SOUND LOCALIZATION AND AUDITORY PERCEPTION BY AN ECHOLOCATING BOTTLENOSE DOLPHIN (*Tursiops truncatus*)

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAIʻI IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

PSYCHOLOGY

MAY 2005

By

Brian K. Branstetter

Dissertation Committee:

Paul E. Nachtigall, Chairperson
Whitlow W. L. Au
Edward P. Chronicle
Kentaro Hayashi
Joseph R. Mobley

ACKNOWLEGMENTS

This work would not have been possible without the hard work from the many interns, volunteers, graduate students and doctors at the Kewalo Basin Marine Mammal Laboratory. I would like to extend special thanks to the "angular resolution interns" who were the backbone of this project from the first day we began to habituate Hiapo to his "eye-cups" to the last day of data collection. In all reality, their names should accompany my own on the title page. Thanks to Matthias Hoffmann-Kuhnt for the flashy 3-D renderings of the experimental set up. Thanks to Dr. Lou Herman for providing me with the opportunity to conduct the empirical studies and for trusting me to run this project with minimal supervision. Thanks to Dr. Adam Pack for his support, interest, and exceptional advice on research training, and experimental methodology. Thanks to Dr. Eduardo Mercado III for support with the computer simulation part of this dissertation. I would also like to thank my five committee members, Dr. Kentaro Hayashi for quantitative support. Dr. Edward Chronicle has been instrumental in streamlining the graduate school process and providing an outside perspective. Dr. Joseph Mobley has served on both my thesis and dissertation committees and has provided me with the most exciting jobs I've ever had. Dr. Whitlow Au provided me with the inspiration to enter this field and has always treated me as if I were one of his students. Special thanks Dr. Paul Nachtigall who has exceeded all expectations as my academic advisor. Without his support, I would have certainly dropped out of graduate school. Finally I would like to thank my family for their patience and support over the long years. This dissertation is dedicated to the memory of the dolphins Hiapo, Elele, Akeakamai and Phoenix.

TABLE OF CONTENTS

LISTS OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

Chapter 1: Dolphin (*Tursiops truncatus*) Echoic Angular Discrimination: Effects of Object Separation and Complexity

A bottlenose dolphin was tested on its ability to echoically discriminate horizontal angular differences between arrays of vertically oriented air-filled PVC rods. The blindfolded dolphin was required to station in a submerged hoop, 2 radial meters from the stimuli and indicate if an array with 2 rods (S+) was to the right or the left of a single rod (S-). The angular separation between the two rods (θ_w) was held constant within each experiment while the angle between the S+ and the S- stimuli (θ_b) varied to produce angular differences ($\Delta\theta = \theta_b - \theta_w$) ranging from 0.25 to 4 degrees. In experiment I, θ_w was maintained at 2 degrees and in experiment II, θ_w was maintained at 4 degrees. Resulting 75 percent correct thresholds (method of constant stimuli) were 1.5 and 0.7 degrees respectively. The results from experiment I are in good agreement with previous results from Branstetter et al., (2003) in which both experiments had $\theta_w = 2$ degrees. The fact that the latter experiment had twice as many targets does not appear to affect the discrimination threshold. A spatial filter model is proposed to account for the varying results. Taken as a whole, the experiments suggest that dolphins have a well developed ability to resolve spatial information through sonar.

INTRODUCTION

A growing body of evidence suggests that dolphins can perceive the spatial structure or shape of complex objects via echolocation (Azzali, Manzini, & Buracchi, 1995; Harley, Putman, & Roitblat, 2003; Harley, Roitblat, & Nachtigall, 1996; Pack & Herman, 1995; Pack, Herman, & Hoffmann-Kuhnt, 2004; Pack, Herman, Hoffmann-

Kuhnt, & Branstetter, 2002). The mental representation of an object appears to be of sufficient spatial detail to allow the dolphin to spontaneously match an object initially interrogated through vision, to one of several alternatives later interrogated echoically (or visa versa; e.g., Pack & Herman, 1995). Herman et al., (1998) suggested that to perceive the spatial structure of a complex object through echolocation, the dolphin would be required to resolve a temporal-spatial array of echoes from the various object features, and integrate this information to engender a holistic three-dimensional representation. The dolphin's ability to resolve fine spatial details from an object will be dependant at least upon the physical characteristics of the object, the characteristics of the dolphin's sonar signal, the dolphin's auditory signal processing of the returned echoes, and the dolphin's ability to localize echoes from the various reflecting points. The focus of the present study is to determine the limit of the dolphin's echolocation acuity in the horizontal plane.

Background

The incident sonar signal from a bottlenose dolphin[1] can be described as a broadband transient with peak frequencies between 40-140 kHz, durations as short as 40 μsec, and peak-to-peak sound pressure levels often exceeding 220 dB re: 1 μPa (Au, 1993). The concave shape of the dolphin's skull, coupled with the focusing properties of the lipid melon, focus the signal forward in a tight beam, elevated about 5 degrees above the rostrum, with a 3 dB beam width of approximately 10 degrees (Au, 1980). Consequently, targets directly forward of the melon will be highly ensonified, increasing the signal-to-noise ratio relative to peripheral targets. The high frequency, short duration

---

[1] Henceforth, the word "dolphin" will be used to refer to the bottlenose dolphin (*Tursiops truncatus*).

properties of individual clicks contribute to the dolphin's ability to resolve fine spatial details. Higher frequencies allow for echo returns from small targets, provided the target's spatial extent is greater than the wavelength of the incident signal. High frequency components also allow dolphins to detect small structural features within a target that may facilitate target identification. Short duration clicks can improve range resolution by limiting overlapping echoes from closely spaced objects in the same line of acoustic propagation. In addition, fine temporal resolution, demonstrated by an auditory critical interval of 264μs (Moore et al., 1984), allows the dolphin to resolve echoes occurring in close temporal proximity.

Despite the environmental challenges imposed by an aquatic environment (e.g., sound traveling approximately 4.5 times faster in water than in air), several of the standard terrestrial mammalian localization mechanisms also appear to be well developed in dolphins. When a sound source originates off the midsagital plane, a differential distance from each ear to the sound source is produced. The differential distance results in arrival time differences between the two ears know as interaural temporal differences (ITDs). A differential distance between the ears and a sound source will also produce an intensity difference between the ears know as interaural intensity differences (IIDs). Moore et al., (1995) investigated the dolphin's ability to utilize ITDs and IIDs. Jaw phones (i.e., hydrophones embedded in rubber suction cups and attached to the right and left lower jaws) were used to provide precise control over binaural stimulus presentation to the dolphin. ITD thresholds were as small as 7 μsec for a 30 kHz pulse. For sonar type pulses (with higher peak frequencies above 60 kHz) ITDs were between 17 and 18 μsec.

IID thresholds were below 1.0 dB for all frequencies tested. Thus, both ITDs and IIDs appear to be very salient cues a dolphin could employ to localize sound sources.

As noted, in order to echoically perceive the spatial structure of complexly shaped objects a dolphin must be able to echoically resolve angular differences between closely spaced object features in different planes. Recently, Branstetter et al. (2003) investigated a dolphin's ability to echoically resolve angles between multiple targets in the horizontal plane. A blindfolded dolphin positioned its head within a stationary, vertically oriented hoop, two radial meters from the stimuli, and then indicated whether an array with four rods (S+) was to the left or the right of an array with two rods (S-), by pressing a corresponding paddle. The angular separation between the rods within each array (θ_w) was maintained at two degrees but the angular separation between the two arrays (θ_b) was varied to produce angular differences ($\Delta\theta = \theta_b - \theta_w$) ranging between 0.25 degrees to 4 degrees. A modified method of constant stimuli used to test for angular discrimination ability, yielded a psychometric function with a 75% correct threshold of 1.6 degrees (see Figure 11). The results were consistent with passive hearing studies that determined the smallest possible angle (minimum audible angle or MAA) between two sound sources that would allow the two sources to be perceived by the dolphin as discrete. Renaud and Popper (1975) reported the horizontal MAA for 30, 60, and 90 kHz pure tones as 2.5, 3.0, and 3.0 degrees respectively. Horizontal and vertical MAAs for click stimuli (35 μs duration, 64.35 kHz peak frequency) were 0.9 and 0.7 degrees respectively.

The results from Branstetter et al. (2003), suggested that the dolphin's echoic spatial acuity was sufficient to support cross-modal matching of objects between vision and echolocation as shown by Pack and Herman (1995) and Herman et al. (1998).

Because the dolphin in these studies interrogated relatively large objects (approximately 43 cm^2) at relatively short distances (< 1 m), the angle the objects subtended was much greater than the angular differences threshold of 1.6 degrees.

The current study builds on the findings of Branstetter et al., (2003). One goal of Branstetter et al. was to generalize the findings to the cross modal paradigm (i.e., if the dolphin can echoically resolve fine spatial details from multiple echoes in the angular discrimination experiment, the dolphin can probably echoically resolve fine details in the cross-modal experiments). Because the objects in the cross-modal experiments (Pack et al., 2003) were three dimensional with multiple features that produced multiple echoes, the stimuli in Branstetter et al., (2003) were composed of multiple targets (4 rods vs. 2 rods) thus producing multiple echoes as well. However, the resulting complex stimuli may not have been optimal to produce the smallest echoic angular discrimination threshold. In humans, reducing the number of distracting sound sources can aid in localization (Wightman et al., 2001). We employed this tactic in Experiment 1 of the current study by reducing the S+ and S- stimuli to two rods and one rod respectively.

In the Branstetter et al. (2003), experiment, the angular difference between the rods within each array (θ_w) was always 2 degrees. The proportion correct at an angular difference of 2 degrees was 83 percent compared to 93 percent at 4 degrees. This suggests when $\theta_w = 2$ degrees, the perceived positions of the rods were subject to spatial blurring. The echoes from each rod could have effectively masked each other. In humans, an improvement in the ability to localize sound sources occurs due to increased spatial separation (Langendijk et al., 2001). We adopted this strategy in Experiment II by increasing θ_w to 4 degrees. If all the rods were more fully resolvable, perhaps the dolphin could make better decisions about the rod positions relative to one another.

