SCAN WITH E-MAIL AND ODYSSEY REQUESTS

University of Hawai'i at Mānoa **LIBRARIES**

External Services Program

Hamilton Library Room 101 • 2550 McCarthy Mall • Honolulu, HI 96822

Phone/Voicemail: 808-956-5956

Fax: 808-956-7109

Email: esp@hawaii.edu

NOTICE: The copyright law of the United Sates (Title 17, U.S.C.) governs the making of photocopies or other reproductions of copyrighted materials. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than in private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order, if, in its judgment, fulfillment of the order would involve violation of copyright law.

3834

UNIVERSITY OF HAWAII LIBRARY

AUDITORY FILTER SHAPES IN AN ATLANTIC BOTTLENOSE DOLPHIN

(TURSIOPS TRUNCATUS)

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI'I IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

PSYCHOLOGY

AUGUST 1999

By
David Westlee Lemonds

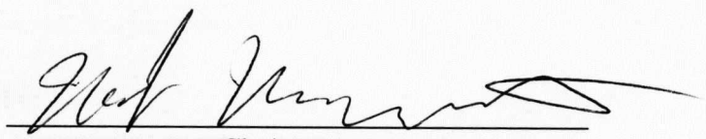
Dissertation Committee:

Herbert L. Roitblat, Chairperson
Whitlow W.L. Au
Daniel D. Blaine
Karl Minke
Paul E. Nachtigall

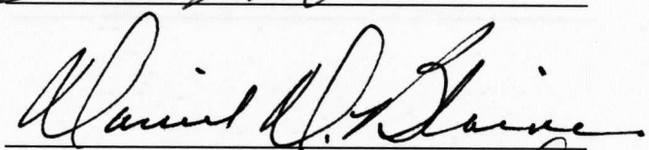
3834

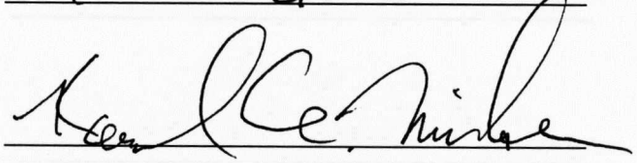
We certify that we have read this dissertation and that, in our opinion, it is satisfactory in scope and quality as a dissertation for the degree of Doctor of Philosophy in Psychology.

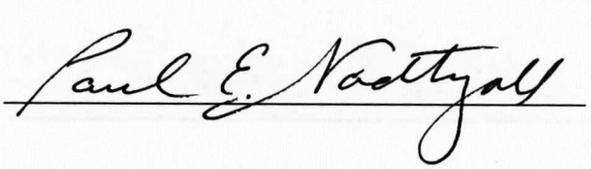
DISSERTATION COMMITTEE



Chairperson







We certify that we have read this dissertation and that, in our opinion, it is satisfactory in scope and quality as a dissertation for the degree of Doctor of Philosophy in Psychology.

DISSERTATION COMMITTEE

_____

Chairperson

_____

_____

_____

_____

ACKNOWLEDGMENTS

This work was made possible by the generous assistance and support of many people. I would like to offer my sincere thanks to Stephanie Vlachos, who has been my partner in this research endeavor from the beginning, for all her enthusiasm and hard work. I would also like to acknowledge my academic advisor, Herb Roitblat, who has been my toughest reviewer, my biggest supporter, and a good friend. I feel very fortunate to have had fantastic advisors including Paul Nachtigall and Whitlow Au; both were essential for making this project a reality. I would like to recognize Marlee Breese and Paka Nishimura, who supported this project as curators of Sea Life Park Hawaii, and Jeff Pawloski, who helped tremendously with animal training. To Dan Blaine and the Monday night statistics group, thanks for the great discussions and camaraderie. For her love and support I thank Kristen Stokes Lemonds.

ABSTRACT

High frequency auditory filter shapes were calculated for a 20 year old female Atlantic bottlenose dolphin, *Tursiops truncatus*. Thresholds were determined for tones of 40, 60, 80, and 100 kHz masked by notched noise. Auditory filter shapes were determined by fitting the integral of the *roex(p,r)* filter shape to the functions relating masked tonal threshold to notch width. Filter shapes were found to be approximately symmetric at the moderate noise level used. The equivalent rectangular bandwidths of the auditory filters ranged from 16% of center frequency at 40 kHz to 11% of center frequency at 100 kHz. There was very little change in the bandwidths of the filters between 60 and 100 kHz, indicating that relative tuning sharpness increases as a function of frequency in this range. Efficiency of processing after the periphery was found to be maximal at 40 and 60 kHz (better than –12 dB SNR) and to decrease gradually above 60 kHz. The efficiency estimates allowed for the reanalysis of critical ratio data collected previously for the same animal (Lemonds et al., 1997); the bandwidth estimates from the present and earlier study were found to be in good agreement.

TABLE OF CONTENTS

Acknowledgments iii

Abstract iv

List of Tables vii

List of Figures viii

Chapter 1: Introduction 1

Physical characteristics of sound 1

Psychoacoustics 3

Auditory sensitivity 3

Auditory frequency selectivity 5

Filters 5

Frequency selectivity 7

Bandpass filter bank 8

Cochlear biomechanics 11

Auditory filter shapes 14

Power spectrum model of masking 15

Critical ratio 17

Notched noise masking 20

Off-frequency listening 23

Filter asymmetry 25

Roex filter shapes 27

Roex(p,r) filter shape 29

Fitting filter shapes 29

Purpose of this study 31

Chapter 2: Methods 33

Subject 33

Electronic apparatus 33

Ambient noise levels 35

Stimulus and masking noise 36

Experimental apparatus 38

Procedure 39

Chapter 3: Results 44

Threshold as a function of notch width 44

Estimating the filter shapes 45

Filter shape results 46

Chapter 4: Discussion 53

Frequency discrimination and selectivity 54

Critical ratios revisited 56

Auditory modeling 58

Conclusions 61

Appendix A: Bandwidths in Hz and dB 63

Appendix B: Constant-Q functions 64

Appendix C: Simulating filter asymmetry 66

References 71

LIST OF TABLES

Table Page

1. Interaction between filter asymmetry and notch asymmetry 27
2. Go/no-go decision matrix 40
3. Tonal threshold as a function of notch width 44
4. Asymmetric notch results 44
5. Summary of filter shapes 52
6. ERB results for CR, "corrected" CR, and filter shape 58

LIST OF FIGURES

Figure Page

1. Pure tone (500 Hz) 2
2. Spectrum of a 5 kHz tone 2
3. White noise in the time and frequency domains 2
4. Human audiogram 4
5. Pure tone detection thresholds for the bottlenose dolphin 4
6. Low-pass filtering 6
7. Bandpass filter 6
8. –3 and –10 dB bandwidths 7
9. ERB of a bandpass filter 8
10. Bandpass filter bank 8
11. Impulse response of a filter bank 10
12. Filter bank output from a 2-tone complex 11
13. Traveling wave 13
14. Filter bank, noise, and tone 16
15. Critical ratios and Q for human 18
16. *Tursiops* critical ratio and Q 19
17. Notched noise 21
18. Notched noise masking of a tone 22
19. Off-frequency listening 24
20. Asymmetric filtering 26

21. *Roex(p)* filter shape 28

22. *Roex(p,r)* filter shape 29

23. *Roex(p,r)* tail integral 30

24. Electronic apparatus 34

25. Ambient noise level in experimental tank 35

26. Notched noise (g=±0.1) at 80 kHz 36

27. Notched noise (g=±0.2) at 80 kHz 37

28. Notched noise (g=±0.3) at 80 kHz 37

29. Notched noise (g=±0.4) at 80 kHz 38

30. Experimental apparatus 38

31. Threshold calculation based on reversal levels 41

32. Error surface for *roex(p,r)* fit at 40 kHz 47

33. Thresholds and integral fit at 40 kHz 47

34. Auditory filter at 40 kHz 47

35. Error surface for *roex(p,r)* fit at 60 kHz 48

36. Thresholds and integral fit at 60 kHz 48

37. Auditory filter at 60 kHz 48

38. Error surface for *roex(p,r)* fit at 80 kHz 49

39. Thresholds and integral fit at 80 kHz 49

40. Auditory filter at 80 kHz 49

41. Error surface for *roex(p,r)* fit at 100 kHz 50

42. Thresholds and integral fit at 100 kHz 50

43. Auditory filter at 100 kHz 50

44. Auditory filter shapes in *Tursiops* 51

45. Frequency discrimination in *Tursiops* 55

46. Frequency limens and model predictions for *Tursiops* 56

47. Third order polynomial estimate of K 59

48. Auditory filter ERB estimates for *Tursiops* 60

49. Plots of constant-Q functions 64

50. Simulated filter bank 68

51. Simulation results 69

52. Simulation contour plots for 60 and 80 kHz 70

CHAPTER 1: INTRODUCTION

The purpose of this study is to characterize the auditory frequency processing capabilities of an Atlantic bottlenose dolphin (*Tursiops truncatus*). In a controlled psychoacoustical experiment, sensitivity to tonal signals masked by noise was measured as a function of changing spectral characteristics of the noise. A theoretical model was derived that explained the experimental results, compelled reexamination of earlier studies (including Lemonds, Au, Nachtigall, Vlachos, & Roitblat, 1997), and supported the specification of a computational auditory model of frequency selectivity in *Tursiops*.

Physical characteristics of sound

The psychology of audition is concerned with understanding auditory sensations as a function of the physical nature of sound, the stimulus for hearing. This task entails describing the physical characteristics of acoustic energy and the auditory system; the former will be described first. Sounds are oscillations in the pressure of a medium over time. Two fundamental physical properties of sound are amplitude and frequency. Amplitude is the magnitude of change in pressure during an oscillation and is perceived as loudness, while frequency refers to the rate of oscillation in pressure and is perceived as pitch.

A simple sound is a sine wave, also known as a sinusoid, which has the waveform (the pattern in pressure/time or the time domain) shown in Figure 1.

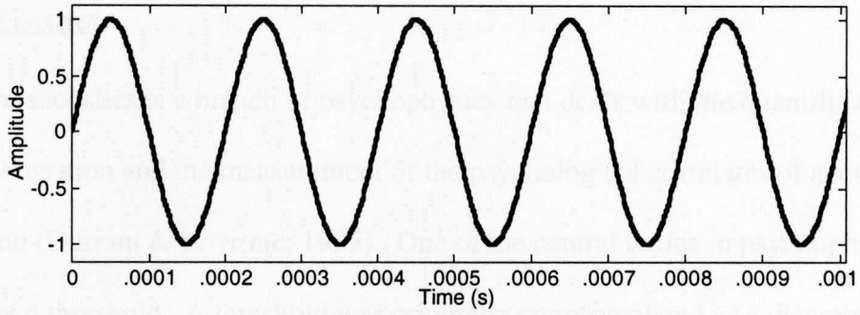


Figure 1. Pure tone (5000 Hz).

Figure 2 shows the spectrum of the 5000 Hz (5 kHz) tone shown in Figure 1. The spectrum of a sinusoid is simply a straight line because all of its energy is concentrated at a single frequency.

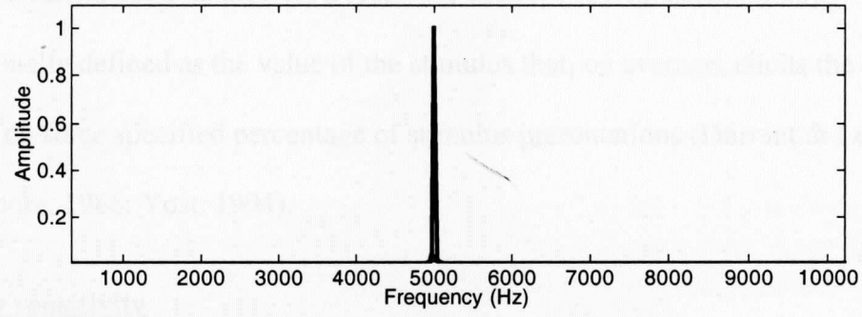


Figure 2. Spectrum of a 5 kHz tone.

Complex signals have energy at more frequencies; a completely random "white" noise is shown in Figure 3 that has equal amplitude at every frequency.

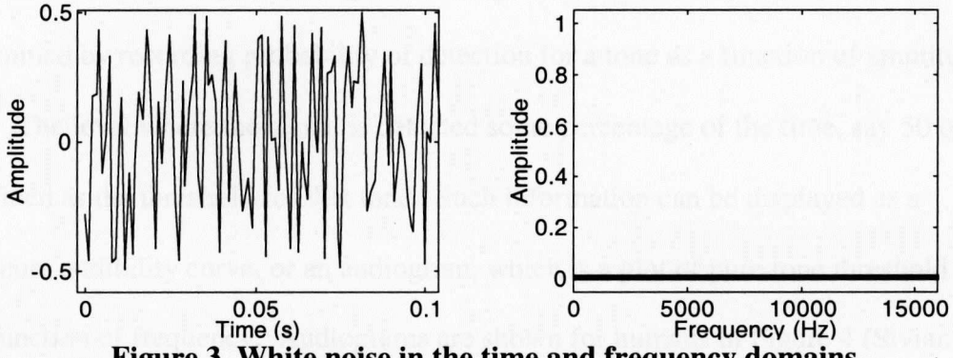


Figure 3. White noise in the time and frequency domains.

Psychoacoustics

Psychoacoustics is a branch of psychophysics that deals with the quantification of auditory sensation and the measurement of the psychological correlates of acoustical stimulation (Durrant & Lovrinic, 1995). One of the central topics in psychophysics is the concept of a threshold. A threshold was originally conceptualized as a discrete point along a physical continuum of a stimulus above which a response is always generated by the stimulus and below which a response is never generated by the stimulus. However, experimental results do not support this notion of a discrete point, rather, response to a stimulus is statistical in nature. As a result the concept of a threshold is now conventionally defined as the value of the stimulus that, on average, elicits the desired response on some specified percentage of stimulus presentations (Durrant & Lovrinic, 1995; Moore, 1986; Yost, 1994).

Auditory sensitivity

Specifying a subject's hearing range and sensitivity is often the first step in a psychoacoustic study. This is typically accomplished by determining tonal detection thresholds over the frequency range of hearing. Tonal detection thresholds are determined by recording probability of detection for a tone as a function of amplitude or level. The level where the signal is detected some percentage of the time, say 50 or 75%, is defined as the threshold for that tone. Such information can be displayed as a minimum audibility curve, or an audiogram, which is a plot of pure tone threshold level as a function of frequency. Audiograms are shown for humans in Figure 4 (Sivian &

White, 1933), and Atlantic bottlenose dolphins *(Tursiops truncatus)* in Figure 5 (Johnson, 1967; Lemonds, 1997).

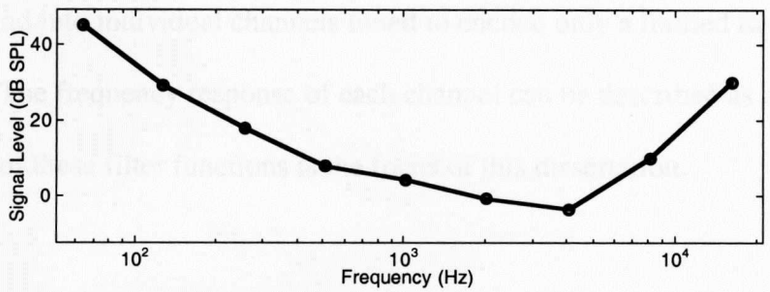


Figure 4. Human audiogram *(Sivian & White, 1933).*

The "U" shaped sensitivity curve evident in these figures is typical for all vertebrates tested, though the frequency range over which this general curve extends can vary greatly between species (Fay, 1988). Tonal detection thresholds for the Atlantic bottlenose dolphin extend over a wider frequency range than that for humans and most other mammals. Dolphins can hear waterborne sounds up to 150 kHz and possibly beyond. The dolphin audiogram looks much like the human audiogram shifted upward in frequency by a factor of ten (Johnson, 1967; Au, 1993).

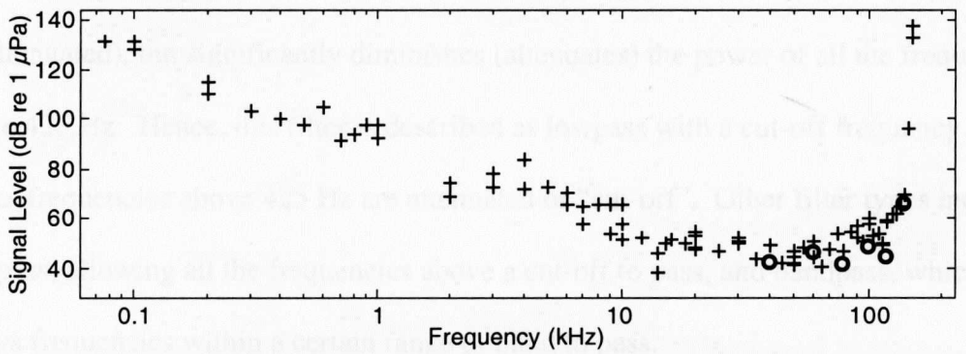


Figure 5. Pure tone detection thresholds for the bottlenose dolphin.
(+ Johnson, 1967; **o** Lemonds, 1997)

Auditory frequency selectivity

The auditory system selectively encodes frequency information, that is, it separates an incoming sound into individual channels tuned to encode only a limited range of frequencies. The frequency response of each channel can be described as a filter; the specification of these filter functions is the focus of this dissertation.

Filters

A filter can be described in either the time or frequency domain, but it is easier to initially conceptualize the workings of a filter in the frequency domain. A filter is a device that responds differentially to the frequency content of an input signal. An example is a lowpass filter, shown schematically in Figure 6. The input to this filter is shown in the left subplot; it consists of a high frequency noise band and a three tone complex. The filter is shown in the center (this is known as the *frequency response* of the filter), and the output is shown on the right; all these stages are shown in the frequency domain. Note that this filter allows the lower frequency content to pass at full power (unattenuated), but significantly diminishes (attenuates) the power of all the frequencies above 425 Hz. Hence, this filter is described as lowpass with a cut-off frequency of 425 Hz, as frequencies above 425 Hz are attenuated or "cut-off". Other filter types include: highpass, allowing all the frequencies above a cut-off to pass, and bandpass, which allows frequencies within a certain range or band to pass.

