF} representing characteristic frequency in Hz, and x being place expressed in percent length from the stapes along the basilar membrane, is

$$f_{CF} = 165.4(10^{2.1(1-\frac{x}{100})} - 0.88). \tag{2.3}$$

2.4.2 Distortion Products and the Second Cochlear Map

The zero location of the transduction filter is inferred from Allen and Fahey's second cochlear map [15].

Distortion Products

When two tones are presented simultaneously, other tones which are not part of the stimulus can be heard as a result of nonlinearities in the system. These tones are known as distortion products. They have been shown to propagate back to the outer ear where they can be measured [16, 17]. If the two stimulus tones are at frequency f_1 and f_2, with $f_2 > f_1$ and f_2 fixed, Allen and Fahey (and Wilson [18] and Brown [19]) have shown that as f_1 is varied, the frequency at which the distortion products are maximum is a function only of f_2 and is approximately independent of other variables such as amplitude and f_1.

The second cochlear map

Allen and Fahey [15] showed the existence of a second cochlear map using distortion product data. This second cochlear map is important because they suggest that it defines the zero location for the transduction filter.

The second cochlear map is the map of maximum amplitude DPs as a function of f_2. This equation matches the location where the tip meets the tail of the neural tuning curves, so is also the zero of the second filter, or f_z. Thus, the two curves are given by the same equation

$$f_z = 0.08 f_{CF}^{1.22}, \tag{2.4}$$

in the cat, where f_z is the zero location for the filter at characteristic frequency f_{CF}. Using distortion product data from human subjects, and reasoning that tnhe map derived from the DP data should also describe a second cochlear map in the human, Allen and Fahey arrive at

$$f_z = 0.5 f_{CF}^{1.04}, \tag{2.5}$$

as the equation for the second cochlear map in humans.

Thus, the zero location f_z for the human model is given by Eq. 2.5.

Parameter	Cat Value	Human Value	Units
L	2.2	3.5	cm
h	0.1	0.159	cm
w_{par}	0.16	0.256	cm
f_{cf}	$456(10^{2.1(1-\frac{x}{100})} - 0.8)$	$165.4(10^{2.1(1-\frac{x}{100})} - 0.88)$	Hz
f_z	$0.08 f_{cf}^{1.22}$	$0.5 f_{cf}^{1.04}$	Hz
f_p	$1.3 f_{cf}^{0.98}$	$1.15 f_{cf}^{0.98}$	Hz
ζ_z	$0.3845 e^{\frac{3.0x}{L}}$	$0.3845 e^{\frac{3.0x}{L}}$	
r_{bm}	$121.7 e^{\frac{-1.4x}{L}}$	$121.7 e^{\frac{-1.4x}{L}}$	g/cm^2·s
w_{bm}	$0.011 e^{\frac{1.19x}{L}}$	$0.0176 e^{\frac{1.19x}{L}}$	cm
G	$0.5 e^{\frac{3.0x}{L}}$	$0.5 e^{\frac{3.0x}{L}}$	
m_0	0.04	0.064	g/cm^2
m_t	0.02	0.032	g/cm^2

Table 2.1: Summary of Model Parameters

2.4.3 The Transduction Filter Pole

The final tuning parameter is the pole location of the transduction filter. Unfortunately, we have no intuition for finding a way to calculate it. The location of the pole was estimated by assuming the same functional form that was used for the zero. Experimenting with the model showed that changing the location of the pole had an effect on both the shapes of the simulated tuning curves and on the bandwidths of the curves. Therefore the pole location was chosen to provide reasonably shaped tuning curves based on the rather ad hoc criterion of visual aesthetics.

2.5 Results

A summary of the parameters is given in Table 2.1 for both the cat and the human models discussed in this chapter. Fig. 2-2 shows the human model tuning curves (actually transfer functions) based on these parameters and for comparison, tuning curves computed from the cat model are shown in Fig. 2-3. The human tuning curves essentially look like scaled versions of the cat tuning curves, with differences in breakpoints of the tails of the curves. These differences represent the constraints forced by the cochlear model based on the human geometric and tuning parameters.

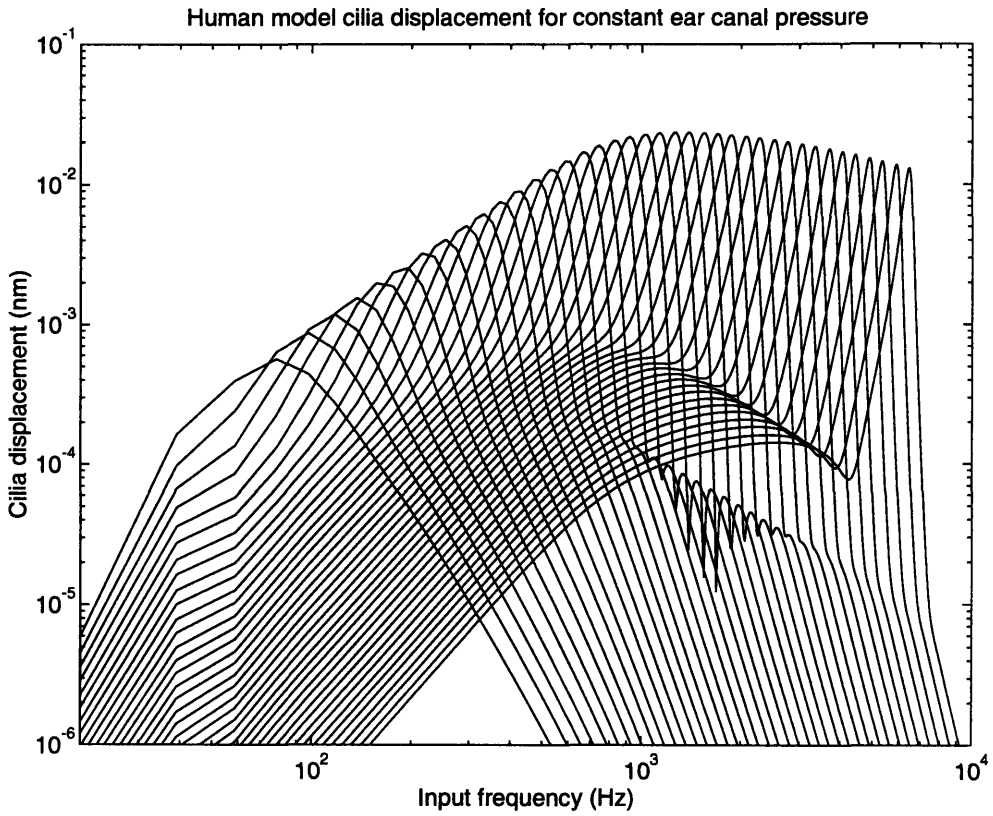

Figure 2-2: Human model transfer functions.