GENERAL METHODS

Subject

The subject for the study was a 16 year-old, male, Atlantic bottlenose dolphin (*Tursiops truncatus*) named Hiapo. Hiapo was the same dolphin used in the Branstetter et al. (2003) study. Hiapo was housed in two interconnected seawater tanks (each circular tank was 1.8 m deep with a diameter of 15.2 m) with three female dolphins at the Kewalo Basin Marine Mammal Laboratory in Honolulu, Hawaii. During the experimental procedures, Hiapo was maintained alone or with one other dolphin. Approximately 11.8 kg of herring, capelin, smelt and squid were fed to Hiapo daily, a portion of which (2.9 kg) was fed during the experiment.

Apparatus

The same experimental apparatus (see figure1) employed from Branstetter et al. (2003) was employed in all of the experiments presented. The dolphin was trained to wear latex rubber suction cups (eye cups), over its eyes to prevent the use of vision during the experiment. A submerged, vertically oriented, hollow, PVC stationing hoop was attached to the tank wall. The hoop's inner surface was wrapped in a thin 2mm layer of insulation foam for the dolphin's comfort when it stationed in the hoop. Perforations allowed the hoop to fill with water that helped the hoop approach neutral buoyancy. The inner diameter of the hoop, including the foam layer, was 40.5 cm in diameter and specifically tailored for Hiapo's dimensions. The hoop was submerged approximately 55 cm below the surface of the water and was 71 cm from the tank wall (measurements taken from the center of the hoop). Two submerged response paddles were connected to the periphery of the stationing hoop allowing the dolphin to exit the hoop and respond to either the left or right paddle. The paddles were constructed of acoustically reflective fishing net floats and were an equal distance (approximately 65 cm) from the center of the hoop. The stimuli were presented using a customized aluminum rack, which was positioned 2 m from the tip of the dolphin's rostrum. The distance was accurately maintained by two PVC pipes connecting the top of the stationing hoop to the corners of the aluminum rack.

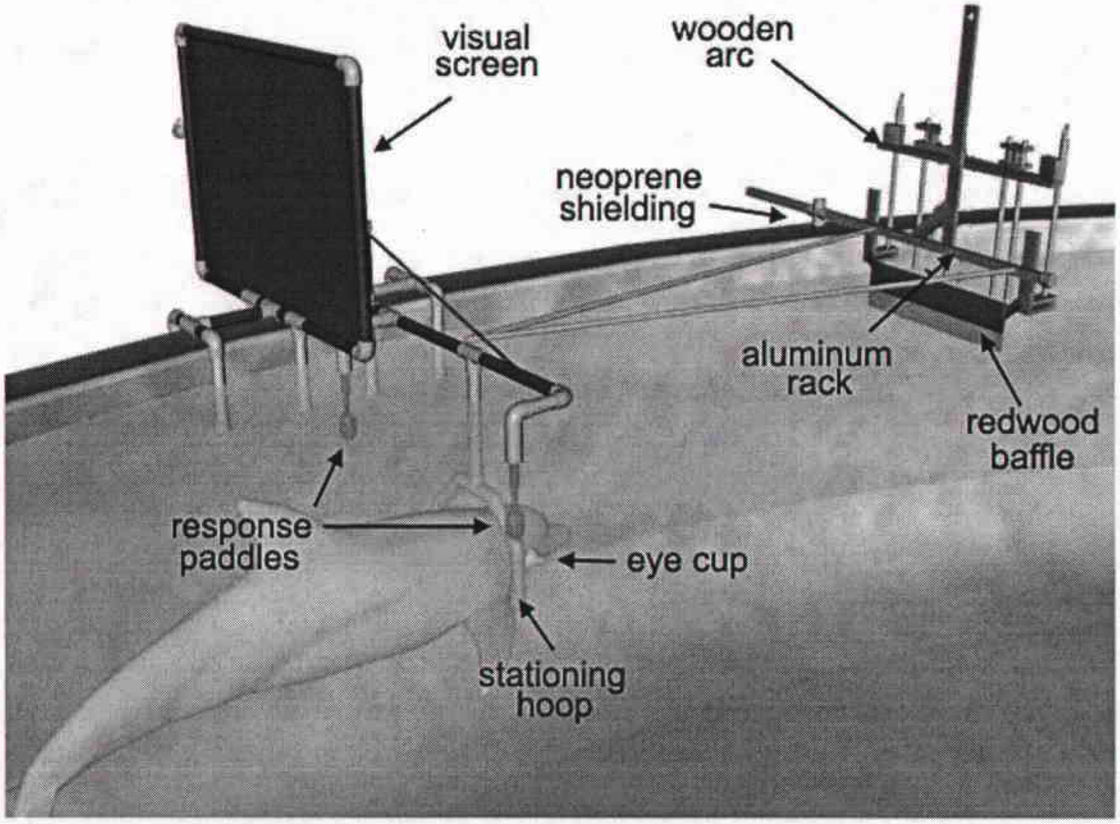

Figure 1. Apparatus and architecture of experiments

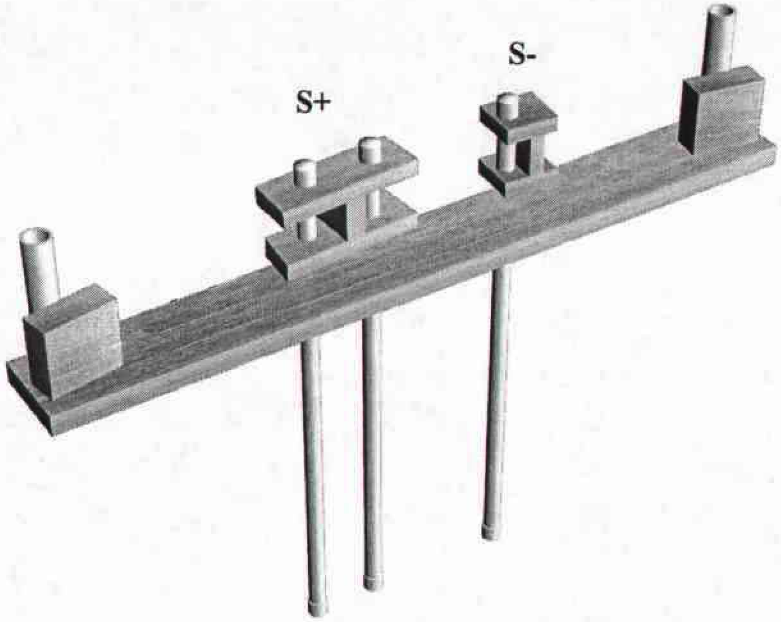

Figure 2. Experimental stimuli

The top of the aluminum rack was fitted with two slots to allow positioning of the stimuli (figure 1). A redwood baffle at the bottom of the rack occluded inspection (visually or echoically) of the tips of the stimuli when they were positioned on the rack prior to full immersion in the water. Finally, neoprene shielding covered any part of the aluminum rack that came into contact with the tank wall. The shielding reduced potentially distracting sound conduction from the rack, into the tank wall, and into the water.

The experimental stimuli consisted of vertically oriented, schedule 40, PVC rods (figure 2). Each rod was air filled to provide a high impedance mismatch to water making them highly reflective to dolphin sonar signals. In all experiments, the positive stimulus to respond to (S+ stimulus) consisted of two rods and the negative stimulus not to respond to (S- stimulus) consisted of one rod. Each PVC rod was 1.9 cm in outer diameter, 1.7 cm in inner diameter and 75 cm in length. PVC caps were placed on each end of each rod (2.7 cm outer diameter, 2.3 cm length) to ensure they were air-tight. A total of 12 PVC rods with caps were used to construct three S+ stimuli and three S- stimuli. Each rod and cap was constructed from the same type of PVC with the same batch number ensuring the material composition was nearly identical. However, as a precaution, before each session, each rod and each cap were randomly assigned to an S+ or an S- array to control for the possibility that Hiapo could learn to identify a particular rod artifact associated with a particular array.

The rods in S+ stimulus were separated from each other by θ_w degrees from a 220 cm distance. This distance represented the radial distance from the arrays to the midpoint between Hiapo's mandibular windows on his lower jaw. An S+ and an S- stimulus were positioned on wooden arcs with the angular separation (θ_b) between the stimuli predetermined from a counterbalanced schedule. The angular difference can be defined as:

$$\Delta_\theta = \theta_b - \theta_w$$

where Δ_θ is the angular difference, θ_b is the separation between the S+ and S- arrays and θ_w is the angle between the rods within each array (figure 3). The arrays were centered on the wooden arc so the far end of the S+ array and the far end of the S- array were an equal distance to the center of the arc and hence, the dolphin's median plane (See Appendix I).

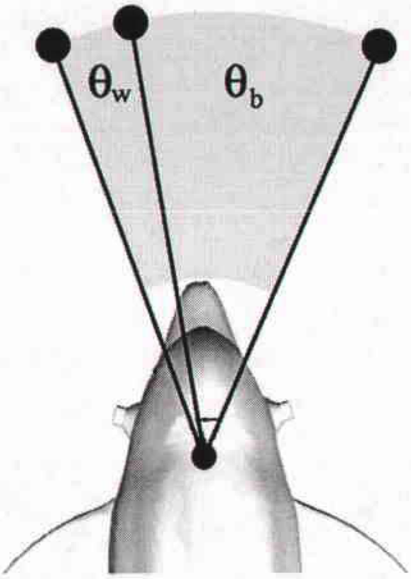

Figure 3. Angular differences

When Hiapo was positioned inside the stationing hoop as far as his pectoral fins would allow, each rod was approximately 200 radial cm from the tip of his rostrum, 220 radial cm from his mandibular window and 245.7 radial cm from the center of the stationing hoop. This allowed Hiapo to pivot his head in the horizontal plane along a 45.7 cm radial arc (measured from the center of the stationing hoop to the tip of his rostrum) while the stimuli maintained a distance of 200 cm from his rostrum tip. The wooden arc was attached to a pulley system on the aluminum rack that was attached to both the tank wall and a wooden stand on the exterior of the tank (Figure 1). The rack and the pulley system provided control over stimuli exposure time by allowing an assistant to lower the arrays when asked. When lowered, the only part of the array submerged was the bottom 55 cm of the six rods. The rack was leveled using a standard bubble level. A visual screen (figure 1) prevented the trainer from seeing the stimuli and inadvertently cueing the dolphin towards the correct response.