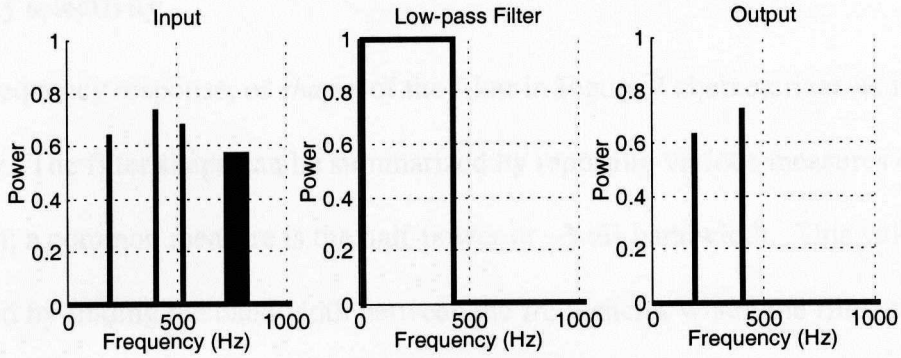


Figure 6. Low-pass filtering.

The filter in Figure 6 has characteristics that can only be achieved as an ideal: real filters do not have infinitely steep slopes and such sharp cut-offs. Rather, real filter responses slope off more gradually. An example of a more realistic filter, in this case a bandpass filter centered at 10 kHz, is shown in Figure 7. Both subplots show the same filter; the ordinate on the left is given in linear power, while the ordinate on the right is given in dBs defined as $10log_{10}$(linear power); it is often convenient to display filters on dB scales. Some helpful conversions between linear and dB measures are 0.5 power (1/2 power) = -3 dB, 0.1 power = -10 dB, 0.01 power = -20 dB, etc.

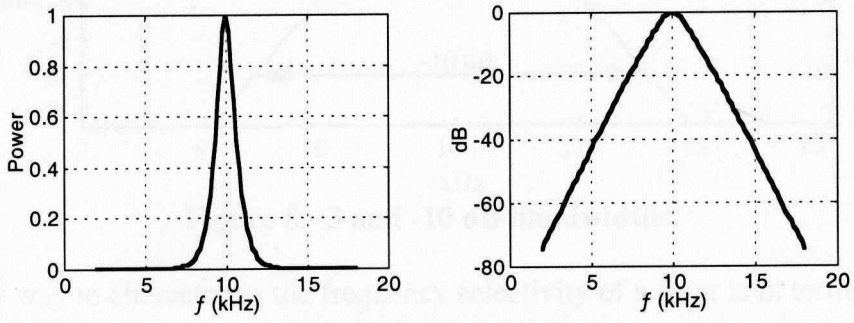


Figure 7. Bandpass filter.

Frequency selectivity

The frequency response, or *shape*, of the filter in Figure 7 characterizes its frequency selectivity. The filter shape can be summarized by reporting various measures of bandwidth; a common measure is the half-power or –3 dB bandwidth. This value is determined by finding the bandwidth between the frequencies where the filter output is decreased by 3 dB relative to its maximum output (Durrant & Lovrinic, 1995). The –3 dBbandwidth and another common measure, the –10 dB bandwidth, for the filter in Figure 7 are shown in Figure 8. The lower frequency –3 dB point is 9.3 kHz, and the higher frequency –3 dB point is 10.7 kHz. Therefore the –3dB bandwidth of this filter is 1.4 kHz (10.7 – 9.3), or 14% of its center or characteristic frequency (f_0) of 10 kHz. The –10 dB bandwidth is approximately 3.2 kHz (11.6 kHz - 8.4 kHz), or 32% of f_0. The –3 and –10 dB bandwidths can also be denoted as $\Delta f_{\text{-3dB}}$ and $\Delta f_{\text{-10dB}}$ respectively.

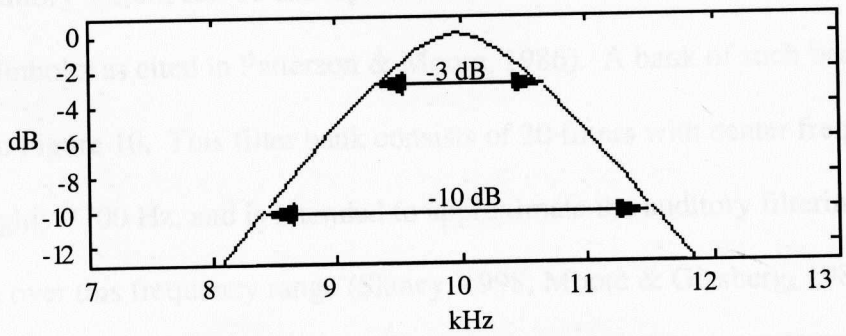


Figure 8. -3 and -10 dB bandwidths.

Another way to characterize the frequency selectivity of a filter is in terms of the area under its frequency response curve. The total area under the bandpass filter described in Figure 7 and Figure 8 is 1.59 (kHz x power), as determined by numerical integration. As the area of a rectangle is simply *h* (height) * *w* (width), a rectangle with *h*=1 and *w*=1.59 has the same area as this filter, which also has height (power) =1. The width of a

rectangle that has the same area and height as a bandpass filter is known as the equivalent rectangular bandwidth (ERB) of that filter. In this example the ERB = 1.59 kHz. The filter and ERB are shown in Figure 9; on a linear power scale the geometric interpretation of equal area beneath the curve and the rectangle is obvious. The ERB as a characterization of frequency selectivity is frequently used in this dissertation.

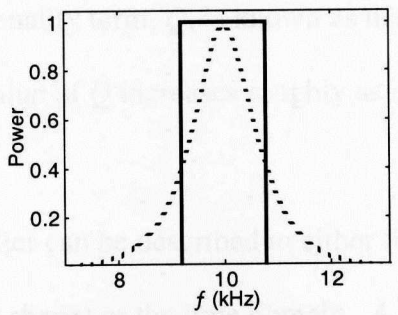


Figure 9. ERB of a bandpass filter.

Bandpass filter bank

The auditory system can be conceptualized as a bank of overlapping, linear, bandpass filters (Helmholtz as cited in Patterson & Moore, 1986). A bank of such bandpass filters is shown in Figure 10. This filter bank consists of 20 filters with center frequencies from 500 to roughly 3700 Hz, and is intended to approximate the auditory filtering capabilities of humans over this frequency range (Slaney, 1998; Moore & Glasberg, 1983, 1990).

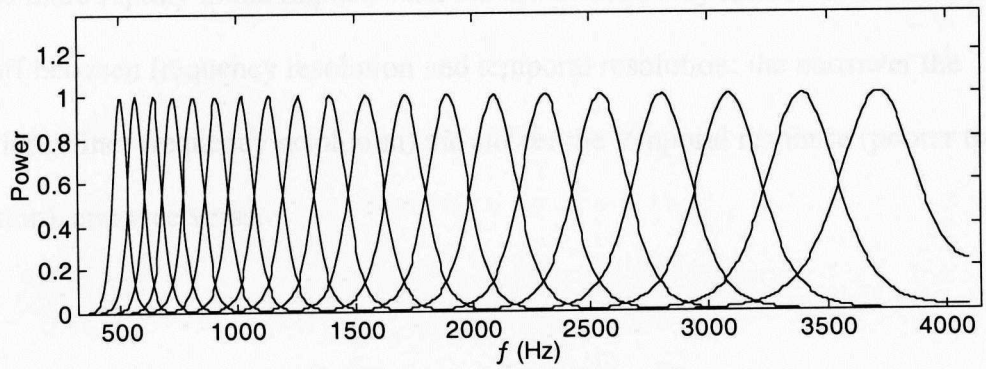


Figure 10. Bandpass filter bank.

Notice that the filter bandwidths increase with increasing frequency. The relationship between bandwidth and center frequency can be described by the following equation:

$$Q = {f_0}/{\Delta f} \tag{1}$$

which simply states that the filter bandwidth (Δf) is proportional (Q) to the filter's center frequency (f_0). The proportionality term, Q, is known as the *quality* of the filter. For the filter bank in Figure 10 the value of Q increases roughly as a logarithmic function of f_o for most values of f_o.

As mentioned above, a filter can be described in either the frequency domain (as a frequency response or a filter shape) or the time domain. A filter can be characterized in the time domain by finding its *impulse response*. An impulse is defined as:

$$f(t) = \begin{cases} 1, & \text{for } t = 0 \\ 0, & \text{for } t > 0 \end{cases} \tag{2}$$

where t is time. The Fourier transform (frequency domain representation) of an impulse is equal to 1 at every frequency. The impulse response for each channel or filter in the bank shown in Figure 10 is shown in Figure 11; notice that the higher frequency filters respond more rapidly to the impulse than the lower frequency filters. There is an inherent trade off between frequency resolution and temporal resolution: the narrower the bandwidth (finer frequency resolution) the slower the temporal response (poorer time resolution), and vice versa.

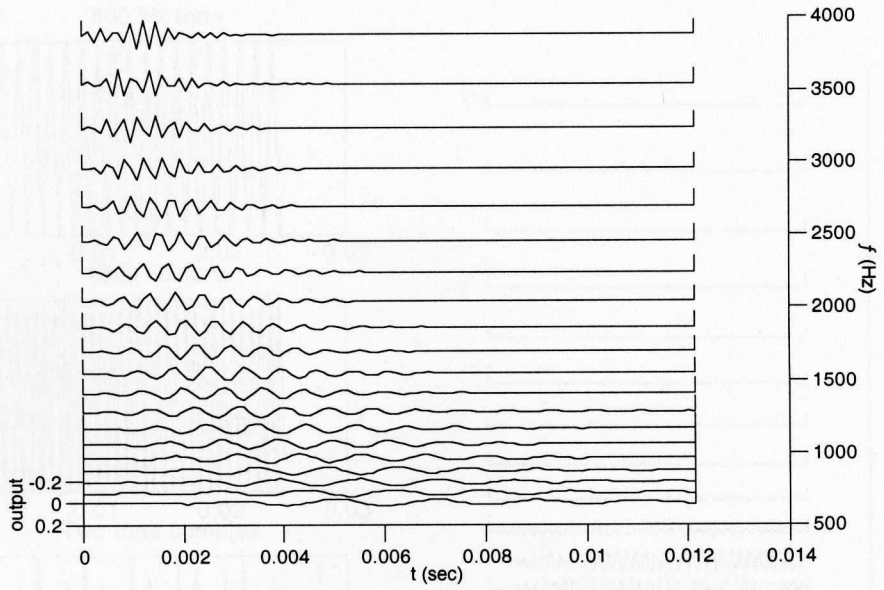


Figure 11. Impulse response of a filter bank.

Filter banks can be used to simulate the response of the auditory system to any given signal. Take as an input signal a two-tone complex containing 800 and 1350 Hz components; the two tones and the complex (their sum) are shown on the left side of Figure 12 as waveforms. A portion of the filter bank output from the tonal complex is shown on the right side of the figure. Notice that substantial output is generated in only 4 or 5 of the 20 filters or channels, those 4 or 5 channels have center frequencies relatively close to one of the two tonal frequencies. The filter bank has encoded the incoming signal such that the frequency information is evident in the activity across the channels. Hence the filter bank, and by analogy the auditory system, converts one representation of sound (a time/pressure waveform) into a another (activity across a bank of channels) which explicitly represents salient features of the original signal (frequency content as a function of time).

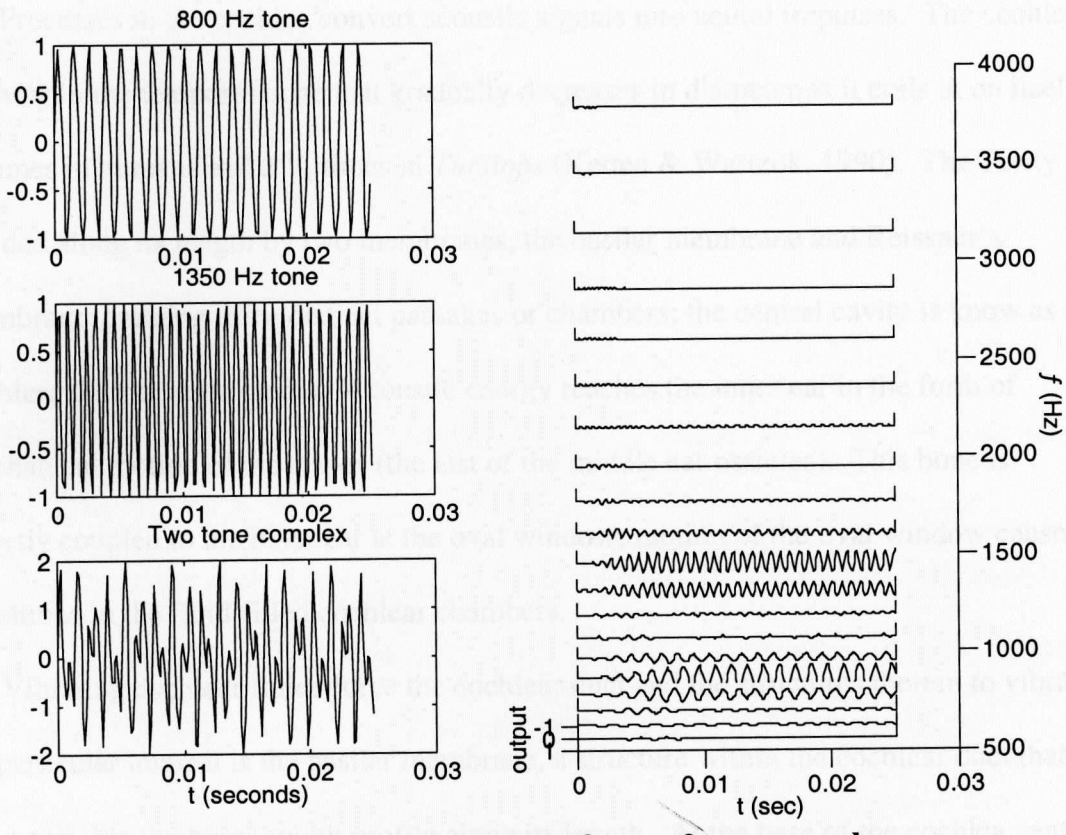


Figure 12. Filter bank output from a 2-tone complex.

Cochlear biomechanics

How is a filter bank biologically implemented in the auditory system? The answer, at least in part, can be found in descriptions of the functional anatomy and physiology of the generalized mammalian auditory system (Durrant & Lovrinic, 1995; Kelly, 1991; Luce, 1993; Pickles, 1986; Yates, 1986; Yost, 1994). Many interesting processes and mechanisms are involved in the transmission and conversion of acoustic energy in the outer and middle ear (Denes & Pinson, 1993) and this is especially the case for the dolphin's aquatically adapted system (Au, 1993; Ketten, 1992, 1994; McCormick, Weaver, Palin, & Ridgway, 1970). It is in the inner ear, and in the cochlea specifically, where peripheral auditory filtering takes place.

Processes in the cochlea convert acoustic signals into neural impulses. The cochlea is a cavity in the temporal bone that gradually decreases in diameter as it coils in on itself 2 $^{5}/_{8}$ times in humans and 2 $^{1}/_{4}$ times in *Tursiops* (Ketten & Wartzok, 1990). The cavity is divided along its length by two membranes, the basilar membrane and Reissner's membrane, creating three distinct passages or chambers; the central cavity is know as the cochlear duct or scala media. Acoustic energy reaches the inner ear in the form of mechanical motion of the stapes (the last of the middle ear ossicles). This bone is directly coupled to the inner ear at the oval window; motion of the oval window causes vibrations in the fluid-filled cochlear chambers.

Vibrating cochlear fluids force the cochlear duct and the structures therein to vibrate. Of particular interest is the basilar membrane, a structure within the cochlear duct that has a variable thickness-width profile along its length. At the base of the cochlea near the oval window the membrane is thick, narrow, and relatively stiff, while at the apex of the coil or the apical end it is thin, wide, and relatively flexible. In *Tursiops* the membrane is approximately 30 x 25 μm (width vs. thickness) at the basal end and 380 x 5 μm at the apical end (Ketten, 1994). When presented with a tonal signal the amplitude of the vibratory pattern of the basilar membrane varies with position. High frequency energy excites the basal end of the membrane maximally, while low frequency energy excites the apical end maximally; hence the amplitude of vibration at any point on the basilar membrane is said to be *tuned* to a limited frequency range. The intensity level of the input stimulus is reflected in the magnitude of displacement, with louder sounds generally producing larger displacements.

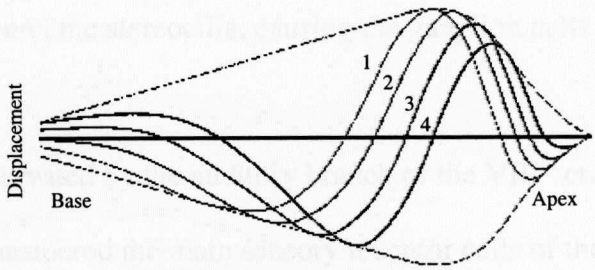


Figure 13. Traveling wave.

The actual pattern of vibration as it develops over time is called a traveling wave because the membrane vibrates in a wave motion that appears to travel toward the apex (von Békésy, 1960). Figure 13 shows a schematic example of the traveling wave generated by a sinusoid. The numbers indicate the temporal ordering of the sequence, while the dashed line represents the maximal displacement occurring at that point on the membrane. When presented with a complex sound, the membrane is displaced maximally at multiple locations corresponding to the frequency components of the complex signal. In this way the biomechanics of the basilar membrane perform a type of spectral analysis on the incoming signal (Yost, 1993; Denes & Pinson, 1993), encoding frequency as a function of location of disturbance along the membrane.

Lining the scala media side of the basilar membrane is the organ of Corti: an ensemble of structures responsible for converting the mechanical motion of the basilar membrane into neural signals. The organ of Corti consists of inner and outer hair cells and a variety of supporting cells (see chapters 7 & 8 in Yost, 1993). Hair cells are named for the stereocilia, or tiny hairs, that extend from the cell body into the scala media. The tectorial membrane, a soft ribbon-like structure, rests on or just above the stereocilia. When the basilar membrane is displaced, shearing forces between the stereocilia and the

tectorial membrane bend the stereocilia, causing the ganglion cells at the base of the hair cells to fire.

Hair cells are innervated by the auditory branch of the VIIIth cranial nerve. Inner hair cells are generally considered the main sensory receptor cells of the mammalian cochlea, while outer hair cells are believed to act primarily as mechanical effectors in a feedback loop that alters the mechanical response of the basilar membrane and, consequently, input to the inner hair cells (Dallos, 1992; Yates, 1986). The outer hair cell feedback loop plays a prominent role in determining the response characteristics of the basilar membrane. This is clearly seen in the fact that tuning in the mammalian cochlea is physiologically vulnerable: damage to outer hair cells results in a loss of sensitivity and a broadening of the frequency selectivity of the basilar membrane (Wilson, 1992; Yates, 1986).

The auditory branch of the VIIIth cranial nerve is organized such that fibers that innervate the apex are in the center of the bundle, while those innervating the base are on the outer edges of the bundle. Hence there exists an orderly connection between hair cells and auditory nerve fibers, such that the patterns of excitation evident in the auditory nerve fibers reflect the frequency and intensity information encoded in the periphery (Yost, 1993; Kelly, 1991).