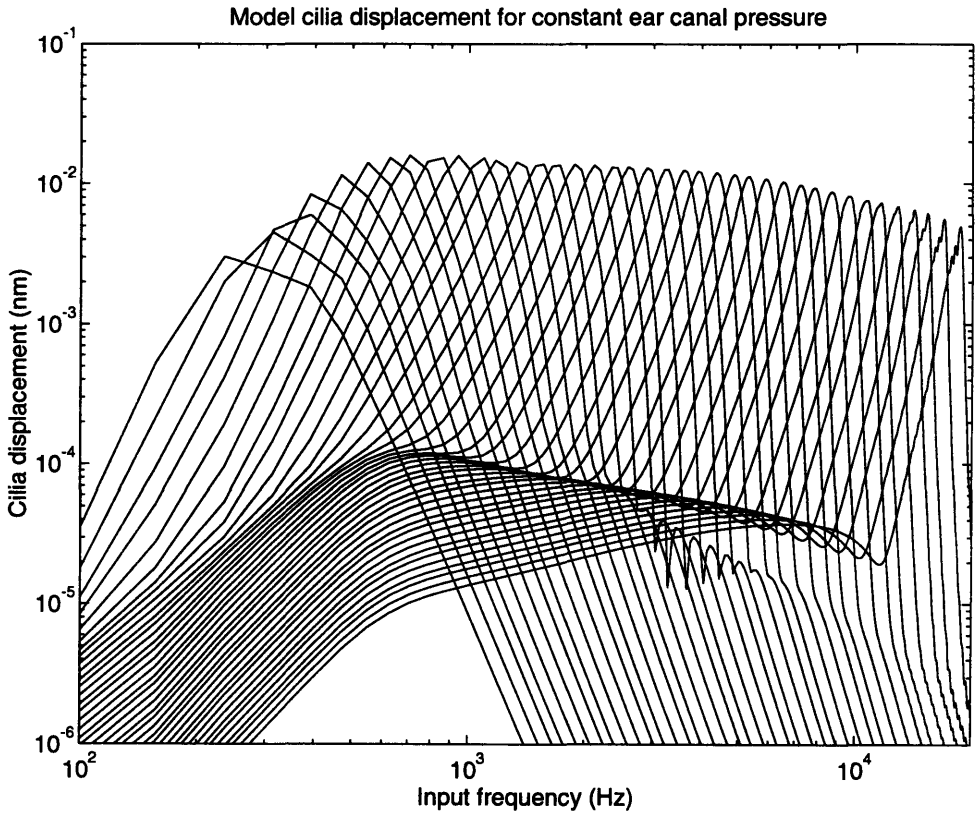

Figure 2-3: Cat model transfer functions.

Now that a human cochlear model has been formulated, Chapter 3 deals with evaluating tuning curves by using them to predict psychophysical masking.

Chapter 3

Predicting Psychophysical Performance from Physiological Models

In this chapter, two psychophysical models for detection of tones in noise are proposed and analyzed using cat data. Human psychophysical masking is then predicted using these psychophysical models.

3.1 Definitions

The *critical band* is the basis for the psychophysical detection models that will be presented here. It is a concept which does not have a consistent definition in the literature and is therefore a common source of confusion. It is known that the masking of a pure tone spreads over a range of frequencies surrounding the frequency of the tone, and from this idea, Harvey Fletcher formed the original concept of the critical band [20, 12, 21]. The following are definitions of the critical band measures and assumptions which are used in this work.

3.1.1 The Critical Ratio

The psychophysical performance metric used in the detection models in this chapter is the critical ratio. The critical ratio is derived from an experiment in which a tone at frequency f is presented to a subject in the presence of wide-band masking noise. The tone is raised in level until it is just detectable, and this signal level is called the detection threshold. The critical ratio is defined as the ratio of tone probe power to masker spectral level at the detection threshold. If the tone power is expressed as $T^2(f)$ (Watts) and the noise spectral level as $N^2(f)$ (Watts/Hz), the critical ratio κ at frequency f is given by

$$\kappa(f) = \frac{T^2(f)}{N^2(f)} \quad \text{(Hz)}. \tag{3.1}$$

Because the units of critical ratio are Hz, it can be considered a critical bandwidth. The critical ratio is frequently reported in dB Hz (e.g. $10 \log \kappa(f)$), as it is computed from a ratio of power measurements. Thus, the critical ratio is a psychophysical measure of critical bands.

3.1.2 Physiological measures of critical bands

Using either the model or measurements of neural threshold tuning curves, the cochlear response $H(f, x)$, which expresses the cochlear filters, can be calculated or measured.

A frequency tuning curve (FTC) is the frequency response of a single point on the basilar membrane. The frequency at which a FTC is maximum is called its characteristic frequency (CF).

On the other hand, the basilar membrane response to a single frequency is called an excitation pattern (EP). The place at which an excitation pattern is maximum is called its best place.

In terms of the cochlear response, the cochlear map is the function that maps each place to its characteristic frequency. This function will be denoted as

$$f_{CF} = F(x). \tag{3.2}$$

The cochlear map may also be thought of as the function that maps each frequency to its best place, and this function will be denoted as

$$x_{BP} = X(f). \tag{3.3}$$

The ERB. The equivalent rectangular bandwidth (ERB) for a FTC at place x is defined as the bandwidth of the rectangular filter with the same amplitude as the peak amplitude of the FTC and bandwidth chosen so its output power is equal to that of the FTC when driven with white noise. The ERB is thus given by

$$\Delta_f(x) = \frac{\int_0^\infty |H(f,x)|^2 df}{|H(F(x),x)|^2} \quad \text{(Hz)}. \tag{3.4}$$

The ERS. The equivalent rectangular spread (ERS) is defined (analogous to ERB) as the width of the excitation pattern for an input frequency f. The ERS is given by

$$\Delta_x(f) = \frac{\int_0^L |H(f,x)|^2 dx}{|H(f,X(f))|^2} \quad \text{(m)}. \tag{3.5}$$

3.2 First Psychophysical Detection Model

The first psychophysical detection model is based on two premises: first, that psychophysical performance for detection of a tone at frequency f in noise is determined by the filter with characteristic frequency equal to f, and second, that the signal-to-noise ratio for that cochlear filter is constant at the detection threshold.