Sonar emissions were recorded with a hydrophone (LAB-core systems, PZ-1A) placed approximately 15 cm in front of the stationing hoop at a depth of approximately 15 cm. When Hiapo was properly stationed in the hoop, the hydrophone was directly above his head and slightly behind his blow hole. The sounds were digitized at 44.1 kHz sampling rate and recorded with a Sony digital video recorder. Because the sonar emissions were already distorted by the low sampling rate, near field distortion was not considered an issue. Hiapo would emit sonar signals prior to array exposure and after his response, only the clicks between the onset of array exposure to when he backed out of the hoop (monitored from recorded video) were considered for analysis. Backing out of the hoop was operationally defined as when his eye cups crossed the hoop threshold on his way out of the hoop. Click trains for each trial were then manually counted by two independent observers from a hardcopy printout. Both observers were blind to which trial condition they were reviewing.

Response time (RT) for each trial was also recorded by an observer with a digital stopwatch. RT was defined as the time between the onset of stimulus exposure to when the dolphin pressed a paddle. The trainer (who, because of the visual screen had no knowledge of the stimulus array) vocally reported a paddle press.

General Procedure

At the beginning of each trial, the eye cups were placed on the dolphin. Hiapo remained at the trainer's station with his jaw held out of the water as an assistant then placed the stimuli on the rack and the tips of the PVC rods were gently submerged a few cm (< 5 cm) below the water surface to prevent a potential passive acoustic splash cue. This reduced the possibility of Hiapo passively listening to the stimuli tips while they were placed in the water. In addition, a water-soaked redwood baffle, attached to the aluminum rack was positioned about 5 cm beneath the surface in front of the tips of the rods (Figure 1). Water soaked redwood is known for its ability to absorb sound underwater (see Johnson, 1967). The baffle served to reduce any sound the stimuli may have made upon entering the water and prevented the dolphin from echolocating on the tips of the stimuli before they were fully lowered. Once the stimuli were in place, an experimenter in an elevated tower overlooking the tank instructed the trainer to signal the dolphin to enter the hoop. The trainer manually assisted the dolphin to insure that he was positioned in the hoop up to his pectoral fins. The experimenter then verbally instructed the assistant to lower the arrays into the water. The dolphin was required to echolocate and identify the location (left or right) of the S+ array by backing out of the hoop and touching the corresponding left or right response paddle. The trainer verbally called out the dolphin's response, either "left" or "right," and an experimenter located on an elevated deck overlooking the pool, identified the choice as correct or incorrect, referring to the pre-planned schedule. The trainer also functioned as a blind observer because the visual barrier prevented the trainer from seeing the stimuli and thus, knowing the correct choice. Correct responses were rewarded with fish, and social reinforcement, followed by an inter-trial interval of approximately 35-40 seconds. The dolphin was called back to station for incorrect responses, did not receive a fish reward and proceeded directly into an inter-trial interval.

Three experiments were conducted. Their presentation was counterbalanced (ABCCBA format) to control for any learning effects. Half of experiment I, experiment II and experiment III were completed in consecutive order. Then, the remaining halves were conducted in reverse order (i.e., Exp III, Exp II, then Exp I). Prior to conducting these experiments, a Pilot study was run to investigate how easily Hiapo would generalize from his traditional arrays of four rods versus two rods (Branstetter et al., 2003) to two rods versus one rod.

PILOT STUDY

In the Branstetter et al. (2003) experiment, the dolphin was required to choose the array with four rods while rejecting the array with two rods. To perform this task the dolphin could have developed one of at least two rules. An *identity rule* would be specific to the context, "choose the stimulus with four rods while not choosing the stimulus with two rods." A *concept rule* would be more generalizable to novel configurations, "choose the stimulus with the most rods while not choosing the stimulus with the least rods." An identity rule is specific only for the stimuli in Branstetter et al.(2003) while a concept rule can be generalized to new stimuli. In the current experiment, the dolphin was required to choose the array with two rods while not choosing the stimulus with a single rod. Because the current stimulus configuration was novel for the dolphin, a short pilot study was conducted to determine if Hiapo would employ an identity rule or a concept rule. Note that in Branstetter et al. (2003) the array with two rods was always the S- and that now it would be the S+. In addition, the results of the pilot study helped designate which angular separations would be presented for experiment I to economically capture the full range of the psychometric function.

Stimuli

The S+ and S- stimuli were composed of a 2-rod array and a single rod respectively (see Figure 2). The angle between the rods within the S+ array (θ_w) was held at a constant 2 degrees and the angle between the S+ and S- array (θ_b) varied to produce angular differences (Δ_θ) of 4.0, 3.0, 2.0, 1.0, and 0.5, degrees.

Procedure

Each session began with 6 warm up trials with an angular difference of 4 degrees. The position of the S+ array (left or right) was randomized within each session. An approximate 25 sec inter-trial interval followed each trial to allow assistants to prepare the next set of stimuli. If the dolphin responded incorrectly on two or more warm-up trials, the testing session was aborted until the next day. If the dolphin responded correctly on five or more warm-up trials ($p < .05$, cumulative binomial), the test session began after a 1.5-2 minute break. The modified method of constant stimuli was used for stimulus presentation. Each angular difference was presented three times in descending order. The last four trials were cool-down trials with an angular difference of 4 degrees. Thus, a total of 6 warm-up trials, 15 test trials, and 4 cool-down trials were conducted each session. If the dolphin was incorrect on 3 or more of the warm-up and cool-down trials combined, the data from the session were eliminated from analysis. Data collection for the pilot study began directly after data was collected for Branstetter et al., (2003).

Results and Discussion

A total of three sessions were conducted. A psychometric function describing the dolphin's performance is plotted in Figure 4. Each data point represents the percentage of correct responses for 9 trials. Using a criterion of 75% correct responses, a threshold of

1.7 degrees was obtained. Hiapo was correct on 27 of 30 "baseline" trials (18 warm-up, 12 cool down). The results strongly suggest Hiapo was employing a concept that was easily generalizable to novel configurations despite the reversal of reward contingencies.

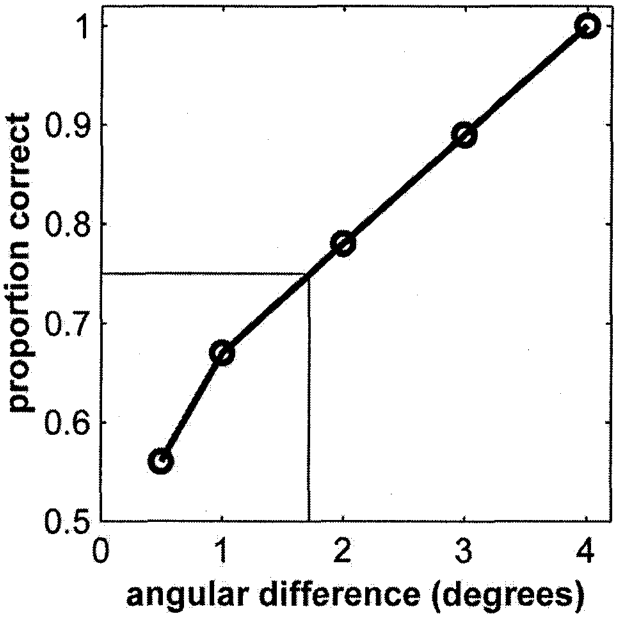

Figure 4. Results from pilot study. A 75% correct threshold of 1.7 was obtained.

EXPERIMENT I

The goal of experiment I was to determine if echoically localizing targets is less difficult when the number of echoes received per incident signal is reduced. In humans, sound localizing performance progressively increases as the number of distracting sounds decreases (Wightman and Kistler, 1997; Langendijk at al., 2001). In the current experiment, the number of potential echoes is reduced to three by decreasing the number of rods presented on each trial to three. If echo complexity, defined as the number of simultaneous echoes received, is a major factor influencing echoic angular discrimination, a threshold decrease is expected in the current experiment relative to the results of Branstetter et al., (2003).

Stimuli

The S+ stimulus was a 2-rod array. The S- stimulus was a single rod (see Figure 2). The angle between the rods within the S+ array (θ_w) was held at a constant 2 degrees and the angle between the S+ and S- array (θ_b) varied to produce angular differences (Δ_θ) of 4.00, 3.00, 2.50, 2.25, 2.00, 1.75, 1.50, 1.25, 1.00, and 0.50 degrees. Ten angular differences were tested compared to only eight in Branstetter et al., (2003) to provide greater resolution estimating the psychometric function.

Procedure

Each session began with 6 warm up trials with an angular difference of 4 degrees. The position of the S+ array (left or right) was randomized with an equal number of left and right presentations within each session. An approximate 25 sec inter-trial interval followed each trial to allow assistants to prepare the next set of stimuli. If the dolphin responded incorrectly on two or more warm-up trials, the testing session was aborted until the next day. If the dolphin responded correctly on five or more warm-up trials ($p < .05$, cumulative binomial), the test session began after a 1.5-2 minute break. The modified method of constant stimuli was used for stimulus presentation. Angular differences were presented in descending order. Each of the 10 angular differences was tested twice per session in two descending sweeps. The left or right position of the S+ stimulus was randomized for the first ten trials. The position of the S+ array in the second ten trials was the opposite of the first ten trials. For example, if the location of the S+ array for an angular separation of 2 degrees occurred on the left within the first ten trials, the position of the S+ array for the same angular separation in the second ten trials would be to the right. This ensured that any potential right or left bias would affect angular differences equally. The last four trials were cool-down trials with an angular difference of 4 degrees. Thus, a total of 6 warm-up trials, 20 test trials, and 4 cool-down trials were conducted each session. If the dolphin was incorrect on 3 or more of the warm-up and cool-down trials combined, the data from the session were eliminated from analysis.

Results and Discussion

A total of 35 sessions were conducted over the duration of experiment I. Two of the sessions were aborted because Hiapo failed to respond correctly to 5 or more of the warm-up trials. One session was aborted to repair the stationing hoop that Hiapo dismantled in mid-session. Two additional sessions were aborted because Hiapo would not respond to his trainer. Therefore, data from a total of 30 sessions were used for analysis. A psychometric function describing the dolphin's performance is plotted in figure 5. Each data point represents the percentage of correct responses for 60 trials. As the angular difference between the stimuli decreased, so did the dolphin’s ability to discriminate the angular difference. Using a criterion of 75% correct responses, a threshold of 1.5 degrees was obtained.