Auditory filter shapes

As discussed above, the peripheral auditory system can be modeled as a bank of overlapping bandpass filters. The characterization of these auditory filters, both in terms

of bandwidth and of actual shape or frequency response, can be accomplished using psychophysical masking experiments.

Power spectrum model of masking

Masking is the process by which the threshold for one sound is raised by the presence of another (masking) sound (Moore, 1989). A popular model of auditory masking assumes that a listener detects a tonal signal by monitoring an auditory filter centered on that tonal frequency. As indicated in Figure 12, filters or channels with characteristic frequencies (f_o) near that of a signal pass more of the signal energy than other filters.

Imagine that instead of the two tones shown in Figure 12, the 800 Hz tone was replaced by random noise. Such a scenario is illustrated in Figure 14. Notice that now there is a substantial amount of activity in a number of channels quite remote from 1350 Hz. The channels centered near 1350 Hz are much less consistent in their output level, and their output is quite noisy as they are now passing both the tone and whatever random noise components align with the bandpass. Hence the output of the channel contains a mix of signal and noise components; this mix is quantified as a ratio called the signal-to-noise ratio (SNR). There is a much lower signal-to-noise ratio in the channel centered near 1350 Hz in Figure 14 than in the same channel in Figure 12, hence adding random noise made the task of detecting the tone at 1350 Hz more difficult. The noise contributes to raising the threshold for, or masking, the tone.

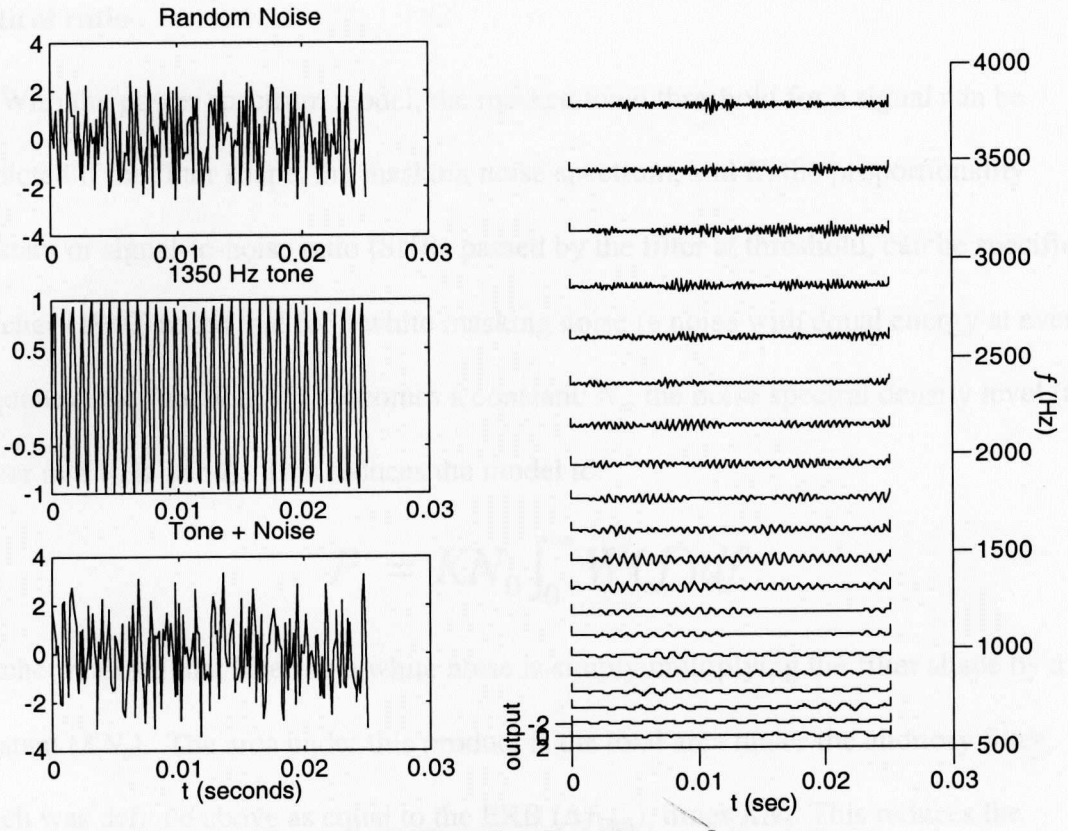


Figure 14. Filter bank, noise, and tone.

The power spectrum model of masking adds the assumption that, at threshold, the power of the signal is some proportion (K) of the noise power passed by the auditory filter (Fletcher, 1940). This can be expressed as:

$$P_s = K\int_0^\infty N(f)W(f)df \tag{3}$$

where P_s is the power of the signal at threshold and K is the proportionality constant. The power of the noise passed by the auditory filter is expressed as the integral of the product of the masking noise, $N(f)$, and the auditory filter, $W(f)$, both expressed in the frequency domain. The power spectrum model is so named because it predicts masked threshold solely on the basis of the power spectra of the noise and the underlying auditory filter.

Critical ratio

With the power spectrum model, the masked tonal threshold for a signal can be predicted if the filter shape, the masking noise spectrum, and K, the proportionality constant or signal-to-noise ratio (SNR) passed by the filter at threshold, can be specified. Fletcher (1940) noted that for a white masking noise (a noise with equal energy at every frequency) the function $N(f)$ becomes a constant: N_o, the noise spectral density level (the power in a 1 Hz band). This reduces the model to:

$$P_s = KN_0 \int_0^\infty W(f)df. \tag{4}$$

Fletcher realized that filtering a white noise is simply multiplying the filter shape by a constant (KN_0). The area under this product is the total area under the auditory filter, which was defined above as equal to the ERB (Δf_{ERB}), times KN_0. This reduces the equation to:

$$P_s = KN_0 \Delta f_{ERB}. \tag{5}$$

Furthermore, Fletcher had previously conducted studies from which he concluded that K=1. This reduces the model to a form known as the critical ratio:

$$\Delta f_{ERB} = P_s / N_0 \tag{6}$$

which expresses the ERB of an auditory filter as the ratio of threshold for a tone masked by white noise to the noise spectral density of the masker.

In practice, the critical ratio threshold is psychoacoustically determined by adjusting the level of a tone masked by a constant level white noise. Results from a critical ratio

study with humans are shown in Figure 15 (Hawkins & Stevens, 1950). Note that the critical ratios are expressed as bandwidths (Δf) in dB re 1 Hz (see Appendix A). The Q values ($Q = f_0/\Delta f$) for the critical ratios are shown on the right subplot (see Equation 1 and Appendix B: Constant-Q functions). The Q values were found to increase as a logarithmic function of frequency up to 1000 Hz, and are roughly constant-Q between 1 and 10 kHz. Over this constant Q range, the bandwidth estimates are approximately 6-7% of center frequency (1/Q).

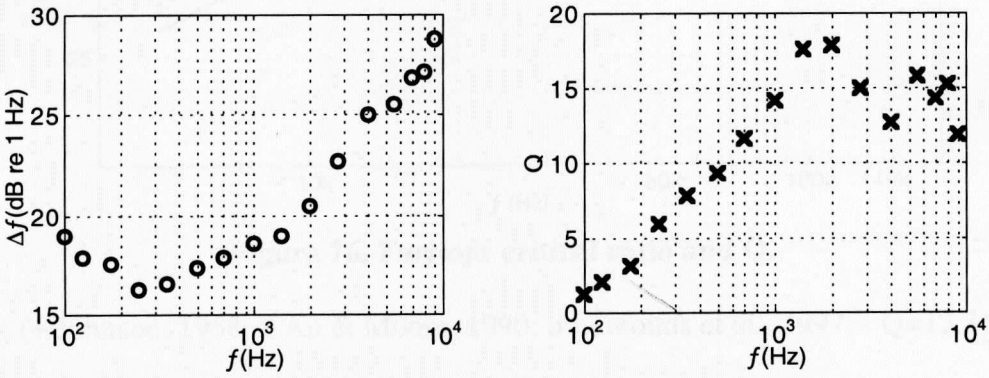


Figure 15. Critical ratios and Q for human.

Critical ratios (CRs) have been measured for a few cetacean species including the bottlenose dolphin (*Tursiops truncatus*) (Johnson, 1968; Au & Moore, 1990; Lemonds, Au, Nachtigall, Vlachos, & Roitblat, 1997), beluga whale (*Delphinapterus leucas*) (Johnson, McManus, & Skaar, 1989), and false killer whale (*Pseudorca crassidens*) (Thomas, Pawloski, & Au, 1990). Results for the three studies conducted on *Tursiops* are shown in Figure 16. Note that results from Lemonds et al. (1997) are from the same subject used in the present experiment.

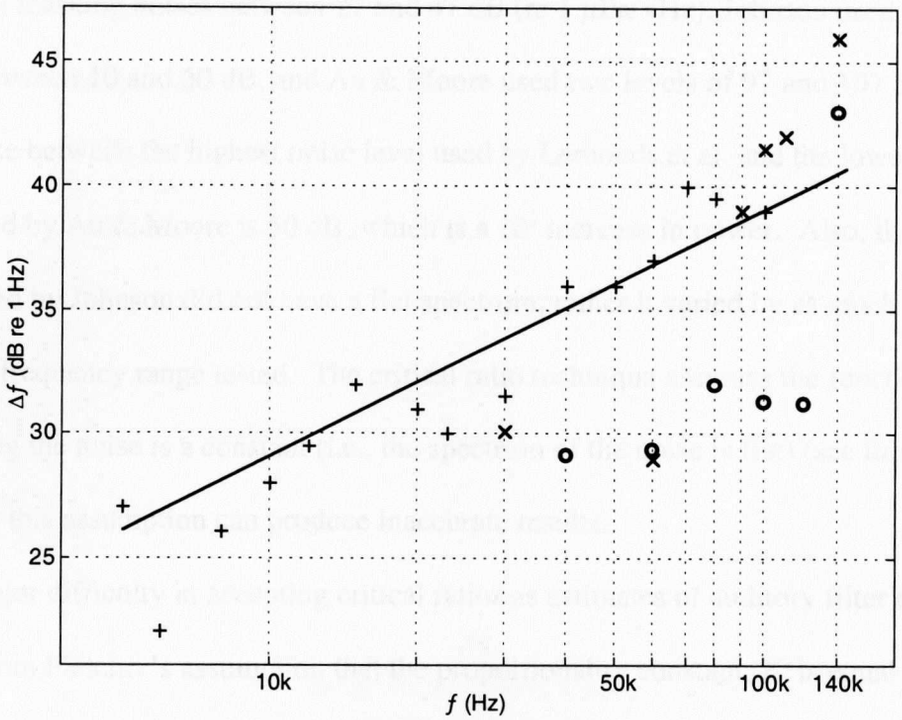


Figure 16. *Tursiops* critical ratio and Q.

(+ Johnson, 1968; x Au & Moore, 1990; **o** Lemonds et al., 1997; - Q=12.3)

The dolphin data are difficult to summarize due to a high level of variability between subjects. The results from Johnson (1968) and Au & Moore (1990) can be roughly approximated as a constant-Q function (Q=12.3), indicating that the estimated filter bandwidths were on average about 8% of center frequency. The results from Lemonds et al. (1997), however, do not fit this general trend. Between 40 and 120 kHz, the critical ratio bandwidth estimates ranged from approximately 1% (at 120 kHz) to 2% (at 40 kHz) of center frequency. The critical ratio at 140 kHz, which was 14% of center frequency, was in good agreement with the other studies at the extremely high frequencies.

The high variability in the results may be due to differences in the experimental conditions. Different masking noise levels were used in each study; Lemonds et al. used

low level masking noises between 27 and 47 dB (re 1 μPa/$\sqrt{\text{Hz}}$), Johnson used noise levels between 10 and 60 dB, and Au & Moore used two levels of 97 and 107 dB. The difference between the highest noise level used by Lemonds et al. and the lowest noise level used by Au & Moore is 50 dB, which is a 10^5 increase in power. Also, the masking noise used by Johnson did not have a flat spectrum, rather it varied by as much as 25 dB over the frequency range tested. The critical ratio technique assumes the function describing the noise is a constant (i.e., the spectrum of the noise is flat) (see Equation 4); violating this assumption can produce inaccurate results.

A major difficulty in accepting critical ratios as estimates of auditory filter bandwidth comes from Fletcher's assumption that the proportionality constant, K, is equal to 1. From Equation 5, K can be interpreted as the signal-to-noise ratio passed by the filter at threshold, or equivalently as the processing efficiency of the detection mechanism after the auditory filter (Patterson, Nimmo-Smith, Weber, & Milroy, 1982). Recent studies with humans have shown that the assumption that K=1 is in error (see Shailer, Moore, Glasberg, Watson, & Harris, 1990 as an example); therefore the logic of using the critical ratio, which assumes K=1, to estimate auditory filter bandwidth is flawed. The critical ratio is a function of both bandwidth and efficiency, and in the form given in Equation 6 there is no way to distinguish between these two factors (Patterson, Nimmo-Smith, Weber, & Milroy, 1982).

Notched noise masking

A technique to measure the effects of not only efficiency and bandwidth, but also the actual auditory filter shape, was derived by Patterson (1976). Starting with the formal

description of the power spectrum model of masking given in Equation 3, Patterson, like Fletcher, simplified the power spectrum model by his selection of masking noise. The type of noise he described was wide-band having a deep notch with sharp edges wherein the noise was removed (or band-stop filtered). An example of notched, or bandstop, noise is shown idealized in Figure 17, with the notch width centered at 500 Hz.

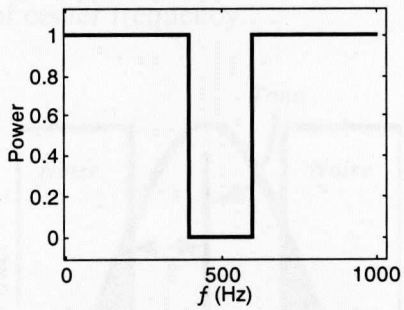


Figure 17. Notched noise.

The cut-off points of notched noise are typically described relative to the center frequency of the notch. The noise in Figure 17 has cut-offs at 400 and 600 Hz; in relation to the center of the notch, 500 Hz, these cut-offs are ±100 Hz respectively. It is convenient to describe these relative cut-offs as a proportion of center frequency: (±100)/500 = ±0.2 of 500 Hz. If a cut-off frequency is denoted as f and the center frequency of the notch as f_0, then f normalized as a proportion of center f_0 can be expressed as g as follows:

$$g = (f - f_0)/f_0. \tag{7}$$

Expressing center and cut-off frequencies in terms of g allows for easy comparison of notches at any frequency; the center frequency expressed as g is equal to 0, while the cut-off frequencies are always given as ± some % of that center frequency. This expression turns out to be very convenient in generalizing the equations that follow.

A notched noise masking experiment employs a notched noise to mask a tonal signal of a frequency within the notch bandwidth. An ideal example of such an experiment is shown in Figure 18. In this particular example the tone (g=0) is centered in the notch band of the noise, and the auditory filter at that characteristic frequency is indicated in the figure as the smooth curve. The notch cut-offs frequencies are given in terms of center frequency g, the proportion of center frequency.

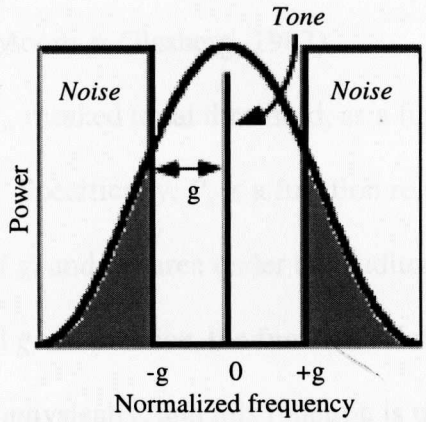


Figure 18. Notched noise masking of a tone.

For a notched noise masker and an auditory filter expressed in terms of g, the power spectrum model of masking (Equation 3) can be expressed as:

$$P_s = KN_0 f_0 \int_{-\infty}^{-g} W(g)dg + KN_0 f_0 \int_{g}^{\infty} W(g)dg. \qquad (8)$$

Note that the two integrals in Equation 8 are shown in Figure 18 as the two shaded areas under both the auditory filter, $W(g)$, and the notched noise; N_0 is represented by the height of the noise power outside of the notch, and P_s is represented by the height of the tone. The f_0 term is multiplied by each integral to convert normalized area into real frequency.

Patterson (1976) observed that if the auditory filter shape, $W(g)$, is assumed to be symmetric on a linear frequency scale, as it is shown in Figure 18, then the areas under the two integrals in Equation 8 are identical, reducing the equation to:

$$P_s = 2KN_0 f_0 \int_g^{\infty} W(g)dg. \qquad (9)$$

For humans the assumption of symmetric auditory filters is generally appropriate for moderate masking levels (Moore & Glasberg, 1987).

Equation 9 expresses P_s, masked tonal threshold, as a function of the integral of the auditory filter shape, $W(g)$. Specifically, P_s is a function relating tonal threshold to notch width expressed in terms of g, and the area under the auditory filter is evaluated as the tail integral of the filter beyond g. In practice, the function relating tonal threshold to notch width is determined psychophysically, and this function is used to determine the filter shape (Patterson, 1976; Patterson et al., 1982).

Off-frequency listening

In the masking models presented to this point the center frequency of the auditory filter has always been assumed to be the same frequency as the tonal signal. This assumption is not always appropriate. As an example, consider the task where a 500 Hz tone is masked by a bandpass noise with a low frequency cut-off at 450 Hz and a high frequency cut-off at 625 Hz, as shown in Figure 19. Two auditory filters, one with a characteristic frequency of 500 Hz (the same as the tonal frequency), and one with a characteristic frequency of 517 Hz, are shown in the figure. In this case the SNR (signal

power/noise power under the curve) at the output of the filter at 517 Hz is actually 1 dB higher than the SNR at the output of the 500 Hz filter. A listener attempting to detect a 500 Hz tone can actually achieve a higher SNR by attending to a filter with a different characteristic frequency than the tone. This is called "off-frequency listening": the listener attends to a filter with a characteristic frequency "off" the tonal frequency. More generally, it is assumed that a listener will attend to the auditory filter that provides the highest SNR at its output (Patterson & Nimmo-Smith, 1980); this is called the "max SNR" assumption.