3.2.1 Effect of Cochlear Filters on Signal and Noise

Tone excitation. If a tone with input power $T^2(f)$ at frequency f is presented to one of these filters at its characteristic frequency, then at place x the signal power after filtering will be equal to

$$\sigma_T^2(f,x) = T^2(f)|H(f,x)|^2 \quad \text{(W)}. \tag{3.6}$$

Noise excitation. On the other hand, if the stimulus is wide band noise with constant spectral level $N^2(f)$, the power at place x after filtering will be

$$\sigma_N^2(x) = N^2 \int_0^\infty |H(f,x)|^2 df \quad \text{(W)}. \tag{3.7}$$

3.2.2 Detection Criterion

The probe to masker ratio (PMR) $c(f,x)$ at the detector is then given by the ratio

$$c(f,x) = \frac{\sigma_T^2(f,x)}{\sigma_N^2(x)}. \tag{3.8}$$

Combining this with Eqs. 3.6 and 3.7,

$$c(f,x) = \frac{T^2|H(f,x)|^2}{N^2 \int_0^\infty |H(f,x)|^2 df}. \tag{3.9}$$

From the definition of κ, Eq. 3.1, and of Δ_f, Eq. 3.4, we get

$$c(f,x) = \frac{\kappa(f)}{\Delta_f(x)}. \tag{3.10}$$

The noise that reaches the detector is the noise spectral level integrated through the cochlear filter, or equivalently, the noise spectral level multiplied by the ERB of the filter. The *detection criterion* is defined as this PMR at the place of maximum excitation. The critical ratio, $\kappa(f)$, is a psychophysical measure, while the ERB, $\Delta_f(x)$, is a physiological measure, so Eq. 3.10 summarizes a relationship between the physiology and the psychophysics.

3.3 Cat Psychophysical and Physiological Data

The psychophysical detection model was analyzed using the cat data described in this section.

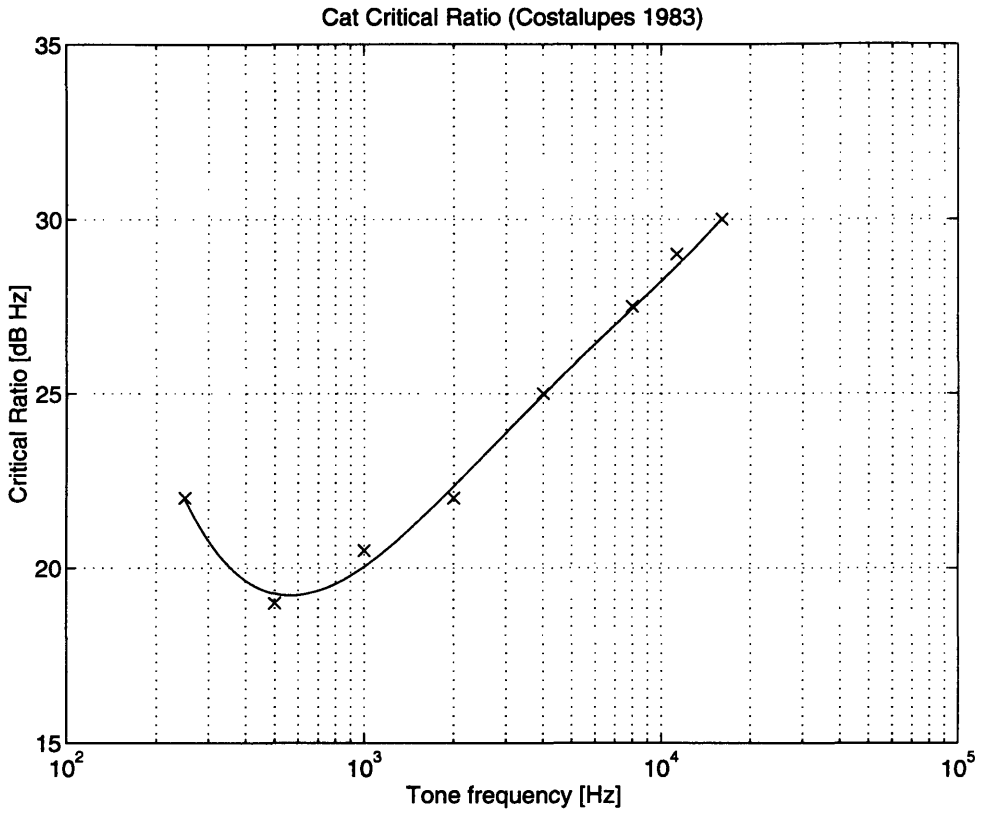

Figure 3-1: Costalupes' critical ratio data. The original data points are indicated with x's. Data were fit with a 4th-order least squares polynomial.

3.3.1 Critical Ratio

Critical ratios were measured in the cat by Costalupes [5]. The experiment consisted of presenting a tone in wide-band noise, and testing if the cat detected the presence of the tone. The cats were trained to initiate trials by touching a panel and releasing when the signal was detected, at which point the cat was rewarded with food. To reduce false alarms, cats were punished for releasing the panel when there was no stimulus by a time-out in the experiment. The threshold was taken as the 50% detection level.

The critical ratio was measured by Costalupes at many different frequencies and different intensity levels. Although he discovered level dependence in the critical ratio at high and low intensities, the data presented here (Fig. 3-1) are from the middle intensity range, where the critical ratio is relatively independent of intensity.