There was no significant linear trend effect of angular difference on the number of clicks the dolphin emitted ($F = 0.97$, $p > 0.05$, $y = -2.75x + 195.45$). On average, the dolphin emitted 190.01 clicks per trial ($SD = 39.28$). The maximum number of emitted clicks on a single trial was 287 while the smallest number was 97. There was a significant linear trend effect of angular difference on the dolphin’s reaction time ($F= 14.42$, $p >$

0.05, y = -0.14x + 4.79). On average, the dolphin's response latency was 4.50 sec (SD = 0.81 sec).

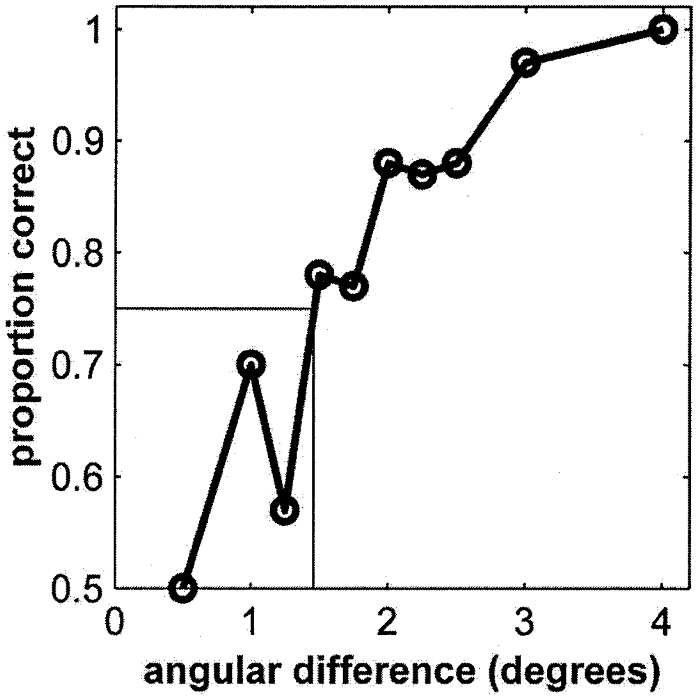

Figure 5. Results from Experiment I. A 75% correct threshold of 1.5 was obtained.

Maximum and minimum response latencies were 9.25 and 2.91 sec respectively. Although the dolphin appears to require significantly more time to respond as the angular separation decreases, the trend appears to be rather weak with a minor negative slope.

The current threshold compares very well with the threshold of 1.6 measured by Branstetter et al., (2003) These data suggest that the dolphin was able to simultaneously localize six sound sources (i.e., Branstetter et al, 2003) approximately as well as it can simultaneously localize three sound sources (i.e., current study). Thus, within the limited stimuli used between Branstetter et al. and the current study, echo complexity does not appear to be a major factor governing echoic angular discrimination.

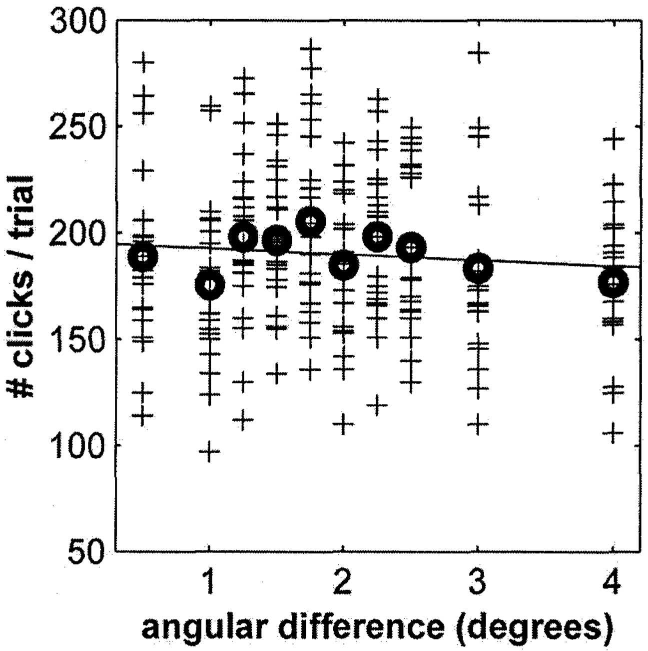

Figure 6. Number of clicks (from experiment I) per trial as a function of angular difference.

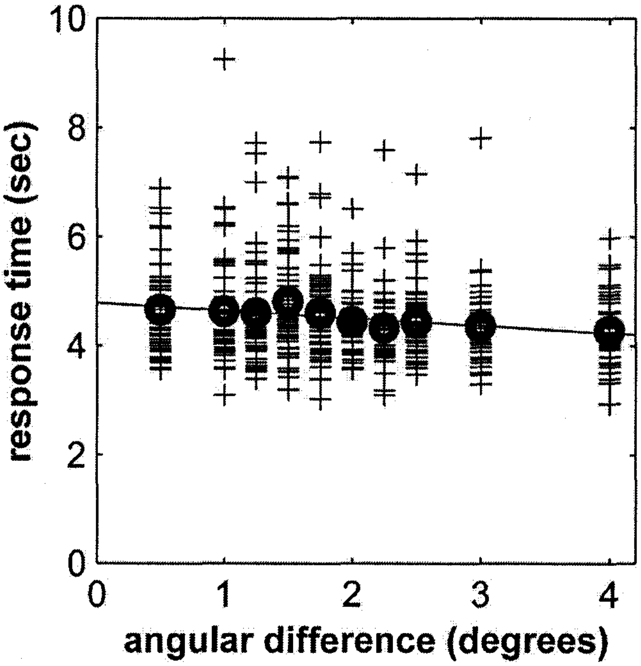

Figure 7. Response time (from experiment I) as a function of angular difference

EXPERIMENT II

The goal of experiment II was to determine if the positions of the rods are less difficult to localize when each rod is more fully resolvable. The threshold from experiment I and Branstetter et al., (2003) were 1.6 and 1.5 respectively. Because θ_w was equal to 2 degrees is both of these experiments, the rods may have been too close together to be fully resolvable. By increasing θ_w to 4 degrees each rod should be more fully resolvable, thus allowing the dolphin to make better judgments about the precise location of each rod relative to one another. If this assumption is true, the dolphin's angular discrimination threshold should be lower than those obtained by experiment I and Branstetter et al. (2003)

Stimuli

For Experiment II, the angle between the rods within the S+ array (θ_w) was 4 degrees and the angle between the S+ and S- array (θ_b) varied to produce angular differences (Δ_θ) of 4.00, 3.00, 2.50, 2.00, 1.50, 1.00, 0.50 and 0.25 degrees.

Procedure

Experiment II employed identical procedure to Experiment I, however there were eight angular differences tested. A subjective decision was made to decrease the number of presented angular differences from 10 (in experiment I) to eight to shorten the experimental sessions. The longer sessions appeared to result in a decrease of motivation by the dolphin. Thus, a total of 6 warm-up trials, 16 test trials, and 4 cool-down trials were conducted each session.

Results and Discussion

A total of thirty eight sessions were conducted. A three month hiatus resulted between the dates of January 8th to March 26th (2001) due to the death of a companion dolphin. The Hiatus took place between session 18 and 19. During this time, a total of 9 practice sessions (half the number of trials / session) were conducted to maintain the behaviors required for the experiment. An *apriori* decision was made to exclude all practice data from analysis. Experiment II resumed after the remaining companion dolphins were trained to facilitate (not interfere with) data collection. Six sessions were aborted because Hiapo failed to respond correctly to 5 or more of the warm-up trials. An additional two sessions were aborted because Hiapo chose to socialize with the other dolphins rather than respond to his trainer. Therefore, data from a total of 30 sessions were used for analysis. A psychometric function describing the dolphin's performance is plotted in figure 8. Each data point represents the percentage of correct responses for 60 trials. Using a criterion of 75% correct responses, a threshold of 0.7 degrees was obtained.

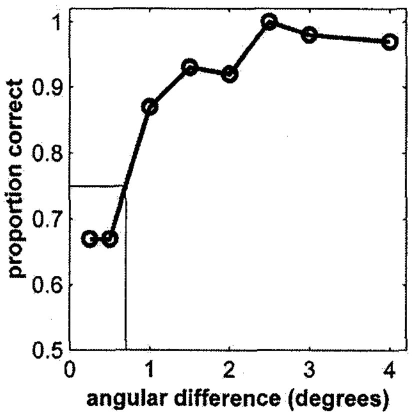

Figure 8. Results from Experiment II. A 75% correct threshold of 0.7 degrees was obtained.

There was no significant linear trend effect of angular difference on the number of clicks the dolphin emitted (F= 0.40, p > 0.05, y = -1.16x + 168.19), nor was there a significant linear trend effect of angular difference on the dolphin's reaction time (F= 3.98, p >0.05, y = -0.05x + 3.91). On average, the dolphin emitted 166.06 clicks per trial (SD = 38.57). The maximum number of emitted clicks on a single trial was 260 while the smallest number was 90. On average, the dolphin's response latency was 3.83. sec (SD = 0.67 sec). Maximum and minimum response latencies were 8.70 and 2.17 sec respectively.

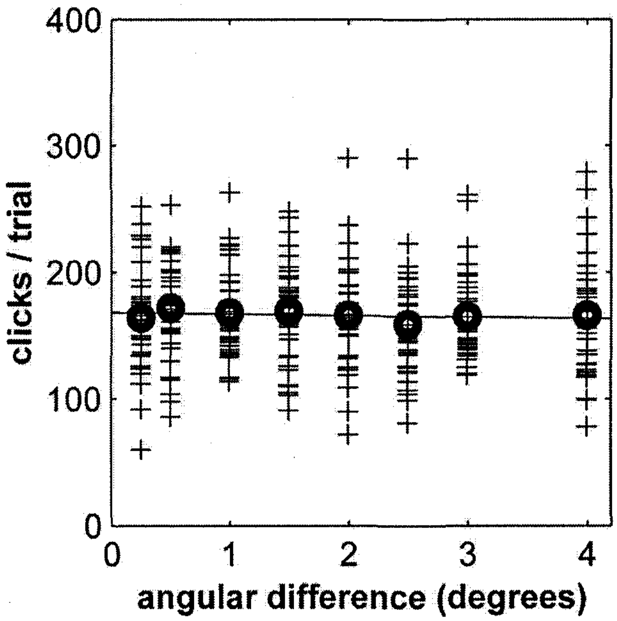

Figure 9. Number of clicks (from experiment II) per trial as a function of angular difference.