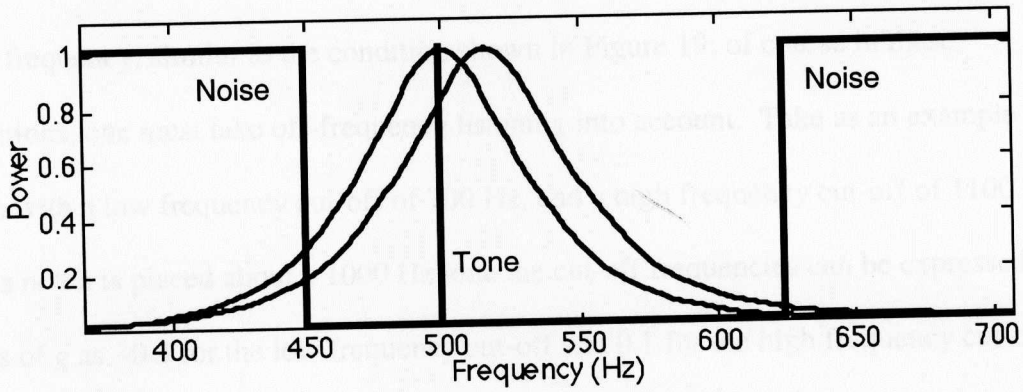


Figure 19. Off-frequency listening.

Patterson (1976) limited the efficacy of off-frequency listening by centering the notch on the tone, that is, the cut-offs of the lowpass and highpass bands were equidistant from the tone on a linear frequency scale. Under such conditions, the reduction of noise power passed by shifting the filter away from the lowpass noise is more than offset by the increase of noise coming in under the high frequency skirt. A shift away from the highpass noise will result in the same outcome, namely an increase in the total noise passed by the low frequency skirt of the filter. Thus, a listener in a centered notched noise task is constrained to use the filter centered at the tone to optimally perform the

detection task. (Patterson & Nimmo-Smith, 1980), which eliminates concerns of off-frequency listening.

Filter asymmetry

Auditory filters in humans have been shown to become increasing asymmetric with increasing masking level (Patterson & Nimmo-Smith, 1980; Moore & Glasberg, 1987); the lower branch of the filter becomes shallower and the upper branch becomes slightly steeper with increasing level. A technique to estimate the asymmetry of the auditory filter consists of masking a signal with the noise notch placed asymmetrically about the tonal frequency, similar to the condition shown in Figure 19; of course in these conditions, one must take off-frequency listening into account. Take as an example a notch with a low frequency cut-off of 700 Hz, and a high frequency cut-off of 1100 Hz. If this notch is placed about a 1000 Hz tone the cut-off frequencies can be expressed in terms of g as -0.3 for the low frequency cut-off and 0.1 for the high frequency cut-off; asymmetric notches are typically referred to by (Δg low, Δg high), so this notch condition is denoted as (0.3, 0.1). The "mirror" condition of this notch can be expressed in terms of g as (0.1, 0.3), which refers to a notch between 900 and 1300 Hz in this case. In both notch conditions a shifted auditory filter will likely provide the best SNR.

Given the max SNR assumption, there is an interaction between direction of filter asymmetry and maximum SNR on the two asymmetric masking noise conditions. As an example take two asymmetric filters: both filters have characteristic frequencies at 1000 Hz, the first has a negative (-) skew, that is most if its area is to the left of the characteristic frequency, while the second has a positive (+) skew, it has most of its area

to the right of its characteristic frequency. The (-) filter that provides the highest SNR in the (0.3, 0.1) condition is shown in the upper left subplot of Figure 20; the (–) filter with the highest SNR in the (0.1, 0.3) condition is shown in the bottom left subplot. Notice that the (–) filter shifted from the tonal frequency a much smaller distance than the (+) filter; the (–) skew is a better match for this noise condition. The best SNR's at the output of the filter for each noise condition are listed in Table 1; the SNR for the (0.3, 0.1) condition is almost 5 dB greater than that for the (0.1, 0.3) condition. The two subplots on the right hand side of Figure 20 show the (+) skewed filter that provides the best SNR on the two masking noise conditions; the best SNR's at the output of this filter shape on both noise conditions are listed in Table 1, with the SNR for the (0.1, 0.3) condition almost 3 dB greater than the (0.3, 0.1) condition.

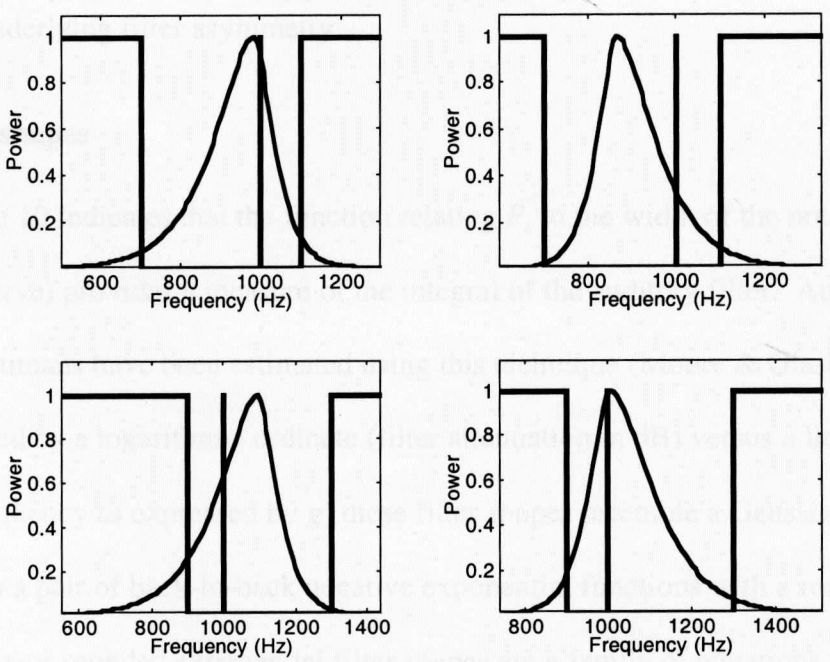


Figure 20. Asymmetric filtering.

Table 1. Interaction between filter asymmetry and notch asymmetry.

	Skew	
Notch	(-)	(+)
(0.3, 0.1)	19.22	16.19
(0.1, 0.3)	14.45	18.94

Note. Values in SNR dB.

There is an interaction between notch condition and skew evident in the SNR values; hence, filter asymmetry can be revealed by comparing the tonal thresholds for "mirror" asymmetric masking conditions (and see Appendix C: Simulating filter asymmetry). In this case, the difference between the (0.3, 0.1) and (0.1, 0.3) notch conditions for the (-) filter is approximately 5 dB, while the difference for the (+) filter is approximately –3 dB. The difference between the two notch conditions varies as a function of the direction and degree of underlying filter asymmetry.

Roex filter shapes

Equation 10 indicates that the function relating P_s to the width of the notch (the threshold curve) provides a measure of the integral of the auditory filter. Auditory filter shapes for humans have been estimated using this technique (Moore & Glasberg, 1987). When defined by a logarithmic ordinate (filter attenuation in dB) versus a linear abscissa (relative frequency as expressed by g) these filter shapes resemble a Gaussian curve or alternatively a pair of back-to-back negative exponential functions with a rounded top.

The Roex or rounded exponential filter shapes are a family of equations that have been developed to model the shape of auditory filters (Patterson et al. 1982; Patterson and Moore 1986). There are three versions of the Roex filter shapes, taking from one to three

arguments respectively. The simplest of the three shapes is the one parameter shape, the *roex(p)*, which is given by:

$$W(g) = (1 + pg)e^{-pg} \tag{10}$$

where the parameter p in the negative exponential determines the width of the passband and rate of fall of the filter skirts. The $(1+pg)$ term rounds off the top of the exponential, producing a filter that is relatively flat at its center frequency (Patterson & Moore, 1986). Figure 21 shows two roex(p) filter shapes back-to-back as they are used to model auditory filter shapes.

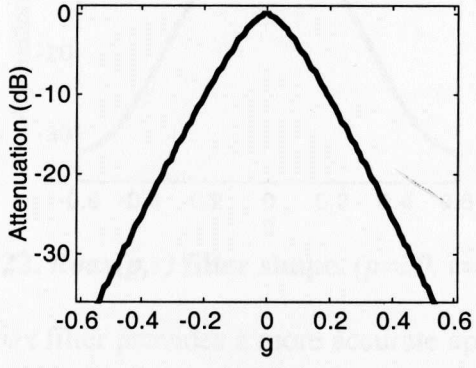


Figure 21. *Roex(p)* filter shape. (p=20)

This function provides a good approximation of actual auditory filters within 25 dB of the tip, but real auditory filters do not continue down indefinitely as this shape suggests. Rather the slope of auditory filters becomes markedly shallower around 35 dB down in response (in humans). To model these shallower tails, Patterson et al. suggested a two parameter roex filter which captures the tail effect by introducing a dynamic range limitation.

Roex(p,r) filter shape

The two-argument version of the roex, the *roex(p,r)* filter, has the form:

$$W(g) = (1-r)(1+pg)e^{-pg} + r. \tag{11}$$

The filter shape is shown in Figure 22. The parameter p, again in the negative exponential, determines the slope of the filter skirt, while r is a constant that limits the range of the function.

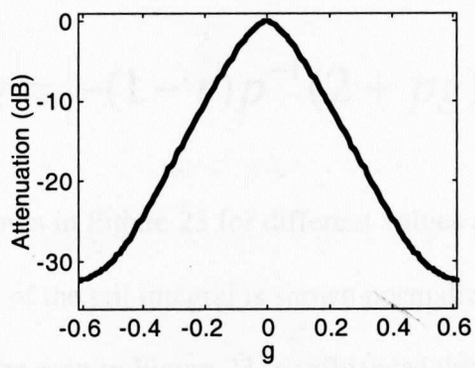


Figure 22. *Roex(p,r)* filter shape. *(p=20, r=0.0005).*

The three parameter *roex* filter provides a more accurate approximation of actual auditory filters but requires more data to find stable solutions. For this reason the *roex(p,r)* filter will be used in the present experiment.

Fitting filter shapes

Equation 9 indicates the curve relating masked tonal threshold to notch width is a measure of the integral of the auditory filter. The auditory filter shape is determined by substituting the integral of the *roex(p,r)* function Equation 9, and fitting this equation to the masked threshold curve.

Because the tail of the *roex(p,r)* function is flat (see Figure 22) the integration of the filter must be domain limited in order to keep the predicted threshold curve bounded. It was empirically found that setting domain limitations to between 0.6 and 1.0 had very little effect on the quality of fits that could be produced with notched noise data; all the fits were very similar. Therefore a domain limitation of $g = 0.8$ is typically applied for convenience (Patterson et al.,1982). The tail integral (the area under the filter beyond g) of the *roex(p,r)* filter is:

$$\int_g^{0.8} W(g)dg = \left|-(1-r)p^{-1}(2+pg)e^{-pg} + rg\right|_g^{0.8}. \quad (12)$$

The tail integral is shown in Figure 23 for different values of p and r. For convenience, the ordinate of the tail integral is shown normalized to the total area and expressed in dB. As can be seen in Figure 23, p influences the slope of the integral curve, while r determines where the tail will flatten with respect to g.

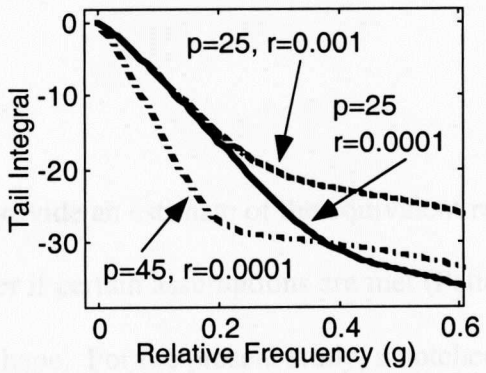


Figure 23. *Roex(p,r)* tail integral.

Substituting the tail integral (Equation 12) into the power-spectrum model for notched noise masking (Equation 9) produces:

$$P_s(g) = 2Kf_0N_0[-(1-r)p^{-1}(2+pg)e^{-pg} + rg]\Big|_g^{0.8} \quad (13)$$

Because $P_s(g)$ is in decibels, ten times the logarithm$_{10}$ of Equation 13 should be fit to the threshold curve (threshold in dB vs. $\Delta f/f$). Only the tail integral on the right hand side of Equation 13 determines the shape of the auditory filter, the remaining terms are simply additive constants when expressed in dB. Therefore the only free parameters in the fitting are p and r, the bandwidth and dynamic range limitation terms. These values are determined by fitting the threshold curve to the tail integral of the roex filter; both these curves are expressed as deviation values from their respective largest elements so that each curve has a value of 0 dB at the 0 notch condition. Once adequate estimates for p and r are found, the only remaining free parameter is K; the values of f_0 (tonal frequency) and N_o (noise power) are constants for any particular filter fit. The final step involves adjusting the value of K to bring the output of the model into agreement with the threshold curve. The fitting procedure used in the present study is described in more detail in Chapter 3.

Purpose of this study

Critical ratio studies provide an estimate of the equivalent rectangular bandwidth (ERB) of the auditory filter if certain assumptions are met (Patterson et al. 1982), but tell you little about the filter shape. For the present study, a notched noise masking experiment was conducted to determine the functions relating masked tonal threshold to notch-width for tonal frequencies of 40, 60, 80, and 100 kHz. A single moderate level of masking noise was used. The dolphin's auditory filter shapes were calculated by fitting

the integral of the *roex(p,r)* filter shape to the function relating tonal threshold to notch width. The relationship between the auditory filter shapes and the previously reported critical ratios for the same experimental subject is discussed. The results were also extended to derive models of processing efficiency and auditory filtering in the species.

CHAPTER 2: METHODS

Subject

The subject was a 20 year-old female Atlantic bottlenose dolphin *(Tursiops truncatus)* located at the Hawaii Ocean Theater at Sea Life Park, Waimanalo, Hawaii. The hearing of the dolphin, Itsi Bitsi, was previously measured in psychoacoustic studies of pure-tone thresholds and masked tonal thresholds (Lemonds, 1997). As shown in Figure 5, she was found to have tonal sensitivity very similar to the animal tested by Johnson (1967). Critical ratios in *Tursiops* are shown in Figure 16; the present subject was found to have better masked tonal sensitivity than the other animals tested (Johnson, 1968; Au & Moore, 1990) over much of the high frequency range.

During the course of this experiment, the animal received a mixed diet of herring and smelt, the quantity of food provided over the course of a day was not contingent upon experimental performance. Itsi Bitsi weighed approximately 425 lbs while the experiment was conducted; she received approximately 16 lbs of fish per day and would receive up to 3 lbs in the course of a single experimental session.

Electronic apparatus

The electronic apparatus used in this experiment is shown in block diagram form in Figure 24. Tonal signals were generated using a Wavetek model 154 programmable waveform generator, and passed through an attenuator (up to 75 dB). Noise generating circuitry produced white masking noise which was also passed through an attenuator (up to 40 dB) and through a custom-built filter capable of bandpass or bandstop (notch)

filtering (this filter had cut-off slopes of approximately 200 dB/octave). For example, on the low frequency edge of a notch centered at 80 kHz, the output of the filter between 60 and 69 kHz dropped approximately 40 dB. This difference is 0.201 octaves (n) as determined in the equation $\log_{10}(f_1/f_2) = 0.301n$ (Moore, 1989) (and 40 dB/0.201 octaves = 200 dB/octave). Custom-built control circuitry gated the signal with an adjustable rise/fall time and controlled the duration of the signal and noise presentation. A control switch allowed for the selection of the trial type, noise alone (N) or signal plus noise ($S+N$), and for the initiation of the stimulus. The stimulus was passed through an equalizing filter to provide a flat frequency response from the piezoelectric transducer (WAU-3), amplified, and projected.

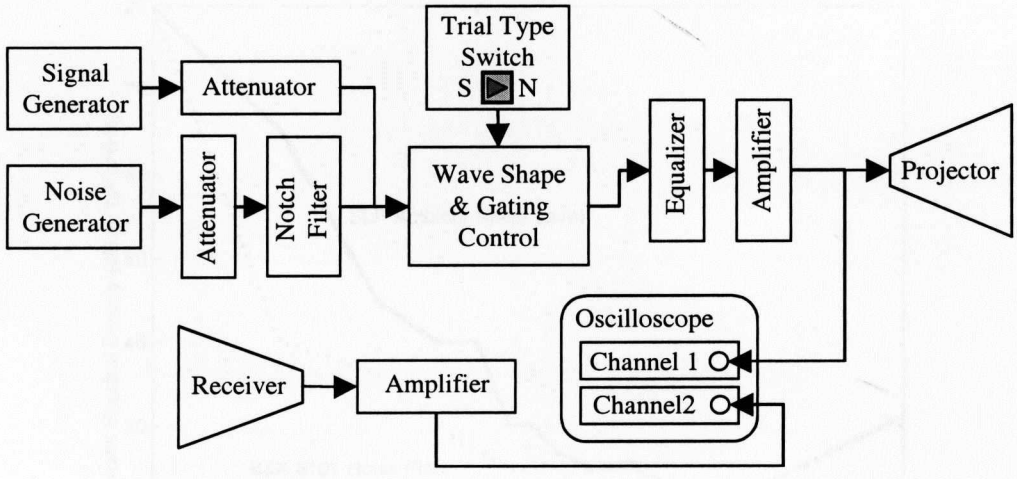


Figure 24. Electronic apparatus.

The projected signal was also monitored on one channel of a Tektronix 2235 100 MHz oscilloscope. Transducer output was monitored by another small piezoelectric hydrophone located 2 m in front of the projector (note the block diagram does not

indicate this relationship spatially). The received signal was amplified and monitored on the second channel of the oscilloscope.

Ambient noise levels

Ambient noise levels in the tank as measured with a Brüel & Kjer 8101 hydrophone are shown in Figure 25. Noise levels were very low, falling at or near the noise floor of the hydrophone between 20 and 50 kHz, and not exceeding thermal noise above 50 kHz. Thermal noise results from molecular agitation and limits the high frequency sensitivity of hydrophones (Richardson, Green, Malme, & Thomson, 1995). For comparison to studies conducted in open water conditions, ambient noise levels in the tank were equivalent to sea-state zero (Mellen, 1952).

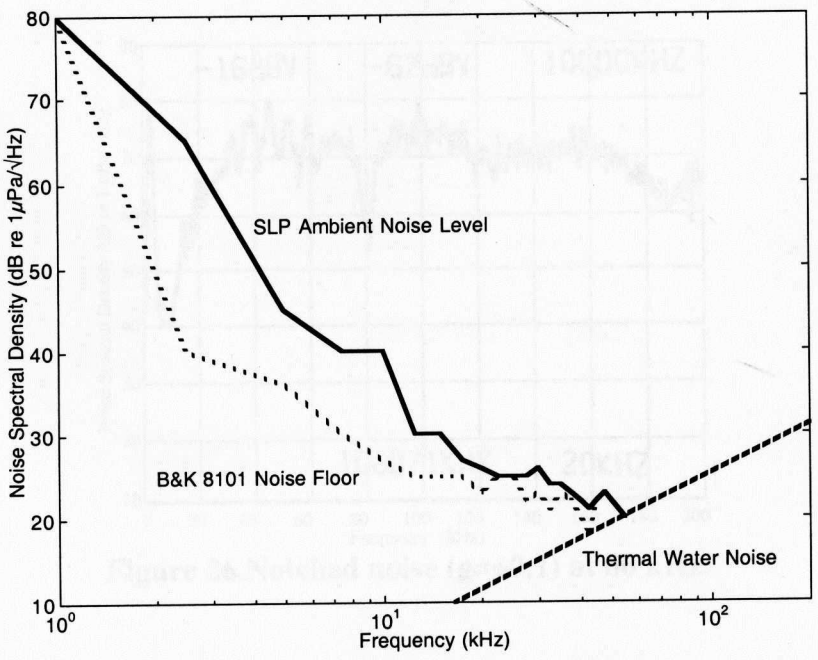


Figure 25. Ambient noise level in experimental tank.