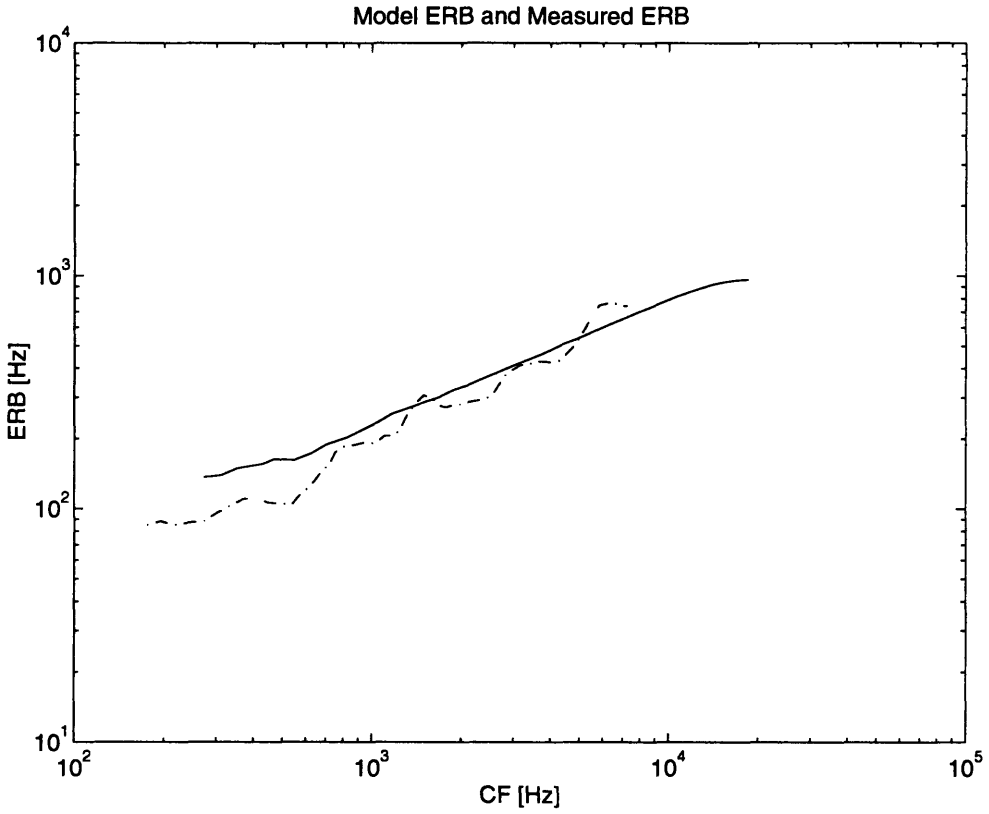

Figure 3-2: Equivalent rectangular bandwidth. The model ERB is shown in the solid line. The measured ERBs are shown in the dash-dot line for comparison.

3.3.2 ERB and ERS

We calculated the cat ERB both from neural threshold tuning curves and from the model described in chapter 2. The results are shown in Fig. 3-2. Subsequent analysis is based on the ERB calculated from the model, because it is smoother.

Although the ERS could in principle be calculated from neural data, this calculation requires combining data from many different neurons. Such a procedure is much more difficult and probably less accurate than the corresponding calculation of the ERB. Therefore the ERS was computed only from the model. The results are shown in Fig. 3-3.

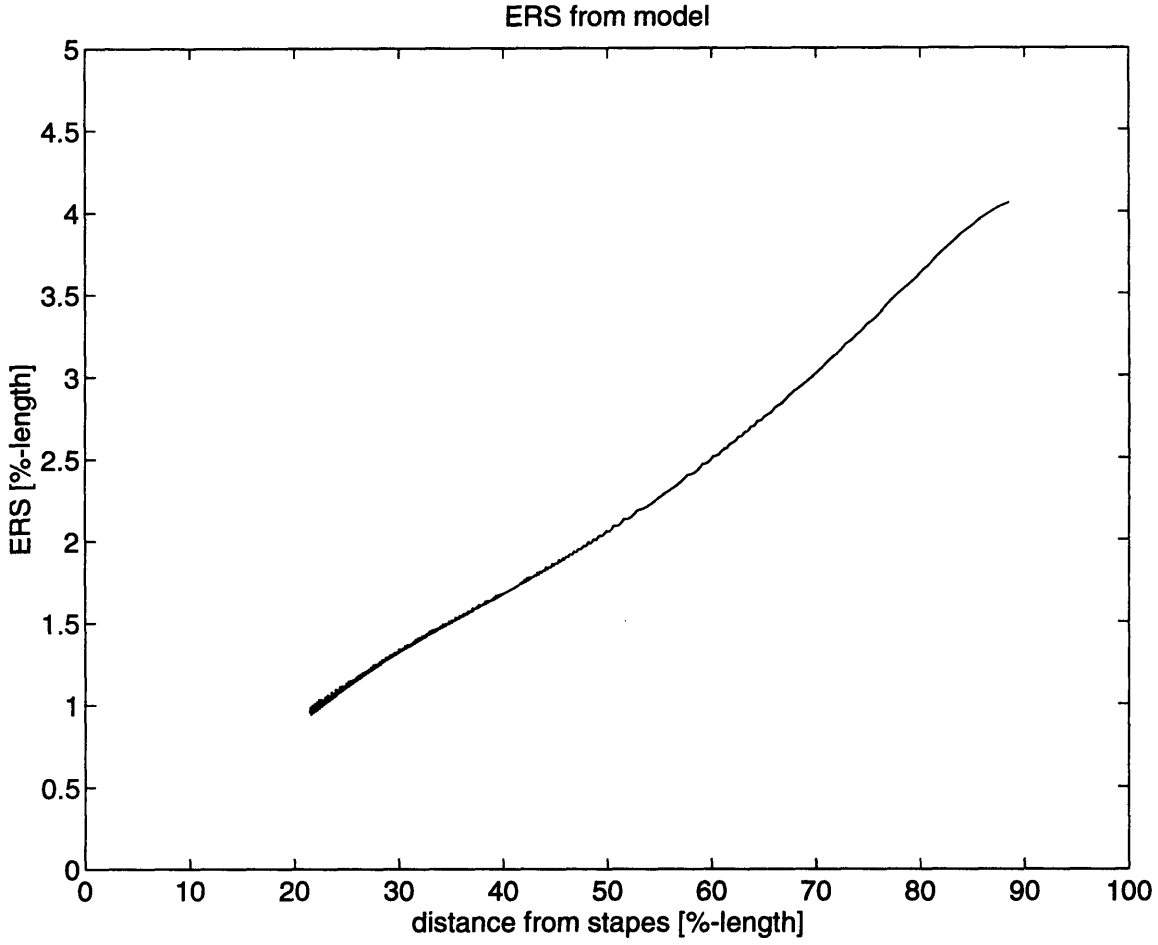

Figure 3-3: Equivalent rectangular spread calculated from the model.