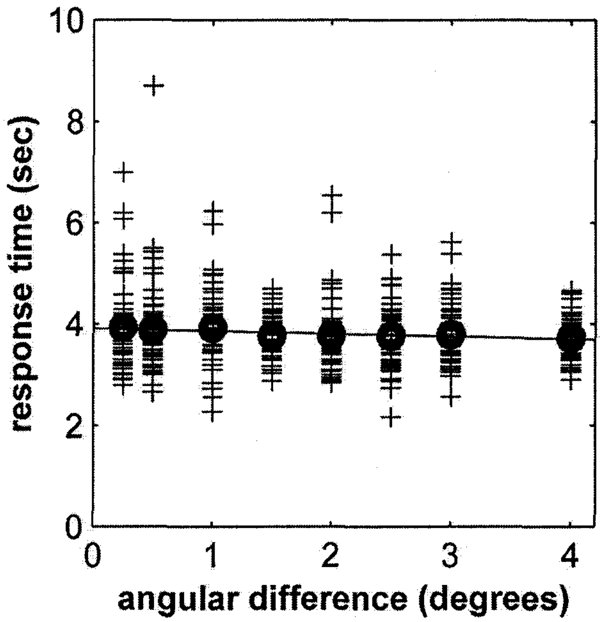

Figure 10. Response time (from experiment II) as a function of angular difference

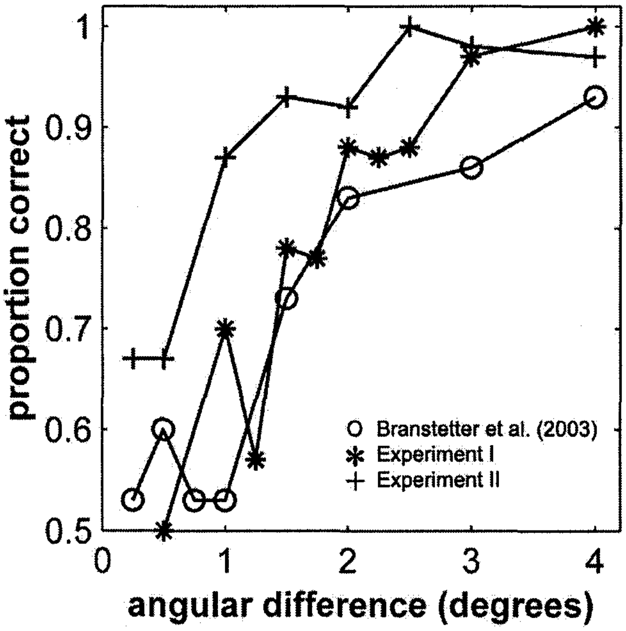

Figure 11. Psychometric functions from Branstetter et al., (2003) and experiments I and II.

Table 1

Comparison Between Psychometric Functions from Exp. I and Exp. II.

$\Delta\theta$	0.50	1.00	1.25*	1.50	1.75*	2.00	2.25*	2.50	3.00	4.00
z	-1.85	-2.22	-4.13	-2.36	-2.40	-0.61	-1.78	-2.73	-0.58	--
p	**0.03**	**0.01**	**0.00**	**0.00**	**0.00**	0.25	**0.04**	**0.00**	0.28	--

A point-by-point comparison between proportions correct from experiment I and II were compared for each angular separation. The angular separations that experiment II did not have in common with experiment I were interpolated by averaging adjacent points and are denoted with an asterisk in Table 1. Differences between points on the psychometric function were large enough to produce significant one-tailed z-scores (indicated by bold p values) for all angular separation except for $\Delta\theta = 2.00$, 3.00 and 4.00. The proportions associated with the last two angular separations are similar due to a ceiling effect. The angular separations surrounding the 75% correct level are all significantly different, thus we conclude the threshold differences are real. A statistical comparison to Branstetter et al. (2003) was not conducted because the results in this study are not counterbalanced with the former study. However, a large leap of faith is not required to notice the function similarity between Exp I. and Branstetter et al.

The number of clicks from experiment I and experiment II were also compared. To prevent any potential biases, only the angular separations the two experiments had in common were chosen for comparison ($\Delta\theta = 0.5$, 1, 1.5, 2, 2.5, 3, and 4). Because there was no significant linear trend in either experiment, a grand mean was calculated for each experiment by pooling the data from the angular separation above. The resulting mean clicks per trial for experiments I and II were 186.52 and 166.35 respectively. There were no significant differences between these means, $t(291)=1.47$, $p > 0.05$ (two-tailed).

Response time was also compared between experiments I and II with only the angular separation in common used for analysis. Despite the fact that there was a significant linear trend for reaction time in Exp. I, a grand mean was calculated for each experiment by pooling the data from the angular separation. There was a significant difference, $t(739) = 12.54$, $p < 0.01$(two-tailed) between the means of 4.5 sec and 3.8 sec for experiments I and II respectively. Because simply pooling data results in the most parsimonious comparison with the least amount of variance accounted for (compared to a linear model with more than a single parameter), the significant result is considered conservative.

The correlation between response time and number of clicks was examined for random subset of trials within each experiment. A moderate low correlation resulted in Exp. I ($R^2 = 0.24$, $n = 99$) and a weak correlation resulted from Exp II ($R^2 = 0.07$, $n=92$).

EXPERIMENT III

The goal of experiment III was to determine if the dolphin may be using an acoustic time-delay confound caused by specular reflection off the tank wall known as the Lloyd-mirror effect (see figure 12). The possibility exists that echoes from the rods may then reflect off the tank wall and arrive at the dolphin with slightly different delays. The reverse is also possible, that the dolphin's incident signal could reflect off the tank wall, and then reflect off the rods, arriving at the dolphin with slight time delays. The dolphin then could learn to associate specific delay patterns with correct left or right responses. To prevent specular reflections from the wall, a redwood baffle was placed along the tank wall between the dolphin and the stimuli. The baffle was constructed from a matrix of redwood wedges glued (hot glue gun) to a series of redwood planks. Each wedge was constructed from a 3 x 3 x 10 cm rectangular block, with an angle cut 5 cm deep along the grain of the wood. The angles were pointed outward so the grain of the wood was along the line of acoustic propagation. The baffle was soaked in water for 3 weeks prior to data collection to ensure saturation. The purpose of the redwood baffle was to absorb and scatter any sounds that could produce specular reflections from the tank wall. Redwood baffles have been used extensively for sound attenuation in aquatic environments (e.g., Johnson, 1967). If the dolphin relied on the pattern of specular reflections from the tank wall to determine the correct choice, the introduction of the redwood baffle should disrupt this cue. The result would be an increase in the angular separation threshold relative to a condition without the baffle. Thus, thresholds for two conditions were estimated, one with the redwood baffle present, and one without the redwood baffle.

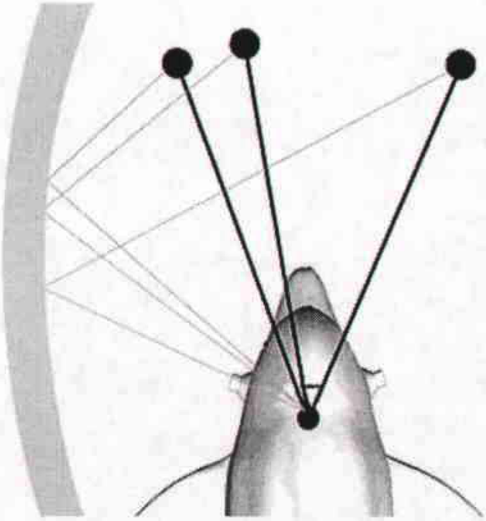

Figure 12. Lloyd mirror effect.

Stimuli

The stimuli for experiment III was identical to experiment II.

Procedure

Because we were only interested in comparing thresholds between two conditions (baffle present, baffle not present), full psychometric functions were not required. We thus employed a 2-down, 1-up, adaptive staircase procedure to estimate thresholds at the 71 percent correct level (Levitt, 1971). Each session began with 6 warm up trials with an angular separation of 4 degrees. If Hiapo made 2 or more errors, the session was aborted until the next day. If Hiapo was correct on 5 or more warm up trials, titration of the angular separation began with initial angular separation of 4 degrees. Step size was determined by table 1. An experimental session ended when a total eight reversals were collected or 30 trials were complete. The first 2 reversals were not included in calculating a threshold. A total of 34 reversals were used for calculating each threshold.

Table 2

Angular Separation (Δ_θ)
4.00
3.00
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.128

Results and Discussion

There was no significant difference between threshold levels for the baffle present and the no baffle conditions (one-tailed $t(66) = 1.37, P > 0.5$)[2]. The threshold level for the baffle condition was 1.04 degrees and the threshold level for the no-baffle condition was 1.23 degrees. The results suggest the dolphin was not using specular reflections from the tank wall as a basis for making decisions.

GENERAL DISCUSSION

The two main findings of this study are: 1) decreasing the number of targets from six to three does not aid in localization, and 2), increasing the space between the rods enhances localization. Although the finding from experiment I (that decreasing the number of targets does not aid in localization), departs from human results (Wightman & Kistler, 1997; Langendijk et al., 2001), this may be a selective adaptation for dolphins. During foraging, dolphins are required to echoically detect, track, and capture multiple fish, often schooling in close spatial proximity. Having the ability to localize multiple targets simultaneously would no doubt be advantageous, if not a basic requirement for this type of auditory predation.

The results were consistent with passive hearing studies that determined the smallest possible angle (minimum audible angle or MAA) between two sound sources that would allow the two sources to be perceived by the dolphin as discrete. Renaud and Popper (1975) reported the horizontal MAA for 30, 60, and 90 kHz pure tones as 2.5, 3.0,

---

[2] Statistically comparing the two means is actually irrelevant because the dolphin performed slightly better with the baffle present. The hypothesis that the dolphin employs specular reflections and removing the reflections would increase the threshold can be rejected without any statistical test.

and 3.0 degrees respectively. Horizontal and vertical MAAs for click stimuli (35 μs duration, 64.35 kHz peak frequency) were 0.9 and 0.7 degrees respectively.