Stimulus and masking noise

Levels of the test signals and the masking noise were measured at 2 m from the projector using an H-52 hydrophone and a Tektronix 7L5 spectrum analyzer. The unattenuated signal levels varied less than 5 dB for tonal signals at 40, 60, 80, and 100 kHz. Figure 26 through Figure 29 show examples of the notched noise used to the mask 80 kHz signal. The floor for most notches was between –35 and –40 dB relative to frequencies outside the bandstop. The spectrum of the masking noise outside of the notches was essentially flat (± 5 dB) at 70 dB re 1 μPa/$\sqrt{\text{Hz}}$ from 30 to 160 kHz. Both the notched-noise and pure-tone sound fields were free of distortion due to standing waves and/or surface reflections.

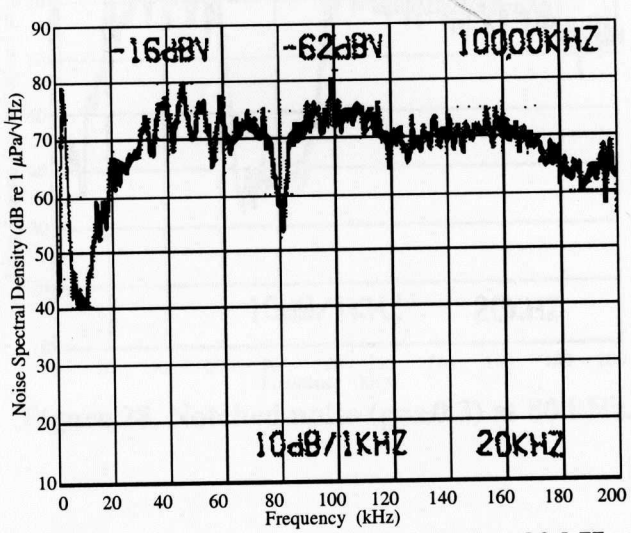


Figure 26.Notched noise (g=±0.1) at 80 kHz.

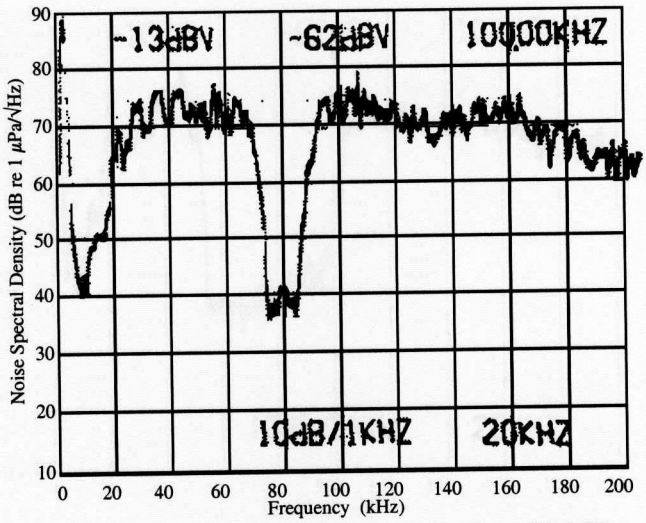


Figure 27. Notched noise (g=±0.2) at 80 kHz.

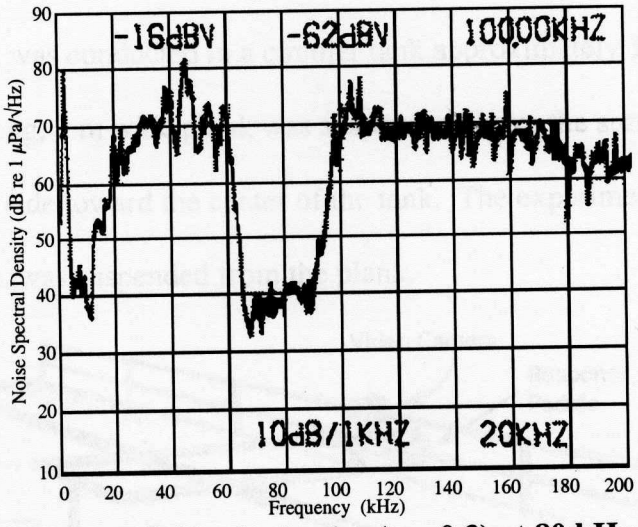


Figure 28. Notched noise (g=±0.3) at 80 kHz.

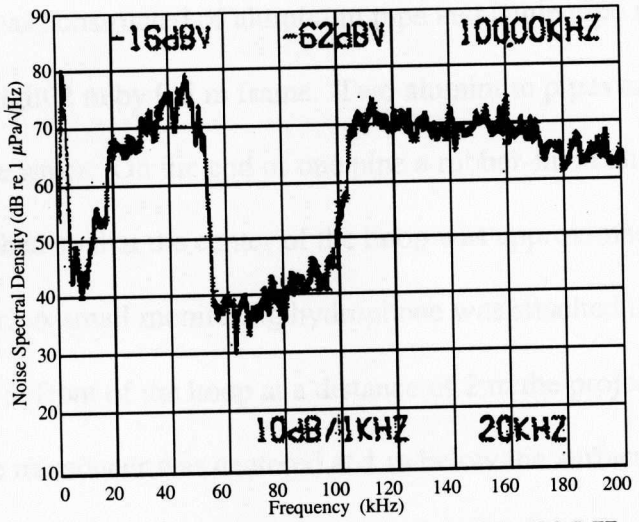


Figure 29. Notched noise (g=±0.4) at 80 kHz.

Experimental apparatus

The experiment was conducted in a circular tank approximately 15 m in diameter and 3 m deep. A 7 m long, 1 m wide plank was suspended above the surface of the water, extending from the side toward the center of the tank. The experimental apparatus, as shown in Figure 30, was suspended from the plank.

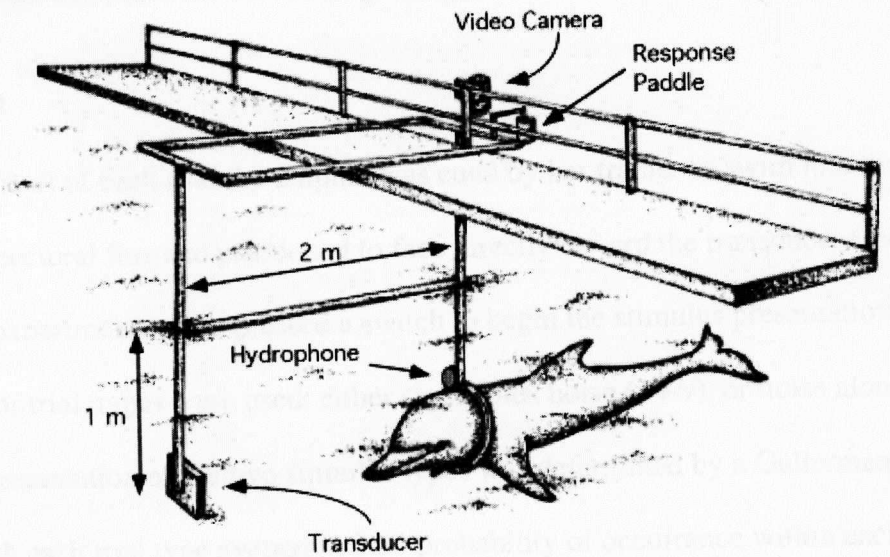


Figure 30. Experimental apparatus.

The apparatus was constructed of aluminum pipe and angle pipe fittings, assembled to form a rigid but light 2 m by 0.5 m frame. Two aluminum pipes extended from this frame down into the water. On the end of one pipe a rubber-lined aluminum hoop (45 cm diameter) was attached so that the center of the hoop was approximately 1 m below the surface of the water. A small monitoring hydrophone was attached to the top portion of the hoop. Directly in front of the hoop at a distance of 2 m the projector was attached to the other pipe. The transducer was centered at 1 m below the surface and was oriented to directly face the center of the hoop. To provide additional rigidity, a 2 m length of PVC pipe (3.75 cm diameter) was connected between the two pipes mentioned above, approximately 1.05 m above the centers of the hoop and projector, placing it just above the surface of the water. A small video camera mounted 2.5 m above the hoop was used to monitor the animal's position in the hoop. A 10 cm ball of foam, mounted on a 0.5 m length of PVC pipe extending laterally from the hoop side of the frame, hung above the surface of the water and served as a response paddle.

Procedure

At the start of each trial the dolphin was cued by her trainer to swim into the hoop as far as her pectoral fins and positioned to face directly toward the transducer (see Figure 30). The experimenter then pressed a switch to begin the stimulus presentation. Two stimulus, or trial, types were used: either signal plus noise (*S+N*), or noise alone (*N*). The order of presentation of the two stimulus types was determined by a Gellerman (1933) series, with each trial type averaging 0.50 probability of occurrence within each 10 trial block. For both trial types, the noise (*N*) was presented for 10 s, and this period was

defined as the duration of the trial. In *S+N* trials, the noise was turned on first, and after a 3 s delay, a 2 s signal (25 ms rise/fall time) was added to the noise. After the signal was over, the noise continued in isolation for the rest of the trial (an additional 5 s).

A go/no-go response procedure was used. The dolphin responded to the presence of tonal signals by backing out of the hoop and touching the response paddle (see Figure 30); this was a "go" response. If no tonal signal was presented, the subject was trained to stay positioned in the hoop until signaled by her trainer that the trial was finished, this was a "no-go" response. When the animal gave the correct response to a given stimulus, the trainer blew a whistle, or "bridge" sound, which was always followed by a fish reward. The whistle had been previously paired with fish rewards using operant conditioning techniques, and acted as a secondary reinforcer. Table 2 shows the possible outcomes when two stimulus conditions are crossed with two possible responses (go/no-go).

Table 2. Go/no-go decision matrix.

	Behavioral Response	
	Go	No-Go
Signal + Noise	Hit	Miss
Noise	False Alarm	Correct Rejection

A correct "no-go" response (correct rejection) occurred during noise only (*N*) trials when the animal stayed in the hoop for the duration of the *N* presentation (10 s) and remained in the hoop until the trainer sounded a whistle. A correct "go" response (hit) occurred during *S+N* trials when the dolphin touched the response paddle following the onset of the signal and before the noise ceased. A "no-go" response to a *S+N* trial is a

miss, while a "go" response to a *N* trial is called a false alarm. The trainer was unaware of the experimental condition until the subject began to respond, at which time the researcher reported the trial type to the trainer to allow correct feedback to the dolphin. False alarms were not rewarded by the trainer, and after a miss the trainer would tap a pipe on the edge of the tank indicating that the trial was over and no reward was being given.

An up/down staircase procedure (Robinson & Watson, 1973) was used to adaptively change the stimulus level as a function of the animal's response to the previous signal stimulus. The noise level on all trials was held constant throughout the experiment. The first 10 trials in a session were considered "warm-up" trials, with the signal level initially set 15 to 20 dB higher than estimated threshold. Following a hit, the signal was attenuated 3 dB. If the animal missed a signal in the warm-up, the signal level was held constant for the following *S+N* trial. If performance on the warm-ups was at 80% correct or better, the session would continue, otherwise the animal would be given a 5 minute time-out before starting the session over again.

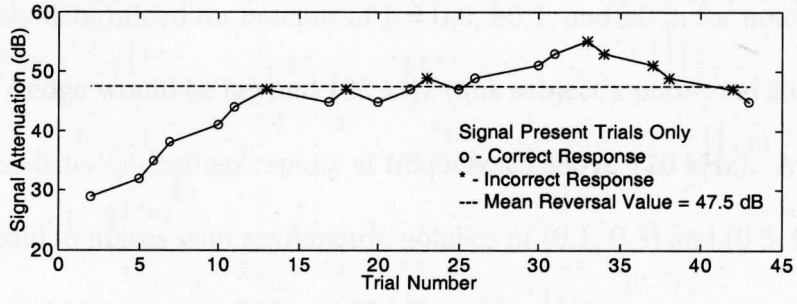


Figure 31. Threshold calculation based on reversal levels.

After the warm-up trials were completed, a hit resulted in the attenuation of the signal by 3 dB steps until the first miss. Following a miss, the signal was raised by 2 dB until

the next hit, and then lowered 2 dB until the next miss, and so on. Incorrect responses to noise only trials (false alarms) did not result in a change in signal level. The transition from miss to hit or hit to miss is a *reversal*, and sessions were ended after eight reversals were collected. An example of a staircase session is shown in Figure 31 as a plot of signal attenuation as a function of trial number. The status of a response, correct or incorrect, is shown by the shape of the data marker given in the legend. Only $S+N$ trials are included to illustrate the stimulus modification technique and response reversals.

The relative attenuation level at threshold was determined by taking the average attenuation value in dB of the eight reversals. To convert this relative attenuation level into signal level at threshold standardized as dB re 1 μPa, the attenuation level was subtracted from the unattenuated signal level of the same frequency.

In the present study, masked thresholds for tonal signals at 40, 60, and 80 kHz were determined for each of five symmetric notch conditions: $g = 0, \pm 0.1, \pm 0.2, \pm 0.3$, and ± 0.4 (where g represents the distance from the center frequency f_c to the edge of the notch (see Equation 7), hence the total notch width is $2g$). Masked tonal thresholds for the 100 kHz signal were only determined for notches of $g = 0.0, \pm 0.1$, and ± 0.2; for notches beyond $+0.2$ the high-f edge would be beyond 120 kHz (this subject's pure-tone audiogram indicates that sensitivity declines rapidly at frequencies above 120 kHz). Additionally, masked threshold to noises with asymmetric notches of (0.1, 0.3) and (0.3, 0.1) were calculated at tonal frequencies of 60 and 80 kHz to check for asymmetries in the filter shapes. Only one signal frequency and one notch size were used in a given session and the order of both was randomized; a data session was accepted if the standard deviation

of the reversal values was less than 5 dB within the session. Sessions at a given frequency and notch condition were run until two consecutive data sessions had mean reversal values within 3 dB; the reversal values for these two consecutive "matching" sessions were averaged to provide a threshold. Thus all the thresholds presented here are based on 16 reversals, eight from each of two matching sessions.

CHAPTER 3: RESULTS

Threshold as a function of notch width

Masked threshold results for symmetrical notches are listed in Table 3. Tonal thresholds were generally found to decrease as the notch in the noise widened. An exception to this occurred at 40 kHz; increasing the notch width from 0.0 to ±0.1 resulted in no change in tonal threshold. Note that no thresholds were determined for the 100 kHz tone beyond g=±0.2.

Table 3. Tonal threshold as a function of notch width.

(For symmetric notches. All thresholds in dB re 1 μPa)

	Tone (kHz)			
Notch (g)	40	60	80	100
0	95	96.9	100.1	105
±0.1	95	95	95.6	95.8
±0.2	78.3	82.1	79.1	79.9
±0.3	67.5	71	69.9	
±0.4	60.9	67.8	65.1	

Table 4. Asymmetric notch results.

	Tone (kHz)	
Notch	60	80
(0.1, 0.3)	87.1	87.8
(0.3, 0.1)	86.6	88.6

Results from the asymmetric notch conditions are shown in Table 4. In these conditions the notch is not centered on the tonal frequency, rather it is shifted either lower

or higher in frequency. Such conditions are denoted by the values of g at the low and high frequency edges of the notch; a notch of (0.1, 0.3) indicates that the low frequency edge is a distance of $0.1g$ below the center frequency (f_c), and the high frequency edge of is $0.3g$ above the center frequency. As an example, a (0.1, 0.3) notch about 80 kHz extends from 72 to 104 kHz. Tonal thresholds for asymmetrical notches were collected at 60 and 80 kHz. The differences in threshold between the notches shifted lower and higher in frequency deviated less than 1 dB, with the (0.1, 0.3) notch threshold slightly larger at 60 kHz, and the (0.3, 0.1) notch slightly larger at 80 kHz. To determine if these differences indicate underlying asymmetry in the auditory filters, a filter model bank simulation was conducted (Appendix C: Simulating filter asymmetry); filters with varying degrees and direction of asymmetry were simulated to perform the two asymmetrical notched noise tasks. If we assume that the ERBs of the filters at 60 and 80 kHz are between 10 and 20% of center frequency then the threshold differences between the two asymmetric notch conditions in the present experiment are consistent with those produced by symmetric filters in the simulation. The conclusion from the simulation is that the observed differences between the asymmetric notch thresholds at both 60 and 80 kHz are consistent with a bank of symmetric filters.

Estimating the filter shapes

As mentioned above, the auditory filter shapes at 60 and 80 kHz appear to be approximately symmetric based on simulations of the threshold results from the asymmetric notches. It is inferred that the filters at 40 and 100 will also be symmetric as there is no reason to assume otherwise. Filter shape results are presented based on fitting

the masked threshold curve determined over symmetrical notches to the right half of Equation 13. The best fitting *roex(p,r)* functions were determined in the following manner: (1) each threshold curve was expressed relative to the threshold level at the (0, 0) notch. This conversion aligns each threshold curve with the form of the *roex(p,r)* tail integral shown in Figure 23, with the output of the filter at g=0 equal to 0 dB. (2) Under the filter symmetry assumption, there are only two free parameters for fitting the tail integral of the auditory filter to the threshold curve: p and r. The fitting task can be described as finding the values of these free parameters that minimizes the sum of the squared error between the curve and the tail integral. The error surface of the three-space defined by p, r, and SSE, was calculated for each center frequency, and searched for the error minimum. (3) The free parameters, p and r, were then limited in range around the location of that minimum, and the surface was resampled (using a narrower range for the two free parameters essentially resampled the surface using a finer grid spacing). (4) The process of limiting the free parameter range, calculating the error surface, and finding the minimum was iterated until the parameters and error stabilized.

Filter shape results

The filter shape results for each center frequency are shown below. Three figures are shown for each signal frequency tested: the first shows the actual parameter error surface; the second shows the threshold curve and the predicted threshold curve based on the power spectrum model of masking, and third shows the auditory filter. For each threshold curve the actual reversal levels are shown as (+) and the mean reversal value as (o). The filter shapes at all four frequencies tested are shown together in Figure 44.