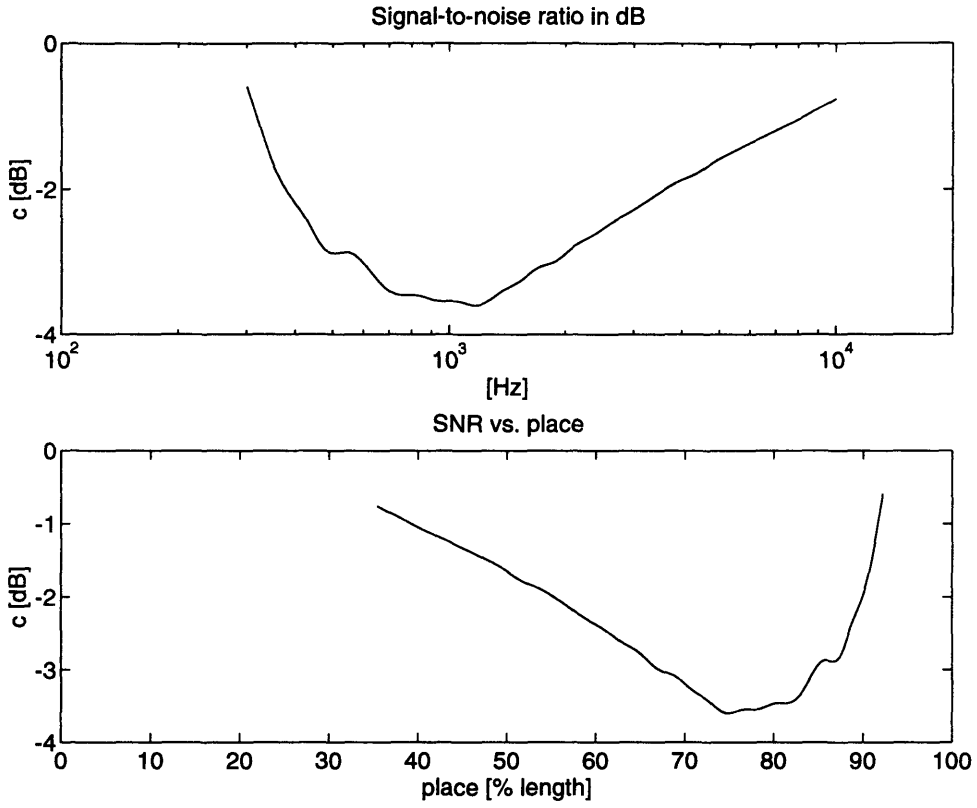

Figure 3-4: Signal-to-noise (Probe-to-masker) ratio at the detector output. The two panels shows c as function of either CF or place.

3.4 Evaluation of First Model

Fig. 3-4 shows c calculated from Eq. 3.10 using κ from Fig. 3-1 and Δ_f from Fig. 3-2. From the plots, $c(f, x)$ varies by about 3 dB along the length of the basilar membrane, with a minimum at the 1 KHz place (75% distance from stapes).

According to the first model, $c(f, x)$ should be constant. Therefore the original model is wrong, and it is not true that psychophysical detection occurs for a constant signal-to-noise ratio in the best cochlear filter.

Besides signal power, other cues may possibly be exploited psychophysically. For example, *salience* would suggest that the tone should be less detectable when the noise "sounds" like a tone. Using this argument, it would be expected that the signal-to-noise ratio at the detection threshold should decrease with increasing bandwidth because the noise which passes through a broad filter is less like a tone than is noise which passes through a narrow filter. This may explain the increase in c at low

frequencies in Fig. 3-4 but not at high frequencies.

On the other hand, *synchrony* is also a possible explaination. Above 1 KHz, synchrony decreases due to the low pass effect of the hair-cell and synapse dynamics. [22]. However, this argument fails to explain the upswing below 1 KHz, as it would predict the curve to be flat there.

While each of these cues explains some of the variation in signal-to-noise ratio, neither fully explains the variations. In addition, the effects of these cues is difficult to explain quantitatively.

3.5 Fletcher's Detection Model

An alternative detection model was derived from the work of Harvey Fletcher. This model assumes that the psychophysical detection threshold results when the signal to noise-per-length ratio is constant. From a derivation by Allen [23], the signal to noise-per-length ratio is expressed by

$$\begin{aligned} C(f) &= \int_0^L c(x,f)dx && (3.11)\\ &\approx c \cdot \Delta_x && (3.12)\\ &\approx \kappa \frac{\Delta_x}{\Delta_f} \ \text{(cm)}. && (3.13) \end{aligned}$$

This equation differs from Eq. 3.10 only by the Δ_x term.

The interpretation of C is that the brain integrates signals over a patch of neurons to detect a tone in noise rather than a just looking at the signal from a single neuron as in the first detection model. This is significant because with a hair cell spacing of 12 μm, one ERS generally covers a region containing about 40 inner hair cells. Thus the Fletcher detection model considers all of the information present from the tone input in the cochlea.

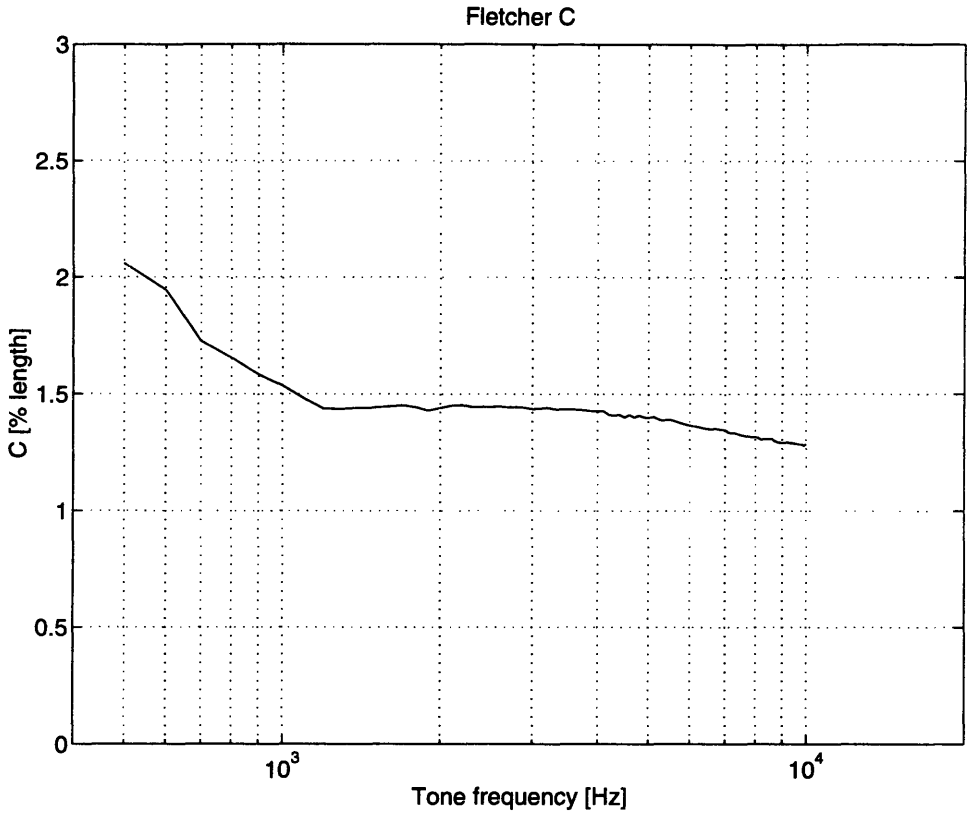

Figure 3-5: Signal to noise-per-length ratio vs. tone frequency at the detector output.