The results from the current study, along with Branstetter et al. (2003), suggest that the dolphin's echoic spatial acuity is sufficient to support cross-modal matching of objects between vision and echolocation as shown by Pack and Herman (1995) and Herman et al. (1998). Because the dolphin in these studies interrogated relatively large objects (approximately 43 cm^2) at relatively short distances (< 1 m), the angle the objects subtended was much greater than the angular differences threshold reported here.

The question remains, how do dolphins segregate individual echoes among several closely spaced echoes? What type of mechanisms may be involved? One possible explanation for these findings may be related to the phenomenon of "summing-localization blur" caused by the presents of simultaneous sound sources. For human listeners, if two *coherent* sound sources are presented simultaneously at an equal distance from the midsagital plane, a single "phantom" sound is perceived half the distance between the sound sources. Coherence can be defined as two signals that are identical and the degree of coherence can be calculated by the normalized-cross correlation function. There are several variables that can disrupt the percept of the phantom sound, one of which is degree of sound coherence As the degree of coherence decreases, summing-localization blur decreases (Jeffress et al., 1962). Although the PVC rods in the experiment were identical, during ensonification, they will not produce coherent echoes. The incident signal of the bottlenose dolphin is not only directional with respect to amplitude, but also highly directional and asymmetric in the frequency domain. Au (1980) demonstrated that peak frequency of the incident signal was 122 kHz directly in front of the dolphin but was 38 kHz only 10 degrees to the right and 115 kHz, 10 degrees to the left of the dolphin. Thus, targets separated in space will produce echoes with different peak frequencies and thus, varying degrees of coherence. Although Au (1980) did not measure peak frequencies for small angles off the median, figure 13 illustrates what the peak frequencies might be when a 4th order polynomial is fit to the data from Au (1980). For the rods at 2 and 4 degrees, the peak frequencies of the echoes would be 109 and 94 kHz respectively. Thus, echoes off the median plane would have spectral differences of 13 kHz and 28 kHz for 2 and 4 degrees respectively. Although a polynomial is overly simplistic, spectral differences are nevertheless likely to be larger for rods separated by 4 degrees compared to 2 degrees. These spectral differences decrease signal coherence and may aid the dolphin in degrading summing-localization blur for closely spaced objects.

In addition, lower frequencies will not reflect well off of a small object compared to higher frequencies due to their relatively large wavelengths. For example, two incident signals with a 38 kHz and a 122 kHz peak frequencies will have 4.0 cm and 1.2 cm wavelengths respectively. Because the PVC rods have a 1.9 cm outer diameter, the 38 kHz signal will produce a much more attenuated echo than the 122 kHz signal. Thus only a small portion of the array will probably produce salient echoes per each individual outgoing click. During echoic interrogation, Hiapo would pivot his head in the horizontal plane, suggesting the dolphin was ensonifying only a portion of the stimuli at a time. Adding multiple targets (e.g., 4 vs. 2 rods from Branstetter et al., 2003) would then not effect his localization ability because peripheral objects will produce increasingly attenuated echoes.

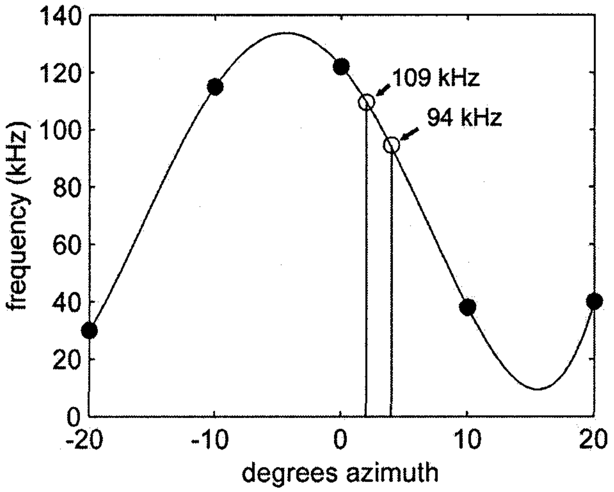

Figure 13

Another mechanism that may aid in producing non-coherent echoes will be the dolphin head related transfer function (HRTF). The dolphin's head behaves as a position dependant spectral filter (Aroyan, 2001; Supin and Popov, 1993; Ketten, 2000) as well as a shaded receiver (Mohl et al., 1999). Echoes from one position in auditory space will be spectrally different in another position. Unlike the pinna which is responsible for HRTF in mammals, the most likely candidates for dolphin HRTFs are internal anatomical structures such as the lipid jaw channels (Ketten, 2000; Aroyan, 2001).

Simulation of Experiments I and II

The results from experiment I and II are consistent with a model that suggests a psychophysical spatial filter is centered at each sound source location which blurs the sound source's position. When sound sources are close enough to produce spatial filter overlap, ambiguity with respect to the sound source position will occur. This is the main effect which produces the psychometric functions. However, the reason for the threshold differences between experiment I and II require an additional mechanism. One possibility is the theoretical spatial filters progressively broaden as the angular difference decreases. Because θ_w equals 2 degrees for both Branstetter et al., (2003) and experiment I, the filter parameters are identical, thus producing very similar thresholds. However, because θ_w is 4 degrees in experiment II the filter band widths are narrower. The shape of the proposed spatial filter is unknown, but the psychometric functions suggest it may be similar to a normal distribution. The model thus assumes the perceived location of a sound source for every observation is sampled from a normal distribution. Each observation is defined as any echo that results from a single incident signal. Because the model assumes the filter bandwidths change as a function of spatial proximity to adjacent sound sources, a parameter *w* is substituted for the standard deviation resulting in the following variation of the normal probability density function:

$$y = f(x \mid \mu, w) = \frac{w}{\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

where μ is the spatial location of the sound source on the horizontal plane, w is a variable. Although the shape of *f(w)* is unknown, *f(w)* is required to reach an asymptote at larger angular separation where the dolphin is near ceiling level performance (i.e., above 4 degrees) and increases with smaller angular separations. Although, many different functions could be used to model *w*, the following exponential decay function was employed, simply because it fulfills the above requirements:

$$w = ke^{-\theta} + m$$

where*:* *k* and *m* are constants and θ is the angular separation between the two targets under observation. When making a decisions about the location of the S+ stimulus, the model firsts calculates the position of each sound source for a single observation sampled from the normal probability distribution. The model then calculates the distance from the center sound source to each flanking sound source and the sound source with the smallest distance is chosen to be the S+ stimulus. This procedure is repeated for each click within a click train and the proportion correct is calculated.

Figure 14 displays simulation results for experiments I and II. Each data point represents 100 trials where each trail consisted of 100 clicks. The vales for *k* and *m* that produced reasonably good fits were 3 and 0.4 respectively. The auditory mechanisms that govern *f(w)* remain speculative. However, auditory masking and HRTF resolution may be likely candidates.

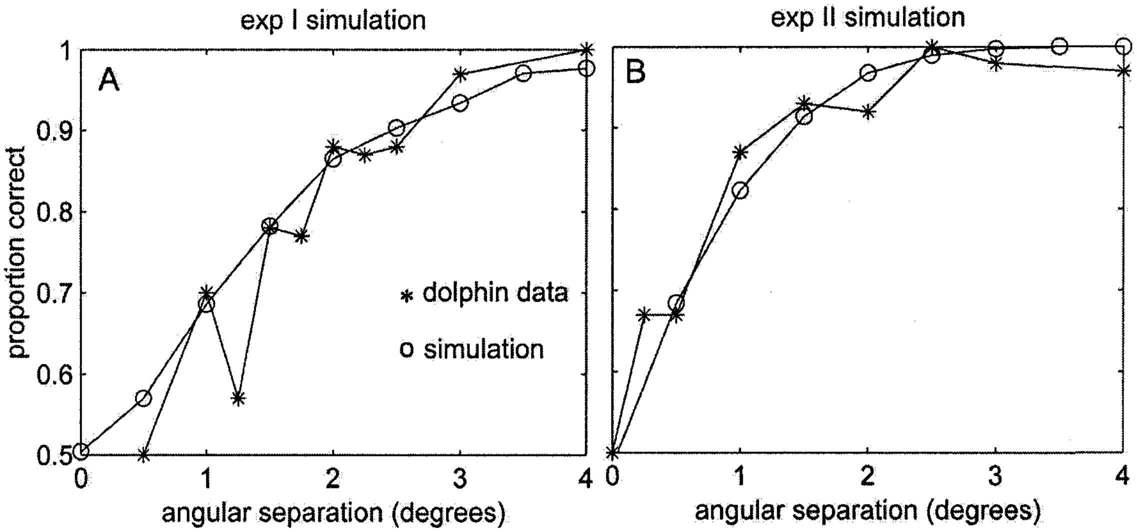

Figure 14. Simulation results compared to dolphin performance in experiment I (A) and experiment II (B).