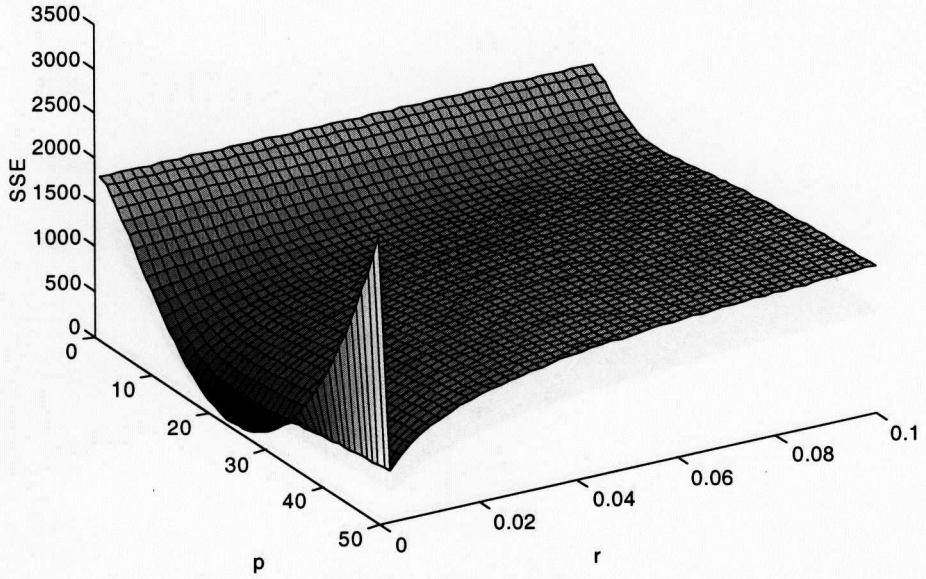


Figure 32. Error surface for *roex(p,r)* fit at 40 kHz.

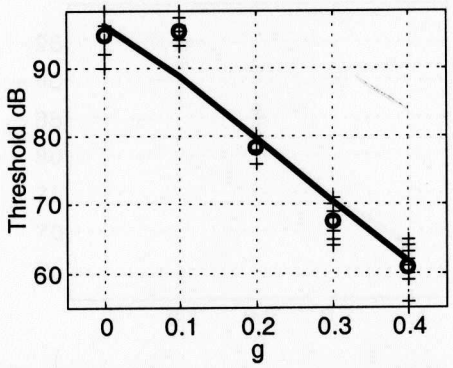


Figure 33. Thresholds and integral fit at 40 kHz.

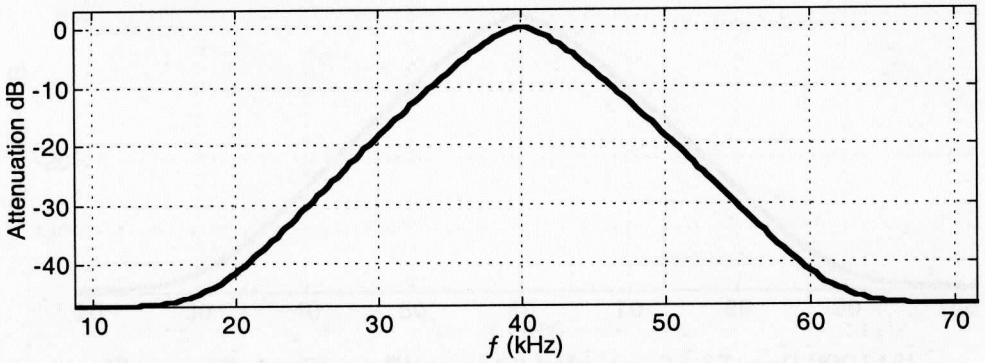


Figure 34. Auditory filter at 40 kHz. *(p=24.90, r=0.000020)*

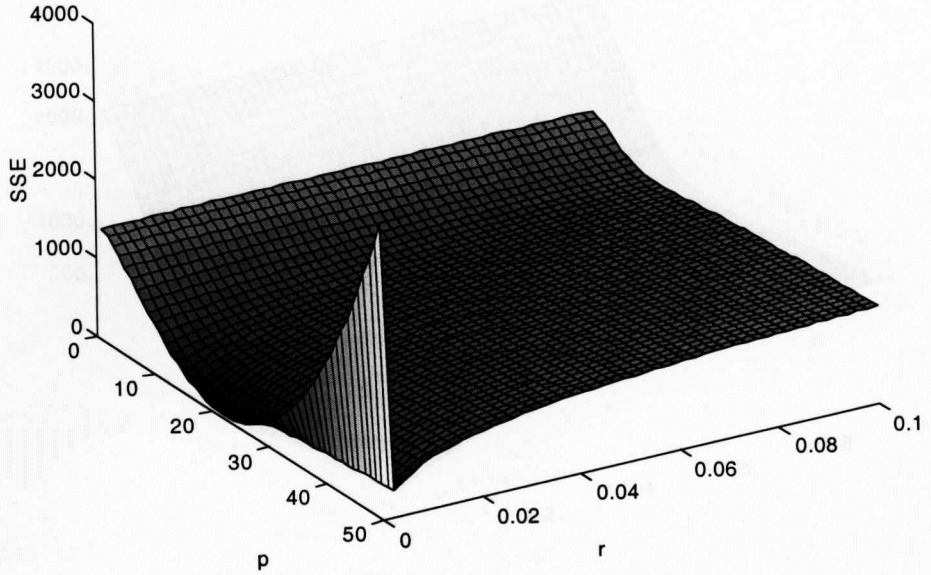


Figure 35. Error surface for *roex(p,r)* fit at 60 kHz.

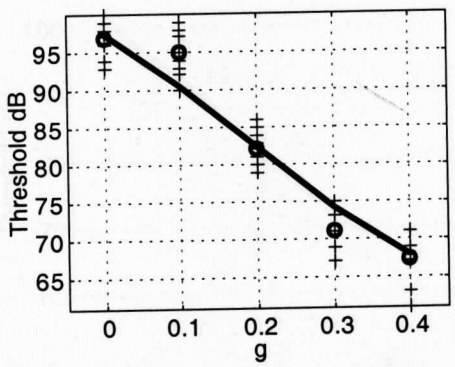


Figure 36. Thresholds and integral fit at 60 kHz.

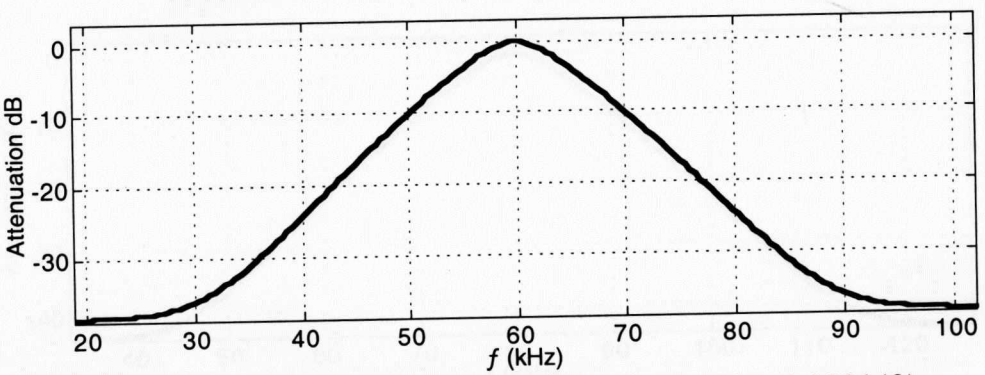


Figure 37. Auditory filter at 60 kHz. *(p=23.57, r=0.000143)*

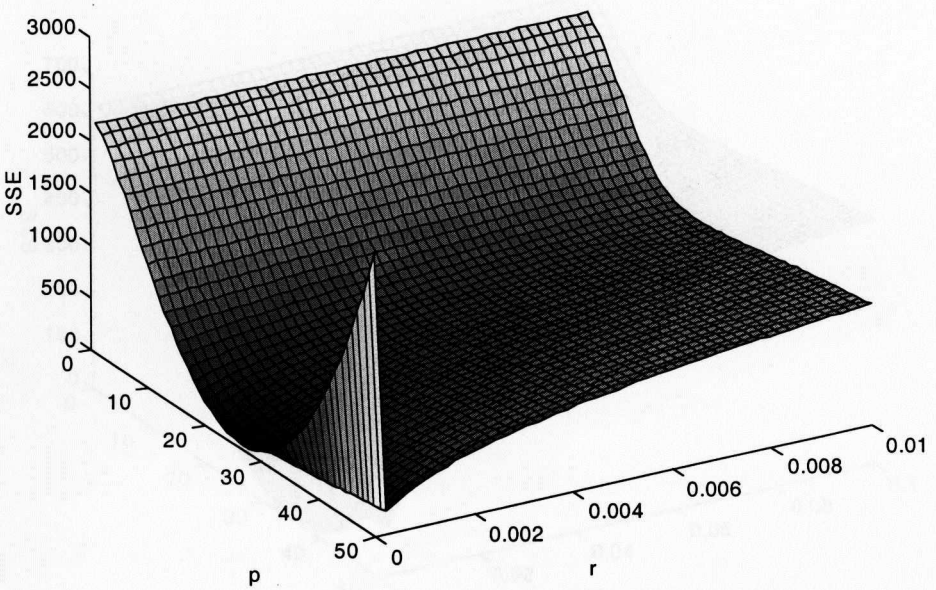


Figure 38. Error surface for *roex(p,r)* fit at 80 kHz.

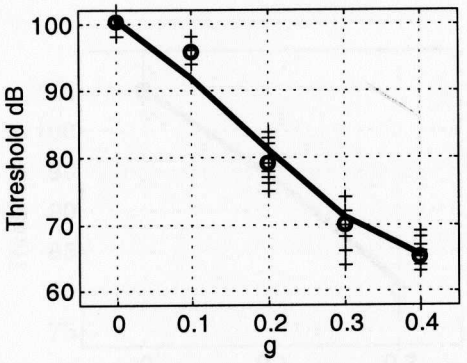


Figure 39. Thresholds and integral fit at 80 kHz.

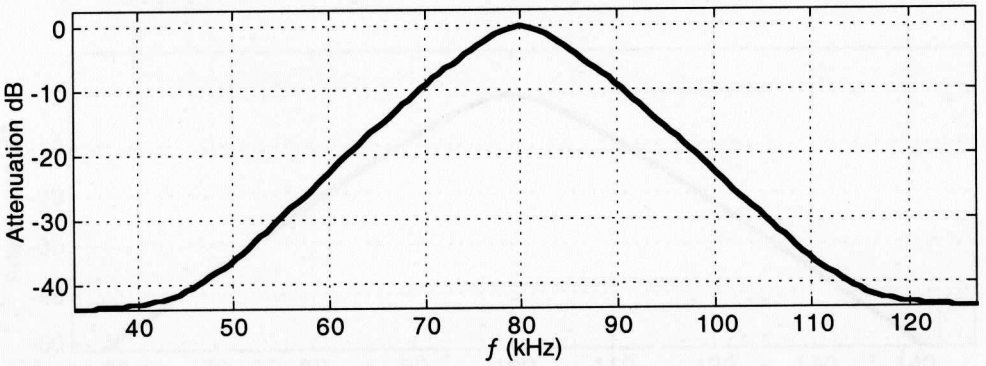


Figure 40. Auditory filter at 80 kHz. *(p=29.49, r=0.000041)*

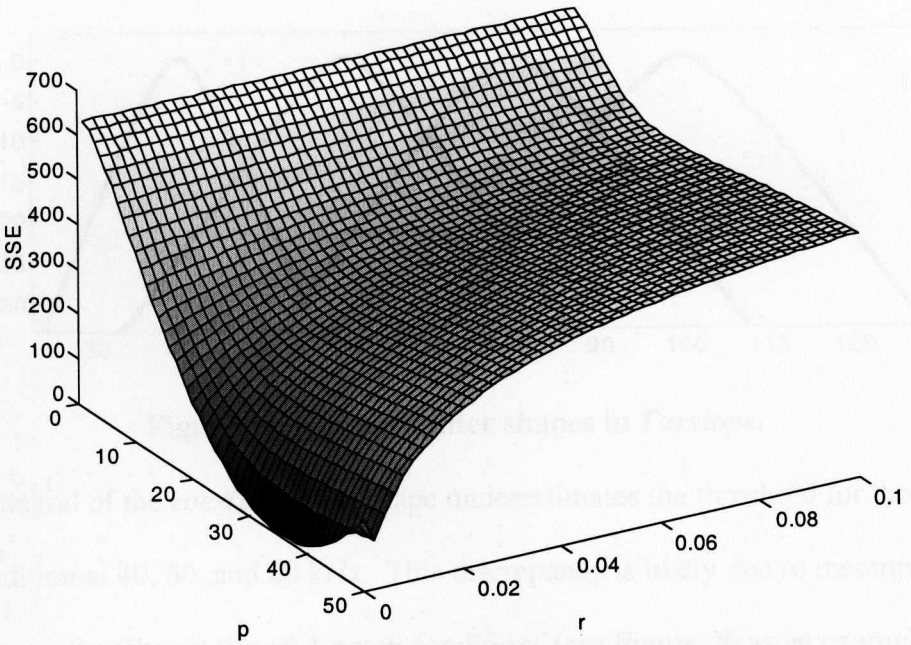


Figure 41. Error surface for *roex(p,r)* fit at 100 kHz.

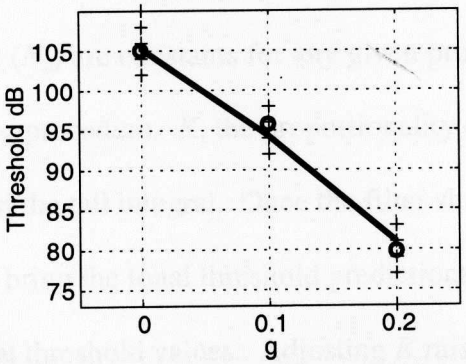


Figure 42. Thresholds and integral fit at 100 kHz.

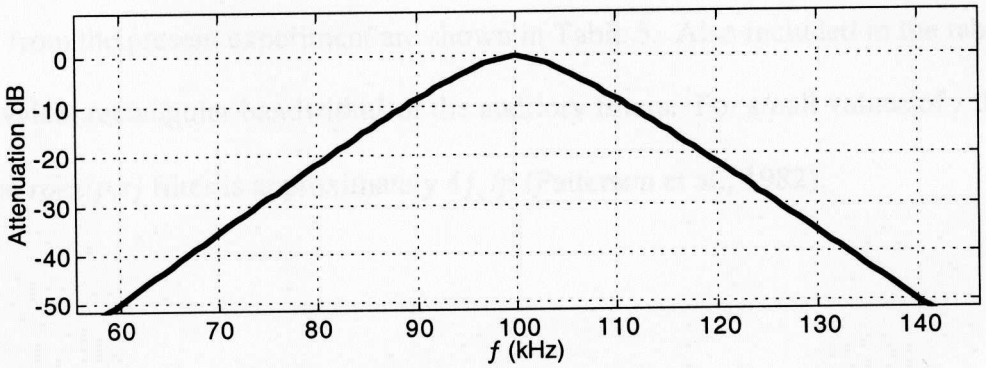


Figure 43. Auditory filter at 100 kHz. *(p=35.57, r=0.000000)*

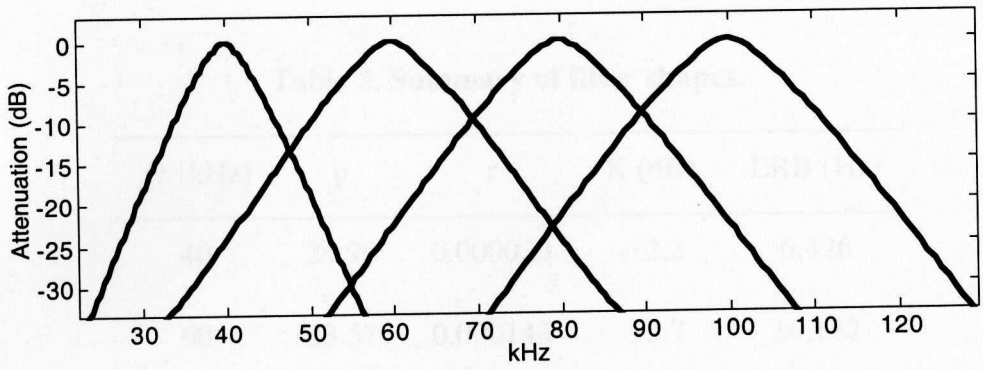


Figure 44. Auditory filter shapes in *Tursiops*.

The integral of the *roex(p,r)* filter shape underestimates the threshold for the ±0.1 notch condition at 40, 60, and 80 kHz. This discrepancy is likely due to incomplete attenuation by the filter at the ±0.1 notch conditions (see Figure 26 as an example).

Of the remaining terms in Equation 13, center frequency (f_0) and noise spectral density outside the notch (N_o) are constants for any given problem and therefore not adjusted during the fitting procedure. *K*, the proportionality constant, is the only "free" parameter not included in the tail integral. Once the filter shape has been determined, the value of *K* is adjusted to bring the tonal threshold predictions from Equation 13 into agreement with the actual threshold values. Adjusting *K* raises or lowers the model predictions uniformly at every value of g. The *roex* parameters; *p* and *r*, and the values of *K* from the present experiment are shown in Table 5. Also included in the table are the equivalent rectangular bandwidths of the auditory filters. For small values of *r* the ERB of the *roex(p,r)* filter is approximately $4f_c/p$ (Patterson et al., 1982).

Table 5. Summary of filter shapes.

*f*c (kHz)	p	r	K (dB)	ERB (Hz)
40	24.90	0.000020	-12.2	6,426
60	23.57	0.000143	-12.7	10,182
80	29.49	0.000041	-9.7	10,851
100	35.57	0.000000	-5.2	11,245

In summary, the *roex(p,r)* filter shape proved an adequate, though not perfect, fit for the collected threshold curves. The discrepancies between the model and the data are very difficult to explain in the context of the power spectrum model of notched noise masking. The estimated filter shapes are symmetric and vary in bandwidth from 11 to 16% of center frequency. Filter bandwidths were not found to increase as a constant proportion of center frequency (constant-Q) and are rather more finely tuned with increasing frequency; there is very little difference in bandwidth between the 60, 80, and 100 kHz filters. The reported *K* values indicate that the most efficient post-filter processing, better than –12 dB SNR, occurred at 40 and 60 kHz, with efficiency gradually decreasing with increasing frequency above 60 kHz.

CHAPTER 4: DISCUSSION

Results from a notched noise masked threshold experiment indicate that, at the moderate noise level used (70 dB re 1 μPa/$\sqrt{}$Hz), the auditory filters between 40 and 100 kHz in *Tursiops* are symmetric and generally well approximated by the *roex(p,r)* filter shape. In contrast to a constant-Q model, the function relating equivalent rectangular bandwidth to center frequency is nearly constant between 60 and 100 kHz; it increases only slightly as a function of frequency. Relative frequency selectivity, therefore, is enhanced with increasing frequency throughout this range. A function is presented that estimates ERB over the entire frequency range of hearing in the species; this function allows for the specification of computational models of the dolphin auditory system.