3.6 Evaluation of Fletcher Model

Since we have data for the terms in Eq. 3.12, we can verify Fletcher's assumption that C is constant. Fig. 3-5 shows C calculated from the ERB in Fig. 3-2, the ERS in Fig. 3-3, and the critical ratio in Fig. 3-1. In the middle of the frequency range, from about 10^3 Hz to 10^4 Hz, C seems to be about 1.4 to 1.5% of the length of the BM.

However, the model fails at low frequencies. This failure may be due to other cues being used at low frequencies, errors in the cochlear model at low frequencies as discussed in Section 2.1.1, or something else entirely.

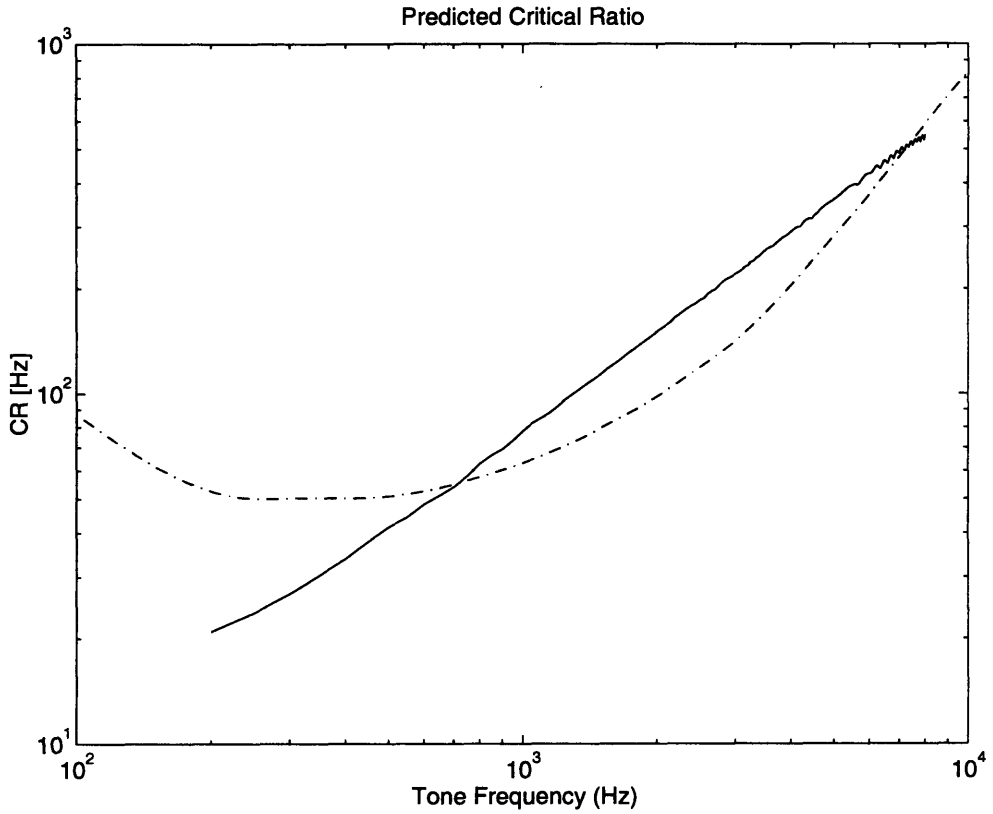

Figure 3-6: Critical ratio predicted using Fletcher detection model and constant C [solid] and measured critical ratio [dash-dot].

3.7 Predictions of Human Critical Ratios

Human critical ratios were predicted using the cochlear model tuning curves and the two detection models discussed in this chapter. Fig. 3-6 shows the prediction made using the Fletcher detection model, assuming a constant for the the detection criterion, C. Fig. 3-7 shows the prediction made using the first model, assuming that c is the same in humans as a function of percent length from the stapes. Both predictions are reasonably close to the measured critical ratio data from Fletcher [21, 4] above 1 kHz, but seem to predict values below the actual critical ratio at low frequencies.

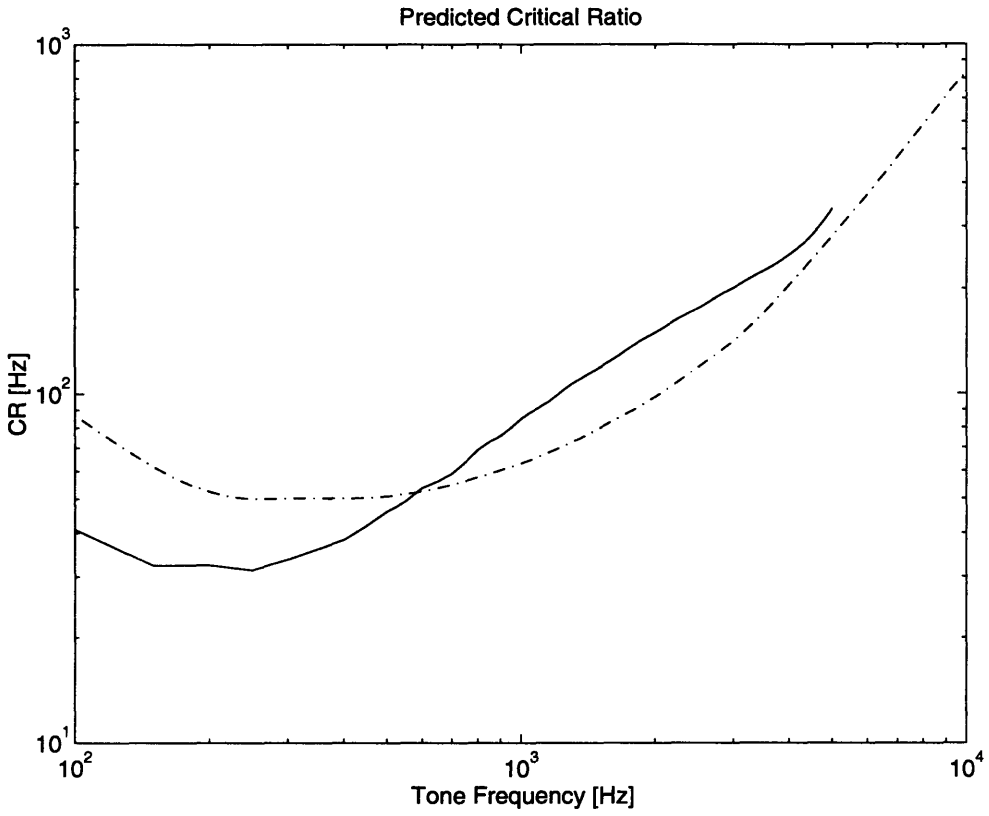

Figure 3-7: Critical Ratio predicted using first detection model [solid] and measured critical ratio [dash-dot].