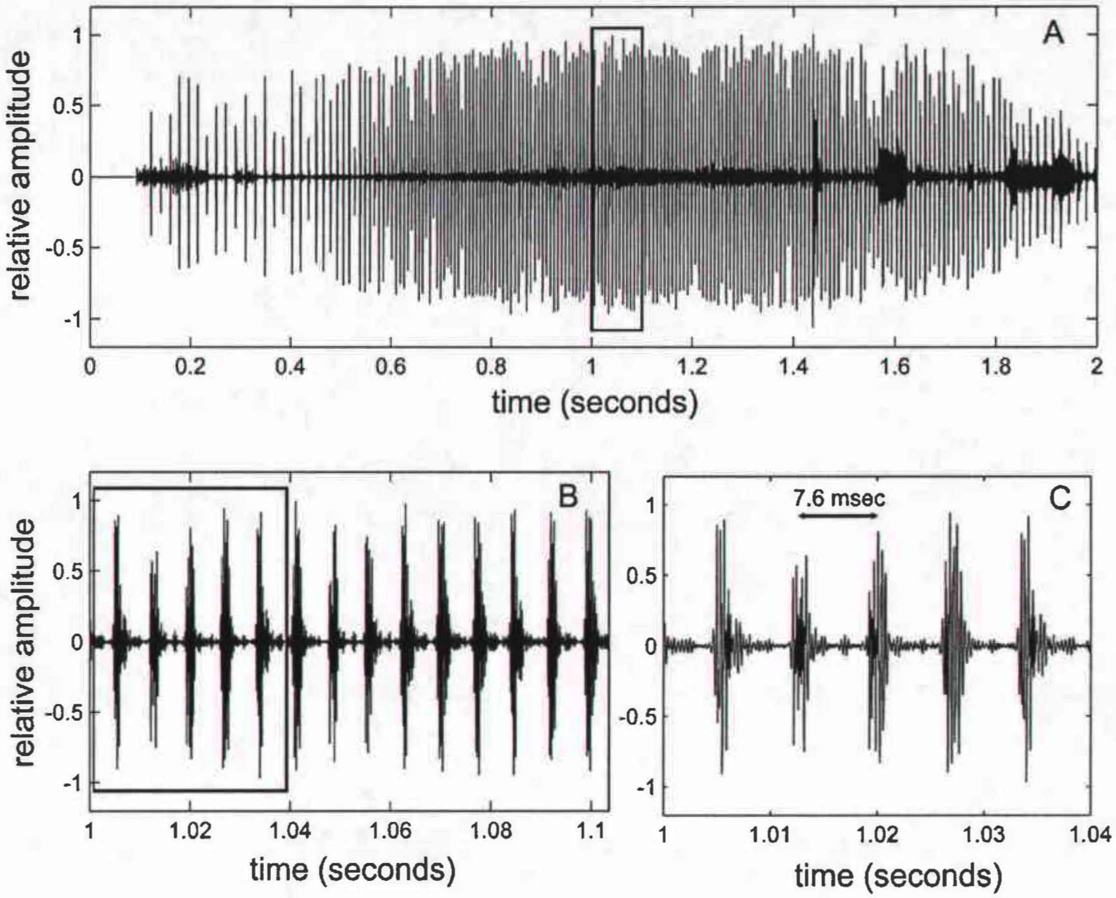

Figure 15. A). A typical click train from a trial. B and C are magnified along the time axis.

Click quantity and response time

Figure 15 is a plot of a typical click train from the current study. The click trains in all of the experiments reported here typically start and end with longer and more variable inter-click intervals (ICI) and lower amplitude clicks. The most stereotyped ICIs have an interval of about 7.6 msec and occur during the majority of the click train. The 2-way sound speed travel time for a target at a 200cm distance is 1.3msec. This suggests the dolphin requires a total of 6.3 msec processing time before emitting a successive click Dolphins typically have processing lag times between 19 and 45 msec for a range of distances (Au, 1993). However, much smaller lag times have been reported for targets at ranges less than half a meter that are more consistent with the current study (Evans and Powell, 1967). Why there appears to be two different processing lag times, one for short distances and one for longer distances remains unknown.

The click-scanning hypothesis makes two predictions: (1) the number of clicks per trial will increase as the angular separation increases, and (2) the mean number of clicks from experiment I would be less than the mean number of clicks from experiment II due to the narrower average angle of the former. Neither predictions were supported. Although the differences in mean clicks between the experiments was not significant, experiment I had an average of 20 more clicks per trial. Therefore, the click-scanning hypothesis can be rejected.

One shortcoming of the current study is the absence of broadband acquisition of the dolphin's sonar signals. Dolphins are known to alter their sonar signals in the amplitude and frequency domain in response to environmental (Au et al., 1985) and sonar task conditions (Au & Penner, 1991; Au, 1980). Because high frequency signals reflect better off of smaller object, the dolphin may selectively increase the higher frequency components of its sonar signal when attempting to resolve fine spatial details. This possibility remains untested.

Auditory mechanisms

Of the auditory mechanisms used for sound source localization, the easiest to evaluate are ITDs. Using a simple two-receiver model (see Appendix II), ITDs were calculated from the obtained angular discrimination thresholds. ITDs for thresholds of 1.6 (Branstettter et al., 2003), 1.5 (Experiment I) and 0.7 (Experiment II) resulted in ITDs of 2.6 μsec, 2.4 μsec, and 1.1 μsec respectively. The smallest dolphin ITD threshold measured by Moore et al. (1995) was 7 μs for click stimuli with a peak frequency of 30 kHz. For click stimuli with peak frequencies between 60-90 kHz, ITD thresholds were between 17-18 μs. The ITD thresholds from the angular discrimination studies are several times smaller than those reported by Moore et al. (1995). In addition, Moore et al. (1995) calculated an ITD, for the dolphin MAA of 1 degree (Renaud and Popper, 1975), would be about 1.3 μs. Again, the MAA ITD was several times smaller than those reported by Moore et al. (1995). There are at least two possibilities for the resulting discrepancies. First, the model used for calculating ITD in the current study was overly simplistic. The model did not take into account reflective or refractive properties of the dolphin's head such as those produced by the skull or the lipid channels in the lower jaw. The internal structures of the dolphin head could possibly increase the ITD threshold. However, it is unlikely that the ITD threshold could be increased by several factor to be consistent with ITDs measured by Moore et al. (1995). Second, the dolphin may not have used ITDs for fine angular discrimination. IIDs, binaural spectral differences or monaural spectral cues may have been used instead. IIDs are more salient at higher frequencies for terrestrial mammals and dolphins (Supin & Popov, 1993). The high degree of sound shadowing produced by the dolphin head (> 20dB; Supin & Popov, 1993) and the high level of IID sensitivity (< 1dB) measured by Moore et al. (1995) suggests IIDs may play a significant and possibly dominant role in sound source localization for fine horizontal angular discrimination. The anatomical structures responsible for producing IIDs and binaural spectral differences remain speculative.

Although IIDs and ITDs can provide the horizontal position of a sound source, these cues typically provide little if any information about source elevation. For terrestrial mammals, vertical and monaural localization are due to the spectral filtering properties of the pinna. The pinna behaves as a position-dependent spectral filter that produces subtle differences in the received sound depending on the location of the sound source. The direction dependent transfer function created by the pinna (and to a lesser degree, the head and torso) is known as the head related transfer function (HRTF). Interestingly, auditory predators often display elaborately convoluted pinnae (e.g., members of the order *Chiroptera*) and asymmetry in external auditory apparatus (e.g., skull asymmetry in the barn owl, *Tyto alba*). Human studies suggests HRTF generated cues are greatest for higher frequencies (Kistler & Wightman, 1992; Middlebrooks & Green, 1992) and broadband sounds. Because dolphin hearing is both broadband with good sensitivity at higher frequencies, HRTFs may be a likely candidate for vertical, and perhaps horizontal sound localization.

Conclusion

The ability to localize fine spatial details through sonar no doubt have several advantages for wild dolphins living in visually restricted environments. For example,. the Amazon River dolphin (*Inia geoffrensis*) feeds in shallow murky waters with very limited visibility. Still others such as the Hawaiian Spinner dolphin (*Stenella longirostris*) are nocturnal feeders (Perrin & Gilpatrick, 1994). Most dolphins typically prey on small fish and squid and are restricted to capturing prey in a serial fashion. Thus, a dolphin may be required to echoically isolate and capture prey tens to hundreds of times within a 24-hour period. This taxing requirement demands a sonar system well adapted for precise target localization through echolocation. Echoic localization may also be important for protection from potential predators such as sharks or other marine mammals. Echoically locating and identifying predators at a distance may provide a dolphin with enough time to evade predation.

Although horizontal echoic angular resolution has been measure for a handful of conditions, only a single passive listening study has measure sound source localization in the vertical plane (Renaud and Popper, 1975). Despite the lack of binaural stimulus differences, the animal's MAA was slightly smaller in the vertical plane. If this result can be corroborated, an even stronger case that dolphins employ HRTFs for fine sound source localization and echoic imaging can be made.

Appendix I.

The arrays were centered on the wooden arc so the far end of the S+ array and the far end of the S- array were an equal distance to the center of the arc and hence, the dolphin's median plane. The center of the stimuli can be calculated by:

$$\frac{\theta_w + \theta_b}{2}$$

As θ_b was varied, the distance from the center of the rack to the furthermost ends of the stimuli also varied. The position of the center of θ_b relative to the center of the arc can be calculated by:

$$D = \left(\theta_w + \frac{\theta_b}{2}\right) - \left(\frac{\theta_w + \theta_b}{2}\right)$$

where D is the distance from the center of the arc to the center of θ_b. The solution for D is the constant $\theta_w/2$.

Appendix II

Because the sound path of the dolphin hearing apparatus is complicated and internal, a simple geometric model was used to estimate time differences between the ears. The distance between the receivers (R_d) in the model was 14 cm, reflecting the measured distance between the estimated position of the left and right mandibular windows of Hiapo's lower jaw. The distance from the sound sources and the midpoint between the receivers (r) was 220 cm. The distance from the left receiver to the sound source (d_L) and the distance between the right receiver (d_R) and the sound source can be calculated using a variation of the law of cosines:

$$d_R = \sqrt{\left(\frac{R_d}{2}\right)^2 + r^2 - 2\left(\frac{R_d}{2}\right) r \cos(90 - \Delta\theta)}$$

$$d_L = \sqrt{\left(\frac{R_d}{2}\right)^2 + r^2 - 2\left(\frac{R_d}{2}\right) r \cos(90 + \Delta\theta)}$$

where $\Delta\theta$ is the angular difference. ITDs can then be calculated by:

$$ITD = \sqrt{\left(\frac{d_L - d_R}{c}\right)^2}$$

Where c was 1520 m/s, the speed of sound in sea water.

REFERENCES

Aroyan, J. L. (2001). Three-dimensional modeling of hearing in *Delphinus delphis*. *Journal of the Acoustical Society of America, 110*(6), 3305-3318.

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

Au, W. W. L., Carder, D. A., Pennerm R. H., and Scronce B. L. (1985). Demonstration of adaptation in beluga wale echolocation signals. *Journal of the Acoustical Society of America,* 77, 726-730.

Au, W. W. L. and Pawloski, D. A. (1992). Cylinder wall thickness discrimination by an echolocating dolphin. *Journal of Comparative Physiology*. A. 172, 41-47.