This study has extended our knowledge of the frequency processing capabilities in *Tursiops*. We can conclude that the auditory filter shapes in the high frequency range of hearing in *Tursiops* are generally similar in form to shapes reported in humans and other mammals at much lower frequencies. This conclusion agrees with the observation made by Au (1993) that auditory spectral filtering appears to function similarly in dolphins and humans. Further, we can conclude that, at least at the noise levels used in the present study, the auditory filters are approximately symmetric.

The present study provides a measure of processing efficiency occurring after the auditory filter (K). A sensible application of this information is for use in "correcting" critical ratio studies as critical ratios do not account for this variable but are certainly influenced by it; this application is discussed in greater detail in a following section.

This study explicitly describes the shape of auditory filters in *Tursiops* and therefore provides accurate measures of filter bandwidth. Auditory filter bandwidth does not seem to increase with frequency in the manner described by a constant-Q filter bank, as has been previously suggested (Au, 1993). Rather there seems to be a frequency region between roughly 60 and 100 kHz in which auditory filter bandwidth is essentially constant, indicating that the relative selectivity of the system is actually increasing with frequency in this range. This is in agreement with filter bandwidth predictions based on a previous critical ratio study with the dolphin used in this study (Lemonds et al., 1997). Accurate measures of filter bandwidth support the specification of models describing the frequency selectivity of the auditory system and allow for predictions of performance on auditory tasks that are assumed to be mediated by frequency selectivity; frequency discrimination is one such task.

Frequency discrimination and selectivity

Frequency discrimination refers to the ability to detect differences in the frequencies of tones, or in the frequency content of other sounds. A typical frequency discrimination experiment has a listener discriminate between two tones or a tone and a frequency modulated (FM) signal which are presented successively. The frequency discrimination limen (FDL) is reported as the smallest detectable difference between the two stimuli. Frequency difference limens from a tone vs. FM signal discrimination study conducted with *Tursiops* are shown in Figure 45 (Thompson & Herman, 1975). The figure is a plot of the frequency discrimination limen vs. frequency; in this figure the frequency limen is displayed as a percentage of tonal frequency. Human listeners show a similar pattern

(Moore & Glasberg, 1986), with frequency shifted down by a factor of ten. The smallest frequency limens reported for humans are just under 0.5% between 1 and 2 kHz, while the best results for *Tursiops* were under 0.3% between 4 and 16 kHz, and under 0.5% between 2 and 64 kHz.

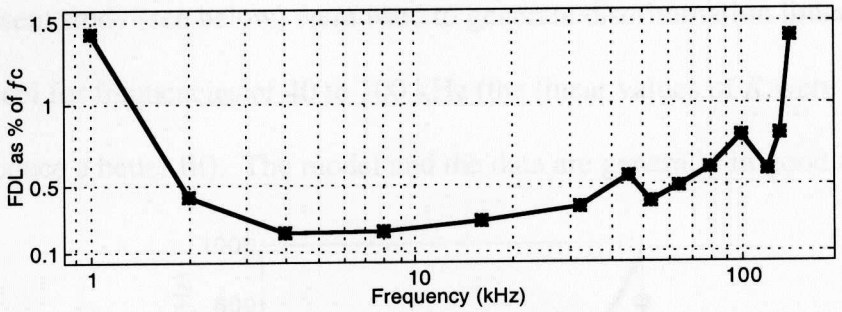


Figure 45. Frequency discrimination in *Tursiops*.

A prominent theory of frequency discrimination holds that the ability is completely determined by the efficiency and spectral filtering of the auditory system. The theory assumes that modulating the tonal frequency produces a detectable change in the distribution of excitation in the peripheral auditory system. In terms of auditory filters it is assumed the subject attends to the auditory filter that exhibits the largest change in excitation between the two stimuli. In a given frequency discrimination task, this is typically not the filter centered on the tonal frequency; rather the filter with the steepest slope at the tonal frequency will exhibit the greatest change. Though the details are beyond the scope of this paper, a model based on the *roex(p)* filter shape has been developed to predict the frequency limen as a function of tonal frequency, the ERB of an auditory filter at that frequency, and the efficiency of processing after the filter (Zwicker (1970) as cited in Moore & Glasberg, 1986).

Based on the ERB and K values from the present study, the model was used to predict the frequency discrimination results from *Tursiops*, as shown in Figure 46. Frequency discrimination limens, in Hz, are shown as open dark circles and the predictions of the model are shown as a solid line. The equations relating K and ERB to center frequency from the present study (see below) were used to generate discrimination limen estimates from the model for frequencies of 40 to 100 kHz (the linear values of K were multiplied by 1.7 to produce a better fit). The model and the data are generally in good agreement.

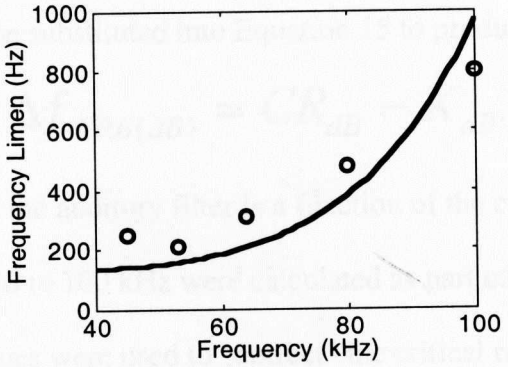


Figure 46. Frequency limens and model predictions for *Tursiops*.
(**o**) frequency limens (Thompson & Herman, 1975), (**-**) model predictions

Critical ratios revisited

This study allows for the comparison of two different measures of frequency selectivity, critical ratios and auditory filter shapes, for the same subject. The animal tested in this experiment was previously used in a study on critical ratios (Lemonds, 1997; Lemonds et al., 1997; and see Equation 6 and Figure 16). As described in Chapter 1, the critical ratio can be derived from the power spectrum model of masking (Equation 3) by assuming that K is equal to one. Beyond this questionable assumption, the logic of

the critical ratio is sound. Without making the assumption that K=1, Equation 6 can be shown as:

$$\Delta f_{ERB} = {}^{P_s}\!/_{KN_0}. \tag{14}$$

Expressing the terms in Equation 14 as dBs produces:

$$\Delta f_{ERB(dB)} = P_{S(dB)} - (K_{dB} + N_{0(dB)}). \tag{15}$$

The classic critical ratio (CR) is defined as $P_s - N_o$ when both terms are in dB. Hence, the classic critical ratio can be substituted into Equation 15 to produce:

$$\Delta f_{ERB(dB)} = CR_{dB} - K_{dB}, \tag{16}$$

which shows the ERB of the auditory filter is a function of the classic critical ratio and K.

The values of K for 40 to 100 kHz were calculated as part of the present experiment (see Table 5). These values were used to "correct" the critical ratios from the same animal (Equation 16); this should provide accurate estimates of auditory filter ERB. The results from this procedure are shown in Table 6, along with the filter shape results for comparison.

Table 6. ERB results for CR, "corrected" CR, and filter shape.

	Frequency (kHz)			
	40	60	80	100
Critical Ratio	29	29	32	31
K	-12.2	-12.7	-9.7	-5.2
"Corrected" CR	41.2	41.7	41.7	36.2
Filter Shape ERB	38.1	40.1	40.35	40.5

The ERB estimates produced by correcting the critical ratios are very similar to the ERB estimates calculated from the auditory filter shapes; the ERB estimates differ less than 5 dB between the two methods. These results lend support to the use of the corrected CR as an accurate predictor of filter shape ERB, and the high level of agreement between the results from the two approaches increases confidence in the accuracy of both studies.

Auditory modeling

Auditory filter shapes were estimated for frequencies between 40 and 100 kHz; it would be useful, however, to predict the frequency selectivity of the dolphin auditory system over a broad portion of the hearing range. A variety of provide evidence about the nature of the auditory filter bank in *Tursiops*. The first has already been discussed: the critical ratio measure. In the previous section it was shown that using *K* to correct critical ratios brought the ERB estimates from both methods into agreement. The only values available for *K* in *Tursiops* come from the present study and *K* as a function of frequency between 0.5 and 4 kHz has been reported for humans (Patterson, 1982). Based

on these two studies, the function relating K to frequency in *Tursiops* will likely increase gradually below 40 kHz and more rapidly above 100 kHz; the values at the extreme frequencies are unknown. (As an estimate assume that K will reach approximately 0 dB at both 1 and 140 kHz). Such a function relating K to frequency was used to fit the 3rd order polynomial that is shown in Figure 47 and expressed as:

$$K = -0.002659x^3 + 0.0079x^2 - 0.599x + 1.0461 \quad (17)$$

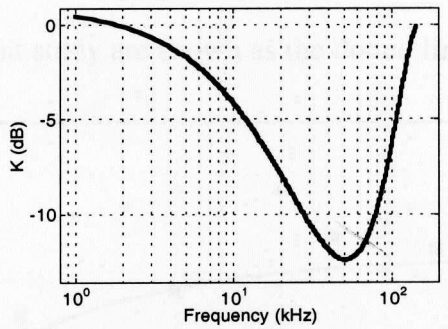


Figure 47. Third order polynomial estimate of *K*.

The polynomial in Equation 17 was used to "correct" the critical ratio data available for *Tursiops*; this correction allows the critical ratio to be interpreted as a measure of the auditory filter bandwidth. The corrected data are shown as (+) for Johnson (1968), (*) for Au & Moore (1990), and (o) for Lemonds et al. (1997) in Figure 48. Note that the studies of Johnson (1968) and Au & Moore (1990) were conducted with animals for which we have no direct measure of efficiency; using results from the present subject to "correct" results from other animals simply provides a rough estimate of filter bandwidth.

The second source of filter bandwidth information comes from the critical bandwidth study of Au & Moore (1990). In this study, tones were masked by bandpass noises

centered at the tonal frequency. A variety of noise bandwidths were used, and masking typically increased with increasing bandwidth up to a point. The "critical" bandwidth refers to that bandwidth past which increasing bandwidth ceases to increase tonal masking (Fletcher, 1940). A major difficulty in interpreting the results of this study from the perspective of the power spectrum model is off-frequency listening: at most noise bandwidths, the highest SNR will be achieved by listening to an off-frequency filter. Despite this, the measure can be interpreted as a measure of auditory filter ERB. Critical bands for *Tursiops* are shown as (x) in Figure 48 (Au & Moore, 1990). Additionally, the ERB results from the present study are shown as the dotted line in the figure.

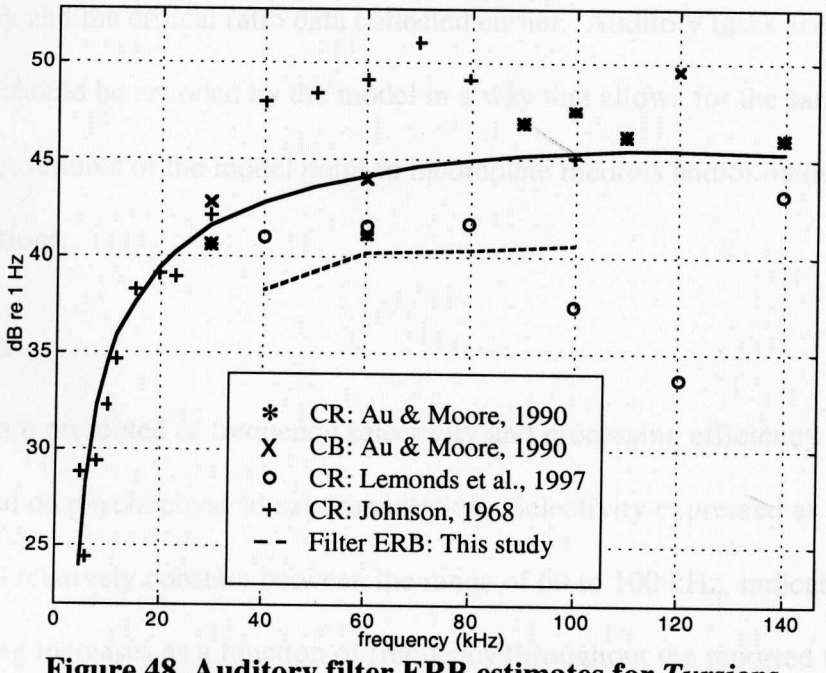


Figure 48. Auditory filter ERB estimates for *Tursiops*.

In addition to the data points, a logistic curve is shown which describes the general trend of the ERB estimates as a function of frequency. The function had to have the following properties: (1) range between roughly 20 and 45 dB, (2) approximately flat

from 60 to 140 kHz, and (3) logarithmically increasing before the flat section. The following equation is not intended to have theoretical significance; it was merely found to fit the results well. The equation expresses auditory filter ERB (Hz) as a function of center frequency (x in kHz):

$$ERB_{Hz} = 0.0039x^3 - 3.593x^2 + 703.316x - 3953.6 \quad (18)$$

The specification of a general filter shape, the *roex(p,r)*, combined with a function relating ERB to center frequency constitutes a model of frequency selectivity in *Tursiops*. The utility of such a model is its ability to predict the auditory capabilities in the species; for example, this model can be used to predict the notched noise masking data collected for this study and the critical ratio data collected earlier. Auditory tasks accomplished by the dolphin should be encoded by the model in a way that allows for the same level of performance; failures of the model point to incomplete theories and/or incorrect implementations.

Conclusions

Models are presented of frequency selectivity and processing efficiency in *Tursiops* that are based on psychophysical experimentation. Selectivity expressed as auditory filter bandwidth is relatively constant between the range of 60 to 100 kHz, indicating that relative tuning increases as a function of frequency throughout the reported frequency range. Efficiency, measured as the signal-to-noise ratio at the output of an auditory filter required for detection, was maximal at 40 and 60 kHz, and decreased gradually with increasing frequency. The efficiency measure can be used to "correct" previous measures of critical ratios for accurate prediction of auditory filter bandwidth. A model

based on both filter bandwidth and efficiency was shown to be in good agreement with frequency discrimination capabilities in the species. Auditory filters in *Tursiops* at the moderate noise level used were well approximated by a symmetric *roex(p,r)* filter shape which is commonly used to describe frequency selectivity in human and other mammalian species; this indicates that the mechanisms underlying frequency selectivity in *Tursiops* appear to be similar to those in other mammalian species.

APPENDIX A: BANDWIDTHS IN Hz AND dB

The decibel (dB) is a logarithmic unit which is used in a number of scientific disciplines. In all cases it is used to compare some quantity with some reference value. In acoustics, the concern is often with an intensity ratio expressed in decibels as:

$$10log_{10}(\text{intensity/reference intensity})$$

However, pressure is easier to measure than intensity. The square of pressure is equivalent to intensity, which provides the common expression of sound pressure level (SPL), expressed in decibels as:

$$\text{SPL} = 20log_{10}(\text{pressure/reference pressure})$$

which is equivalent to:

$$10log_{10}\ (\text{pressure}^2/\text{reference pressure}^2).$$

For underwater acoustics, the reference pressure is standardized as 1 μPa (micro-Pascal), and levels are denoted as dB re 1 μPa (where re is shorthand notation for referenced). Quantities of interest in acoustics (and psychoacoustics) often exhibit such huge ranges of variation that a dB scale is more convenient than a linear scale and more psychophysically relevant.

For the purposes of this paper, bandwidths as expressed in Hz can be converted into dB re 1 Hz, and vice versa. The formula relating the two is:

$$\text{bandwidth in dB} = 10log_{10}(\text{bandwidth in Hz/1 Hz})$$

For example, to convert a bandwidth of 300 Hz to dB, take the log (base 10) of 300 and multiply the value by 10, which equals 24.77 dB. To go back to Hz, divide the dB value by 10 and take the antilog (base 10) of the product

APPENDIX B: CONSTANT-Q FUNCTIONS

The quality factor (or Q) of a bandpass filter is defined as the ratio of center frequency (f_c) to bandwidth (Δf) such that $Q = f_c/\Delta f$. The bandwidth of a filter is typically determined at the half-power or 3-dB cutoff, although 10-dB and other cutoff levels can be used (Yost, 1993). These values would be reported as Q_3 or Q_{10} respectively. Reporting Q values allows comparisons to be made across filters with different center frequencies.

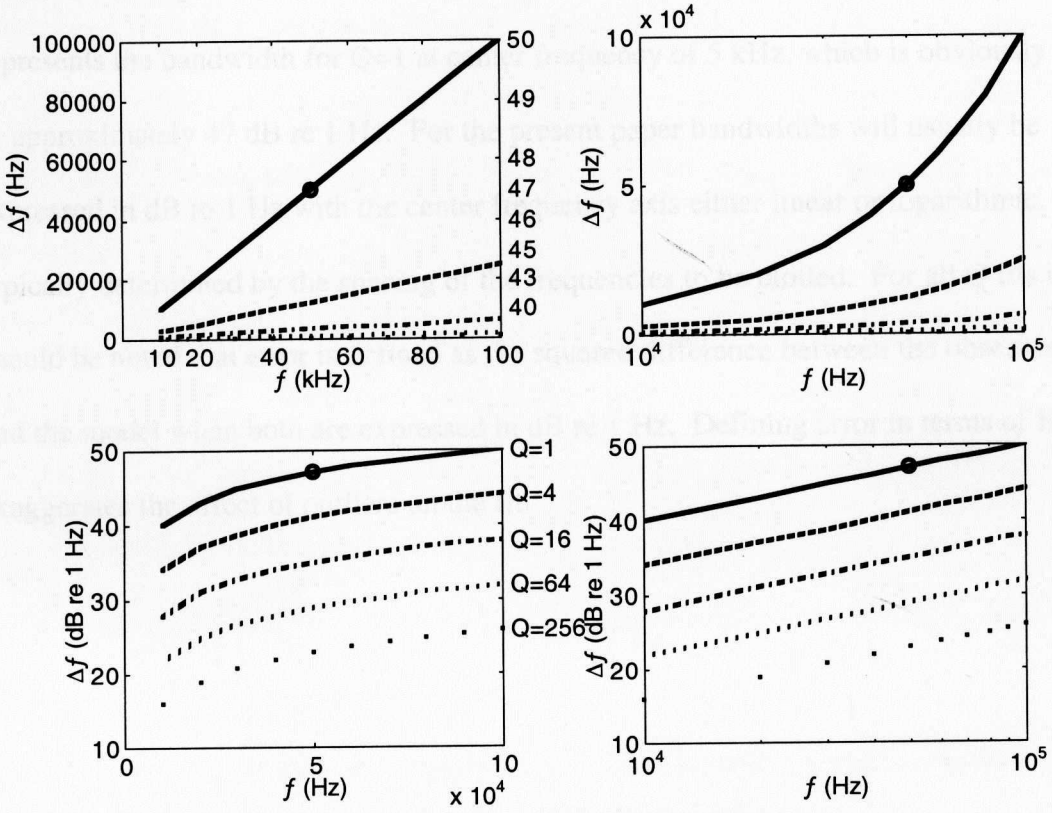


Figure 49. Plots of constant-Q functions.