3.8 Discussion

3.8.1 Validity of Using the Fletcher Detection Model to Test Tuning Curves

It turns out that the Fletcher detection model is rather insensitive to parameter variation in the mechanical model, shown in the following derivation by Goldstein [24].

Begin with the frequency tuning curve $H(f, x_0)$, where x_0 is the characteristic place of a tuning curve, and f is the frequency of the input tone. Also define the cochlear map as

$$f = F(x), \tag{3.14}$$

and its "inverse" as

$$x = X(f). \tag{3.15}$$

An excitation pattern $H(f_0, x)$ is related to a frequency tuning curve by

$$H(f_0, x) = H(F(x), X(f_0)). \tag{3.16}$$

Define the area under the excitation pattern as

$$I(f_0) = \int_0^L |H(f_0, x)|^2 dx. \tag{3.17}$$

Using Eq. 3.16, this becomes

$$I(f_0) = \int_0^L |H(F(x), X(f_0))|^2 dx. \tag{3.18}$$

Let $\alpha = F(x)$, then

$$I(f_0) = \int_0^L |H(\alpha, X(f_0))|^2 \left|\frac{dx}{d\alpha}\right| d\alpha, \tag{3.19}$$

and because of the narrow-band nature of H, $dx/d\alpha$ can be evaluated at f_0 and

brought outside of the integral, giving

$$I(f_0) = \left|\frac{dx}{d\alpha}\right|_{f_0} \int_0^\infty |H(\alpha, X(f_0))|^2 d\alpha. \tag{3.20}$$

$$= \left(\frac{dx}{d\alpha}\right) \cdot (\text{area under tuning curve at } X(f_0)). \tag{3.21}$$

Let $x_0 = X(f_0)$. Substituting the original definition for I from Eq. 3.17, replacing $F(x)$ back in for α, and rearranging, this becomes

$$\left|\frac{dF}{dx}\right|_{f_0} = \frac{\int_0^\infty |H(f, x_0)|^2 df}{\int_0^L |H(f_0, x)|^2 dx}. \tag{3.22}$$

The numerator and denominator can both be normalized to their peak values, since they have the same peak value of $|H(f_0, x_0)|^2$, and using the definitions of the ERB in Eq. 3.4 and the ERS in Eq. 3.5,

$$\left|\frac{dF}{dx}\right|_{f_0} = \frac{\Delta_f(x_0)}{\Delta_x(f_0)}. \tag{3.23}$$

Thus, combining this equation with equation 3.13,

$$\left|\frac{dF}{dx}\right| = \frac{\kappa}{C}. \tag{3.24}$$

Since the critical ratio κ is given, Eq. 3.24 shows that in the Fletcher detection model, the detection criterion C is only sensitive to the slope of the cochlear map, which is actually a model parameter. Experimenting with the cochlear model by wildly varying the other parameters confirmed that this was indeed the case. It appears that this detection criterion is an extremely insensitive test of the cochlear model.

3.8.2 The Cochlear Map

The discussion in the previous section shows that the cochlear map may be an important piece of data pertaining to relating the critical ratio to tuning curves. The cochlear map is also important in transforming the cat cochlear model into a human

model.

The parameter of most interest in Greenwood's function is the exponent, which is 2.1 in both the cat and the human. This parameter is responsible for the slope of the map. Since this was shown to be significant in Fletcher's detection model, a different exponent may invalidate applying this model to humans. In addition, if the slope of the cochlear map is not in fact the same in both the cat and the human, the comparison of the critical ratios originally done in Fig. 1-2 will not hold.

However, the slope of the cochlear map may be determined by the critical ratio if we consider the following. Re-express Greenwood's (or Liberman's) function in the following form, where L is the length of the cochlea:

$$F(x) = Ae^{x/L} - b. \tag{3.25}$$

Rearranging, this becomes

$$F(x) + b = Ae^{x/L}. \tag{3.26}$$

Taking the derivative,

$$F'(x) = \frac{A}{L}e^{x/L}. \tag{3.27}$$

Substituting this back into Eq. 3.25,

$$F(x) = LF'(x) - b, \tag{3.28}$$

and at high frequencies, the constant term becomes negligible leaving us with

$$F(x) \approx LF'(x). \tag{3.29}$$

Finally, using Eq. 3.24, we see that at in the high frequency range,

$$F(x) = \frac{L\kappa}{C}. \tag{3.30}$$

Thus at high frequencies, the cochlear map becomes proportional to the critical ratio.

3.9 Conclusion

While both detection models, when combined with the tuning curves derived in chapter 2, make predictions of masking which are reasonably close to experimental psychophysical data, neither detection model was found to be a rigorous test of the validity of the tuning curves. The Fletcher detection model is insensitivite to variations in tuning curve characteristics, and the first detection model is problematic because its detection criterion is not constant and it is likely that more information is taken into account than from just one cochlear filter.

Chapter 4

Conclusion

In this thesis, human neural tuning curves were derived from a model of the cochlea, and evaluated using psychophysical masking data. The following sections summarize the results from this work.

4.1 The Human Cochlear Model

There were three types of parameters which were considered when converting the cat model of the cochlea into a human model: geometric, tuning, and damping.

Geometric parameters were the simplest to convert, because they were simply scaled from the cat parameters. However, it would be better to use actual measured anatomical data.

The tuning parameters are used to estimate other mechanical element values. More convincing tuning data would lead to more confidence in using the cochlear map. Alternatively, removing the tuning data as inputs to the model and using raw mechanical data would also be beneficial.

Finally, the damping parameters were not changed when converting the model. More knowledge of the damping parameters are needed in order to analyze their effects on the tuning curves predicted by the cochlear model.

The human model for the cochlea presented here shows reasonable tuning curves, and is only marginally different from cat tuning data if the one considers normalizing

to length along the cochlea and normalizing to maximum frequency of hearing for each species.

4.2 The Detection Models

Two detection models were presented in this thesis which relate tuning curve data to psychophysical masking, specifically, the critical ratio experiment.