A constant-Q filter bank contains filters of a constant-Q value at every center frequency. Many of the behavioral data reported in this paper have been fitted using a constant-Q model. Some confusion can arise when displaying these functions due to

variations in how the bandwidths are reported (as dB's or in Hz) and the scaling of the center frequency axis. Figure 49 shows four possible views of five constant-Q functions. The abscissa in the first column of subplots shows frequency (from 10 to 100 kHz) on a linear scale, while the second column shows frequency on a logarithmic axis. Similarly, the ordinate in the first row of subplots shows bandwidth in Hz in a linear fashion, while the second row shows bandwidth in dB re 1 Hz, defined as $10log_{10}(\Delta f$ re 1 Hz). There are five curves shown in each figure with Q values of 1, 4, 16, 64, and 256, and they are labeled in the figure on the bottom right hand corner. The circle on all four plots represents the bandwidth for Q=1 at center frequency of 5 kHz, which is obviously 5 kHz or approximately 47 dB re 1 Hz. For the present paper bandwidths will usually be expressed in dB re 1 Hz with the center frequency axis either linear or logarithmic, typically determined by the spacing of the frequencies to be plotted. For all Q fits it should be noted that error is defined as the squared difference between the observed data and the model when both are expressed in dB re 1 Hz. Defining error in terms of Hz exaggerates the effect of outliers on the fit.

APPENDIX C: SIMULATING FILTER ASYMMETRY

Results from the asymmetric notch conditions are reported in Table 4; the differences between masked tonal thresholds at both 60 and 80 kHz for the (0.1, 0.3) and (0.3, 0.1) notches were less than 1 dB. A *roex(p)* filter bank simulation was conducted to determine if the observed results were consistent with a bank of asymmetrical auditory filters, and if so, to determine the direction and extent of asymmetry. Specifically, the simulation modeled the *differences* between the tonal thresholds for the two asymmetric notch conditions as a function of two dimensions: filter asymmetry and filter bandwidth.

Computational simulations force model constraints and assumptions to be explicitly coded. For this simulation the max SNR assumption was adopted (see the sections on off-frequency listening and filter asymmetry in Chapter 1). For ease of computation, the modeled masking task consisted of a 100 Hz tone masked by both a (0.1, 0.3) and a (0.3, 0.1) notched noise; results from the simulation can be extrapolated to any tonal frequency, as will be explained below. The range of the filter bank was constrained to have characteristic frequencies between the low and high frequency cut-offs of the notch in the masking noise, which in this case were 70 and 110 Hz in the (0.3, 0.1) condition and 90 and 130 Hz in the (0.1, 0.3) condition.

Two dimensions were crossed in the model: filter asymmetry and bandwidth. Due to the relatively limited range of the characteristic frequencies modeled in the bank and the results of an earlier critical ratio study with the same animal (Lemonds et al., 1997), it was assumed that the filter bandwidths over each of these ranges, both around 60 and 80 kHz, would be approximately constant. To allow extrapolation of the simulation results

to the experimental conditions, filter bandwidth was represented as ERB expressed as a percentage of center frequency; this variable was called ERBn. The range of the ERBn variable was constrained between 0.1 and 0.5, or between 10% and 50% of the tonal frequency. Hence the two major assumptions of this simulation are (1) max SNR, and (2) constant bandwidth filters. Filter asymmetry was represented as the ratio of area under the low frequency branch of the filter (the area of the filter below the characteristic frequency) to the total filter area; this variable was called LTR (for Low area to Total area Ratio). The LTR variable was constrained between 0.2 and 0.8. Note that an LTR of 0.5 indicates a perfectly symmetric filter, as one half of the total area of the filter lies below the characteristic frequency.

The actual filter shapes modeled were back to back *roex(p)* functions. In order to determine the *p* values for the lower and upper branches, the areas under the lower and upper branches were calculated:

$$\begin{aligned} Area_{low} &= ERBn f_t LTR \\ Area_{high} &= ERBn f_t (1 - LTR) \end{aligned} \tag{19}$$

where f_t is tonal frequency used to convert ERBn into Hz. Once the two areas are known, the values of *p* for the upper and lower branches can be calculated for a filter with characteristic frequency f_c:

$$\begin{aligned} p_{low} &= {}^{2f_c}\!/_{Area_{low}} \\ p_{high} &= {}^{2f_c}\!/_{Area_{high}} \end{aligned} \tag{20}$$

The simulation consisted of crossing bandwidth (ERBn) with filter asymmetry (LTR) and determining the difference between tonal threshold between the two asymmetric notched masking noise conditions. Figure 50 shows a simulated filter bank (ERBn=0.2, LTR=0.5) performing the (0.3, 0.1) masking task. The top subplot shows the simulated filter bank; the bottom three subplots show, from left to right, the signal power passed by each filter, the noise power passed by each filter, and the SNR at the output of each filter. Under the max SNR assumption, a listener would attend to the filter with the highest SNR at output, which in this example corresponds to the filter with characteristic frequency of 95 Hz, which has an SNR of 16.74 dB.

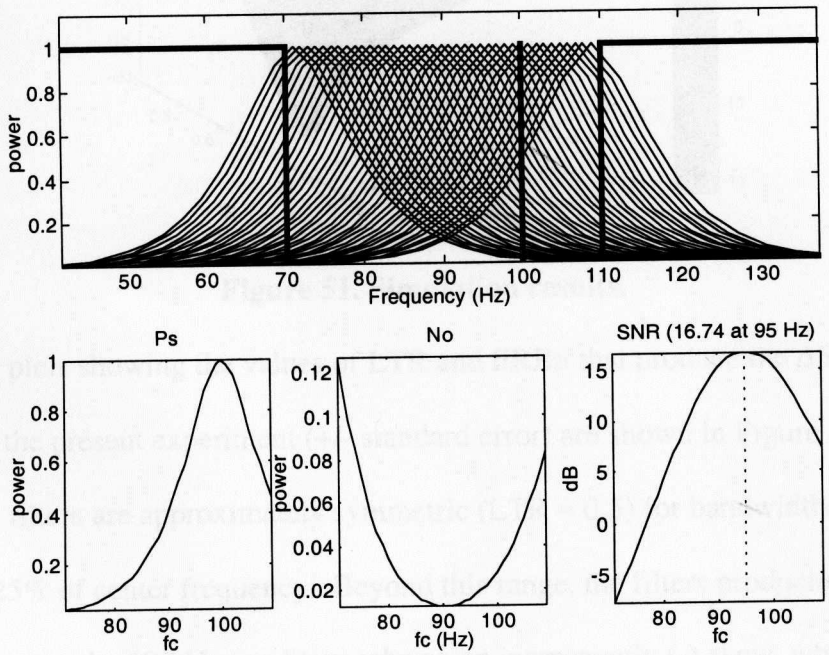


Figure 50. Simulated filter bank.

The simulation then runs the exact same filter bank for the mirror notch condition, in this case the (0.1, 0.3) notch, and determines which filter has the highest SNR. The difference between the two SNRs, one from the (0.3, 0.1) and one from the (0.1, 0.3)

condition, is calculated and returned as a variable ΔSNR. In this manner a value of ΔSNR is determined for every combination of ERBn and LTR. The results of the simulation are shown in Figure 51 as a plot of ΔSNR as a function of both ERBn and LTR. The gray-scale bar to the right of the figure provides an additional indication of ΔSNR.

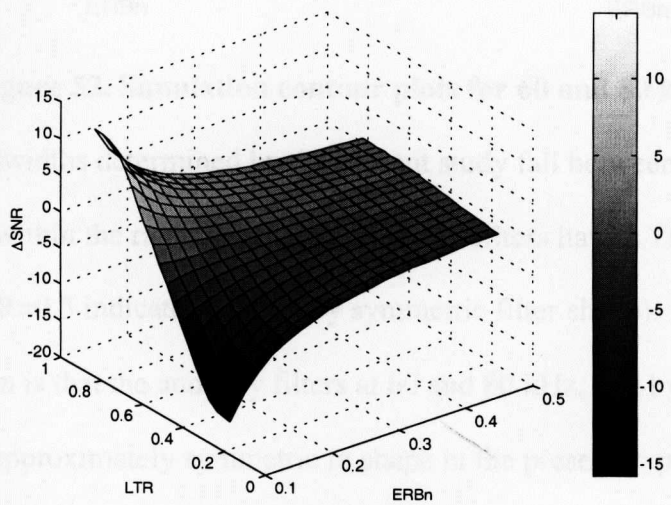


Figure 51. Simulation results.

Contour plots showing the values of LTR and ERBn that produce the ΔSNR values observed in the present experiment (+/- standard error) are shown in Figure 52. In this model, both filters are approximately symmetric (LTR ≈ 0.5) for bandwidths ranging from 10 to 25% of center frequency. Beyond this range, the filters producing the observed data for the 60 kHz condition take on an increasingly (-) skew, while the filters fit to the 80 kHz results take on an increasingly (+) skew.

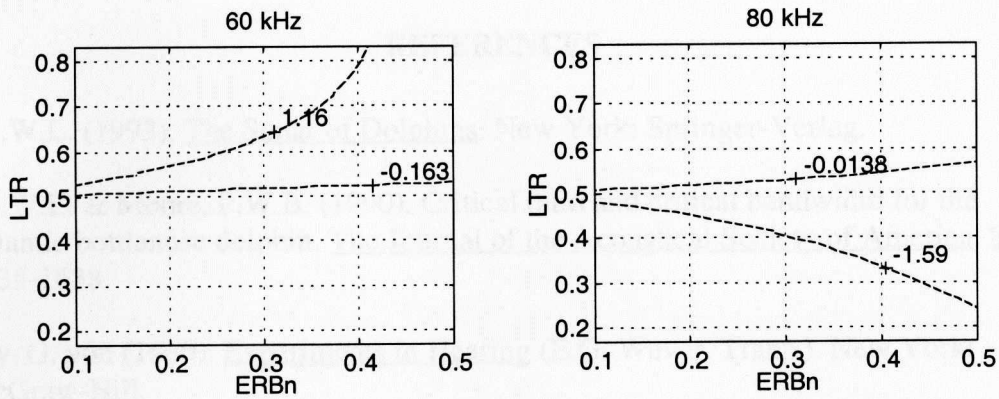


Figure 52. Simulation contour plots for 60 and 80 kHz

The filter bandwidths determined by the present study fall between 11 and 17% of center frequency, within the range where the predicted filters have LTR values between 0.45 and 0.55 (LTR=0.5 indicates a perfectly symmetric filter shape). The conclusion from the simulation is that the auditory filters at 60 and 80 kHz, and by extension at 40 and 100 kHz, are approximately symmetric in shape in the present experiment.

REFERENCES

Au, W.W.L. (1993). The Sonar of Dolphins. New York: Springer-Verlag.

Au, W.W.L. & Moore, P.W.B. (1990). Critical ratio and critical bandwidth for the Atlantic bottlenose dolphin. The Journal of the Acoustical Society of America, 88(3), 1635-1638.

Békésy, G. von (1960). Experiments in Hearing (E.G. Wever, Trans.). New York: McGraw-Hill.

Dallos, P. (1992). Neurobiology of cochlear hair cells. In Y. Cazals, K. Horner, & L. Demany (Eds.), Auditory Physiology and Perception (pp. 3-17). Oxford: Pergamon Press.

Denes, P.B. & Pinson, E.N. (1993). The Speech Chain: The Physics and Biology of Spoken Language (2nd ed.). New York: W.H. Freeman and Company.

Durrant, J.D., & Lovrinic, J.H. (1995). Bases of Hearing Science (3rd ed.). Baltimore: Williams & Wilkins.

Fay, R.R. (1988). Hearing in Vertebrates: A Psychophysical Databook. Winnetka, Illinois: Hill-Fay Associates.

Fletcher, H. (1940). Auditory patterns. Reviews of Modern Physics, 12, 47-65.

Gellerman, L.W. (1933). Chance orders of alternating stimuli in visual discrimination experiments. Journal of Genetic Psychology, 42, 206-208.

Hawkins, J.H. & Stevens, S.S. (1950). The masking of pure tones and of speech by white noise. Journal of the Acoustical Society of America, 22, 6-13.

Helmholtz, H.L.F. von (1863). Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik (1st ed.). Braunschweig: F. Vieweg.

Johnson, C.S. (1967). Sound detection thresholds in marine mammals. In W. Tavolga (Ed.), Marine BioAcoustics (pp. 240-260). New York: Pergamon Press.

Johnson, C.S. (1968). Masked tonal thresholds in the bottlenosed porpoise. The Journal of the Acoustical Society of America, 44(4), 965-967.

Johnson, C.S., McManus, M.W., & Skaar, D. (1989). Masked tonal hearing thresholds in the beluga whale. The Journal of the Acoustical Society of America, 85(6), 2651-2654.

Kelly, J.P. (1991). Hearing. In E.R. Kendel, J.H. Schwartz, & T.M. Jessell (Eds.), Principles of Neuroscience (3rd ed.) (pp. 481-499). New York: Elsevier.

Ketten, D.R. & Wartzok, D. (1990). Three-dimensional reconstructions of the dolphin ear. In J. Thomas & R. Kastelein (Eds.), Sensory Abilities of Cetaceans (pp. 81-105). New York: Plenum Press.

Ketten, D.R. (1992). The marine mammal ear: Specializations for aquatic audition and echolocation. In D. Webster, R. Fay, & A. Popper (Eds.), The Evolutionary Biology of Hearing (pp. 717-754). New York: Springer-Verlag.

Ketten, D.R. (1994). Functional analyses of whale ears: Adaptations for underwater hearing. I.E.E.E. Proceedings in Underwater Acoustics (Vol. 1) (pp. 264-270).

Lemonds, D.W. (1997). Masked hearing capabilities in an Atlantic bottlenose dolphin *(Tursiops truncatus)*. Master's thesis, University of Hawai'i at Mãnoa, Honolulu, HI. (UMI Number: 1387918).

Lemonds, D.W., Au, W.W.L., Nachtigall, P.E., Vlachos, S., & Roitblat, H.L. (1997). Auditory frequency selectivity and masked hearing capabilities in an Atlantic bottlenose dolphin. Journal of the Acoustical Society of America, 102 (5, pt. 2), 3102.

Luce, R.D. (1993). Sound & Hearing: A Conceptual Introduction. Hillsdale, N.J.: Lawrence Erlbaum.

Mellen, R.H. (1952). The thermal-noise limit in the detection of underwater acoustic signals. The Journal of the Acoustical Society of America, 24(5), 478-480.

McCormick, J.G., Weaver, E.G., Palin, G., & Ridgway, S.H. (1970). Sound conduction in the dolphin ear. Journal of the Acoustical Society of America, 48, 1418-1428.

Moore, B.C.J. (1989). An Introduction to the Psychology of Hearing (3rd ed.). London: Academic Press.

Moore, B.C.J. & Glasberg, B.R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. The Journal of the Acoustical Society of America, 74(3), 750-753.

Moore, B.C.J. & Glasberg, B.R. (1986). The role of frequency selectivity in the perception of loudness, pitch and time. In B.C.J. Moore (Ed.), Frequency Selectivity in Hearing (pp.251-308). London: Academic Press.

Moore, B.C.J., & Glasberg, B.R. (1987). Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns. Hearing Research, 28, 209-225.

Moore, B.C.J. & Glasberg, B.R. (1990). Derivation of auditory filter shapes from notched-noise data. Hearing Research, 47, 103-138.

Patterson, R.D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C., & Allerhand, M. (1992). Complex sounds and auditory images. In Y. Cazals, K. Horner, & L. Demany (Eds.), Auditory Physiology and Perception (pp. 429-446). Oxford: Pergamon Press.

Patterson, R.D. & Moore, B.C.J. (1986). Auditory filters and excitation patterns as representations of frequency resolution. In B.C.J. Moore (Ed.), Frequency Selectivity in Hearing (pp.123-177). London: Academic Press.

Patterson, R,D., Nimmo-Smith, I. (1980). Off-frequency listening and auditory-filter asymmetry. The Journal of the Acoustical Society of America, 67(1), 229-245.

Patterson, R.D.., Nimmo-Smith, I., Weber, D.L., & Milroy, R. (1982). The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold. The Journal of the Acoustical Society of America, 72(6), 1788-1803.

Pickles, J.O. (1986). The neurophysiological basis of frequency selectivity. In B.C.J. Moore (Ed.), Frequency Selectivity in Hearing (pp.51-122). London: Academic Press

Richardson, W.J., Greene, Jr., C.R., Malme, C.I., & Thomson, D.H. (1995). Marine Mammals and Noise. San Diego: Academic Press.

Robinson, D. & Watson, C. (1973). Psychophysical methods in modern psychoacoustics. In J.V. Tobias (Ed.), Foundations of Modern Auditory Theory (pp. 99-131). New York: Academic Press.

Shailer, M.J., Moore, B.C.J., Glasberg, B.R., Watson, N., & Harris, S. (1990). Auditory filter shapes at 8 and 10 kHz. The Journal of the Acoustical Society of America, 88(1), 141-148.

Sivian, L.J. & White, S.D. (1933). On minimum audible fields. The Journal of the Acoustical Society of America, 4, 288-321.

Slaney, M. (1998). Auditory Toolbox: A MATLAB Toolbox for Auditory Modeling Work (version 2) [Computer programs and manual]. Interval Research Corporation. (Available from <malcolm@interval.com>; Technical report # 1998-010).

Thomas, J.A., Pawloski, J.L., & Au, W.W.L. (1990). Masked hearing abilities in a false killer whale *(Pseudorca crassidens).* In J. Thomas & R. Kastelein (Eds.), Sensory Abilities of Cetaceans (pp. 395-404). New York: Plenum Press.

Thompson, R.K.R. & Herman, L.M. (1975). Underwater frequency discrimination in the bottlenose dolphin (1-140 kHz) and the human (1-8 kHz). The Journal of the Acoustical Society of America, 57, 943-948.

Wilson, J.P. (1992). Cochlear mechanics. In Y. Cazals, K. Horner, & L. Demany (Eds.), Auditory Physiology and Perception (pp. 71-84). Oxford: Pergamon Press.

Yates, G.K. (1986). Frequency selectivity in the auditory periphery. In B.C.J. Moore (Ed.), Frequency Selectivity in Hearing (pp.1-50). London: Academic Press.

Yost, W.A. (1994). Fundamentals of Hearing: An Introduction (3rd ed.). San Diego: Academic Press.

Zwicker, E. (1970). Masking and psychological excitation as consequences of the ear's frequency analysis. In R. Plomp & G.F. Smoorenburg (Eds.) Frequency Analysis and Periodicity Detection in Hearing. Sijthoff: Leiden.