The first model related the signal-to-noise ratio at the neuron which maximally responded to the frequency of the tone to the critical ratio. This model was evaluated using actual cat data and it was found that the signal to noise ratio was not constant across the range of audible frequencies in the cat.

In an effort to explain this variation, a second model, which has been called the Fletcher detection model, was developed. The detection criterion in this model represented integration of signal-to-noise ratio over a patch of neurons. This quantity was assessed and found to be a constant over a range of frequencies in the cat. Unfortunately, it was shown to be insensitive to most model parameters and therefore appears to be not useful for testing human tuning curves.

4.3 Future Work

4.3.1 Physiological Basis of the Critical Band

A fundamentally important question which needs to be resolved is the physiological basis for the bandwidths of the tuning curves. None of the work done here showed what determines the value of the ERB and the ERS. In order to better formulate a way to test derived human tuning curves, it would be desirable to identify what determines the ERB and the ERS in the cochlea.

4.3.2 Other Psychophysical Data

The critical ratio is just one of the many pieces of psychophysical data available for the human. In this paper, the critical ratio was used because of the signal detection model proposed to test the derived human tuning curves.

However, the derived human tuning curve data should also be compared to other human psychophysical data, including psychophysical tuning curves and critical band measures other than critical ratio. Psychophysical tuning curves would be a good test of curves from a mechanical model, since they can be direcly compared to tuning curves predicted by a cochlear model.

For using other critical band measures, theoretical relationships between the psychophysics and the physical characteristics of the cochlea must be developed before predictions about these critical bands can be made from the model, and these relationships could be far more complicated than those presented here for the critical ratio.

4.3.3 Nonlinear Modeling

Since the analysis of the tuning curves involved computing bandwidths of tuning curves, a major problem which must be taken into account is the compressive nonlinearity of the cochlea. At higher sound levels, the tip of the tuning curve is less sharp. This would then affect the ERB and ERS calculated from those tuning curves, and could possibly change the conclusions of the previous chapter. Further work therefore awaits a nonlinear representation of the resonant tectorial membrane model.

4.4 Finale

The work done on the human tuning curves can be used as a framework for future research in modeling the human cochlea. Of greater interest in the long run might be the research into the detection models. Fletcher's detection model has been analyzed using real cat data and so far holds promise as an explanation of the critical ratio.

Bibliography

[1] J. B. Allen. Cochlear micromechanics–a physical model of transduction. *Journal of the Acoustical Society of America*, 68(6):1660–1670, December 1980.

[2] R. L. Wegel and C. E. Lane. The auditory masking of one pure tone by another and its probably relation to the dynamics of the inner ear. *Physical Review*, 23:266–285, 1924.

[3] J. B. Allen. Unpublished, 1993.

[4] Harvey Fletcher. *Speech and Hearing in Communication.* D. Van Nostrand Company, Inc., Princeton, NJ, 1953.

[5] John A. Costalupes. Broadband masking noise and behavioral pure tone thresholds in cats. *Journal of the Acoustical Society of America*, 74(3):758–764, September 1983.

[6] C. Daniel Geisler, William S. Rhode, and Duncan T. Kennedy. Responses to tonal stimuli of single auditory nerve fibers and their relation to basilar membrane motion in squirrel monkey. *Journal of Neurophysiology*, 37(6):1156–1172, November 1974.

[7] E. F. Evans and P. J. Wilson. Cochlear tuning properties: Concurrent basilar membrane and single nerve fiber measurement. *Science*, 190:1218–1221, December 1975.

[8] J. B. Allen. Modeling the noise damaged cochlea. In P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele, editors, *The Mechanics and Biophysics of Hearing*, pages 324–332. Springer-Verlag, New York, 1990.

[9] J. B. Allen. Unpublished.

[10] Sunil Puria and Jont B. Allen. A parametric study of cochlear input impedance. *Journal of the Acoustical Society of America*, 89(1):287–309, January 1991.

[11] M. Charles Liberman. The cochlear frequency map for the cat: Labeling auditory-nerve fibers of known characteristic frequency. *Journal of the Acoustical Society of America*, 72(5):1441–1449, November 1982.

[12] Harvey Fletcher. Loudness, masking and their relation to the hearing process and the problem of noise measurement. *Journal of the Acoustical Society of America*, 9(4):275–293, April 1938.

[13] Donald D. Greenwood. Critical bandwidth and the frequency coordinates of the basilar membrane. *Journal of the Acoustical Society of America*, 33(10):1344–1356, October 1961.

[14] Donald D. Greenwood. A cochlear frequency-position function for several species–29 years later. *Journal of the Acoustical Society of America*, 87(6):2592–2605, June 1990.

[15] J. B. Allen and P. F. Fahey. A second cochlear-frequency map that correlates distortion product and neural tuning measurements. *Journal of the Acoustical Society of America*, 94(2):809–816, August 1993.

[16] J. L. Hall. Two-tone distortion products in a nonlinear model of the basilar membrane. *Journal of the Acoustical Society of America*, 56(6):1818–1828, December 1974.

[17] D. O. Kim, C. E. Molnar, and J. W. Matthews. Cochlear mechanics: Nonlinear behavior in two-tone responses as reflected in cochlear-nerve-fiber responses

and in ear-canal sound pressure. *Journal of the Acoustical Society of America*, 67(5):1704–1721, May 1980.

[18] J. P. Wilson. The combination tone, $2f_1 - f_2$, in psychophysics and ear-canal recording. In G. van den Brink and F. A. Bilsen, editors, *Psychophysical, Physiological, and Behavioural Studies in Hearing*, pages 43–50. Delft Univ. Press, Delft, the Netherlands, 1980.

[19] A. M. Brown and S. A. Gaskill. Can basilar membrane tuning be inferred from distortion measurement? In P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C. R. Steele, editors, *The Mechanics and Biophysics of Hearing*, pages 164–169. Springer-Verlag, New York, 1990.

[20] Harvey Fletcher and W. A. Munson. Relation between loudness and masking. *Journal of the Acoustical Society of America*, 9(1):1–10, July 1937.

[21] Harvey Fletcher. The mechanism of hearing as revealed through experiment on the masking effect of thermal noise. *Proceedings National Academy of Science*, 24:265–274, 1938.

[22] J. B. Allen. A hair cell model of neural response. In E. deBoer and M. A. Viergever, editors, *Mechanics of Hearing*, pages 193–202. Delft Univ. Press, Delft, the Netherlands, 1983.

[23] J. B. Allen. Harvey Fletcher. Preface to reprinting of Fletcher's 1953 book, *Speech and Hearing in Communication*, 1993.

[24] Julius Goldstein. Unpublished, 1993.