Au, W. W. L. and Penner, R. H. (1981). Target detection in noise by echolocating Atlantic bottlenose dolphin. *Journal of the Acoustical Society of America,* 70, 251-282.

Au, W. W. L. and Turl, C. W. (1983). Target detection in reverberation by an echolocating Atlantic bottlenose dolphin, (*Tursiops truncatus*). *Journal of the Acoustical Society of America,* 73, 1676-1681.

Au, W. W. L. (1980). Echolocation signals of the Atlantic bottlenose dolphin (*Tursiops* *truncatus*). In: R.G. Busnel and J. F. Fish, eds., *Animal Sonar Systems*. New York: Plenum Press, pp.251-282.

Azzali, M., Manzini, A., & Buracchi, G. (1995). Acoustic recognition by a dolphin of shapes. In R. A. Kastelein & J. A. Thomas & P. E. Nachtigall (Eds.), *Sensory systems of aquatic mammals* (pp. 137-156). Woerden, The Netherlands: De Spil.

Branstetter, B. K., Mevissen, S. J., Herman, L. M., Pack, A. A., & Roberts, S. P. (2003). Horizontal angular discrimination by an echolocating bottlenose dolphin *Tursiops truncatus*. *Bioacoustics, 14*, 15-34.

Butler, R. A., Humanski, R. A., & Musicant, A. D. (1990). Binaural and monaural localization of sound in two-dimensional space. *Perception, 19*(2), 241-256.

Evans W. W., and Powell, B. A. (1969). Discrimination of different metallic plates by an echolocating delphinid. In: R. G. Busnel, ed., *Animal Sonar Systems: Biology and Bionics.* Laboratoire de Physiologie Acoustique, Jouy-en-Josas, France, pp. 363-382.

Harley, H. E., Putman, E. A., & Roitblat, H. L. (2003). Bottlenose dolphins perceive object features through echolocation. *Nature, 424*(6949), 667-669.

Harley, H. E., Roitblat, H. L., & Nachtigall, P. E. (1996). Object representation in the bottlenose dolphin (*Tursiops truncatus*): integration of visual and echoic information. *Journal of Experimental Psychology of Animal Behavioral Processes, 22*(2), 164-174.

Herman, L. M., Pack, A. A., Hoffmann-Kuhnt, M. (1998). Seeing through sound: Dolphins (*Tursiops truncatus*) perceive the spatial structure of objects through echolocation. *Journal of Comparative Psychology*, 112, 292-305.

Jeffress, L. A., Blodgett, H. C., and Deatherage, B. H. (1962). Effects of interaural correlation on the precision of centering noise.. *Journal of the Acoustical Society of America*, 34, 1122-1123.

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

Ketten, D. R. (2000). Cetacean ears. In W. W. L. Au & A. N. Popper & R. R. Fay (Eds.), *Hearing by whales and dolphins* (pp. 43-108). New York: Springer.

Kisler, D. J., Wightman, F. L. (1992). A model of head-related transfer functions based on principle components analysis and minimum-phase reconstruction. *Journal of the Acoustical Society of America*, 91, 1637-1647.

Langendijk, E. H. A., Wightman, F. L., & Kistler, D. J.(2001). Sound localization in the presence of one or two distracters. *Journal of the Acoustical Society of America,* 109, 2123-2134, 2001.

Levitt, H. (1971). Transformed up-down procedures in psychoacoustics. *Journal of the Acoustical Society of America*, 49, 467-477.

Middlebrooks, J. C., Green, D. M., (1992). Observations on a principle components analysis of head-related transfer functions. *Journal of the Acoustical Society of America*, 92, 597-599.

Moore, P. W. B., Pawloski, D. A., and Dankiewicz, L. (1995). Interaural time and intensity difference thresholds in the bottlenose dolphin (*Tursiops truncatus*). In: *Sensory Systems of Aquatic Mammals*. Eds. R. A. Kastelein, J. A. Thomas and P. E. Nachtigall. De Spil Publishers, Woerden, The Netherlands

Moore, P. W. P., Hall, R. W., Friedl, W. A., and Nachtigall, P. E. (1984). The critical interval in dolphin echolocation: What is it*?* *Journal of the Acoustical Society of America*, 76, 314-317.

Murchison, A. E. (1980). Detection range and range resolution of echolocating bottlenose porpoise (*Tursiops truncatus*). In: R.G. Busnel and J. F. Fish, eds., *Animal Sonar Systems*. New York: Plenum Press, pp.43-70.

Nachtigall, P. E. (1980). Odontocete echolocation performance on object size shape and material. In: R.G. Busnel and J. F. Fish, eds., *Animal Sonar Systems*. New York: Plenum Press, pp.71-95.

Pack, A. A., and Herman, L. M. (1995). Sensory integration in the bottlenosed dolphin: Immediate recognition of complex shapes across the senses of echolocation and vision. *Journal of the Acoustical Society of America*, 98, 722-733.

Pack, A. A., Herman, L. M., & Hoffmann-Kuhnt, M. (2004). Dolphin echolocation shape perception: from sound to object. In J. A. Thomas & C. F. Moss & M. Vater (Eds.), *Echolocation in bats and dolphins* (pp. 288-298). Chicago: University of Chicago Press.

Pack, A. A., Herman, L. M., Hoffmann-Kuhnt, M., & Branstetter, B. K. (2002). The object behind the echo: dolphins (*Tursiops truncatus*) perceive object shape globally through echolocation. *Behavioral Processes, 58*(1-2), 1-26.

Perrin WF, Gilpatrick JW Jr (1994) Spinner dolphin - Stenella longirostris (Gray, 1828). In: Handbook of Marine Mammals (Ridgway SH, Harrison SR eds.) Vol. 5: The first book of dolphins. Academic Pres, London, pp. 99-128.

Renaud, D. L. and Popper, A. N. (1975). Sound localization by the bottlenose porpoise, Tursiops truncatus. *Journal of Experimental Biology*, 63, 569-585.

Supin, Y. A., and Popov, V. V. (1993). Direction-dependant spectral sensitivity and interaural spectral difference in a dolphin: Evoked potential study. *Journal of the Acoustical Society of America*, 93, 3490-3495.

Wightman, F. L., and Kistler, D. (1995). Sound localization. In*:* *Human Psychophysics*. Edited by W. A. Yost, A. N. Popper, and R. R. Fay. Springer-Verlag.

Wightman, F. L. and Kistler, D. J. (1997). Sound localization in the presence of multiple distracters. *Journal of the Acoustical Society of America, 101, 3105.*

Chapter 2. Computational Models for Dolphin Auditory Discrimination and Localization

INTRODUCTION

The sonar signals of the dolphin appear well adapted for detecting and discriminating between targets (Nachtigall, 1980; Au, 1993), extracting shape information from complex objects (Harley et al.,, 2003; Harley et al.,1996; Pack & Herman, 1995; Pack et al., 2004; Pack et al., 2002) as well as localizing echoes in three-dimensional space (Branstetter et al., 2003). During echolocation, dolphins typically emit a series of "clicks" where the click interval is sufficient for the dolphin to receive an echo before emitting another click. Each click, or incident signal can be described as a broadband transient with peak frequencies between 40-140 kHz , durations as short as 40 μsec and peak-to-peak sound pressure levels often exceeding 220 dB re:1 μPa (Au, 1980). the signal is projected forward in a tight beam, elevated about 5 degrees above the rostrum having a 3 dB beamwidth of approximately 10 degrees (Au, 1980). Figure 1 illustrates a single dolphin click in the time domain, and its power spectral density. Such visual representations, as figure 1, have been invaluable for describing the physical attributes of dolphin phonations, given that much of the sounds they produce are well out of the human frequency sensitivity range. Because dolphins produce and hear high frequency broadband sounds (Au, 1980; Johnson, 1967), it is likely that these sounds are biologically significant. However, the visual representations typically employed may not accurately reflect how a dolphin may represent or perceive that sound.

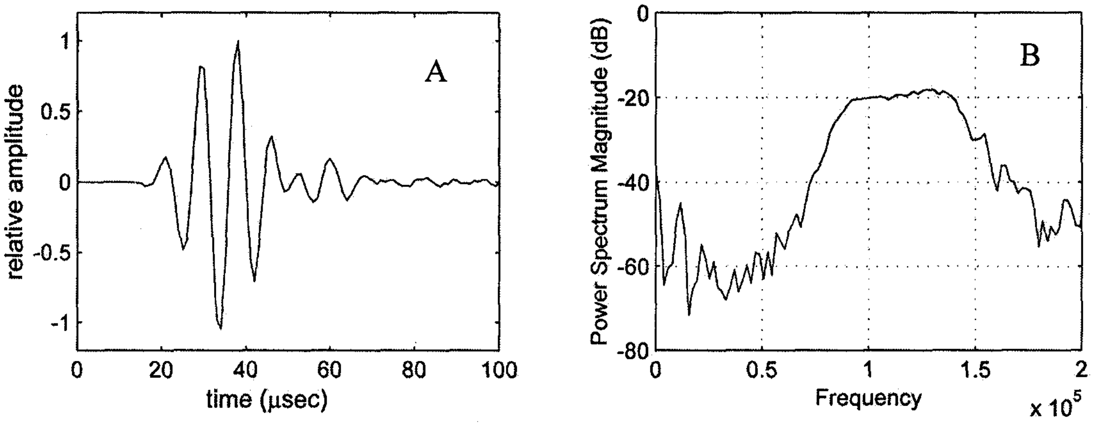

Figure 1. (A) Time amplitude representation of a bottlenose dolphin "click" with (B) the power spectrum magnitude of the click.

For example, Lammers (2003) demonstrated that although spinner dolphin (*Stenella longirostris*) whistle harmonics are lower in amplitude than the fundamental frequency, when a correction factor is applied that takes into account their frequency hearing sensitivity, the first and second harmonics are actually "perceived" louder than the fundamental. Because the higher amplitude harmonics are more directional, dolphin listeners may be able to use these harmonics to determine not only the location of conspecifics, but also, which direction they are traveling. The directional high frequency harmonics may be important for maintaining group cohesion in a visually restricted environment and their importance may have gone unrecognized using traditional acquisition and visualization techniques.

Mercado (1998) applied a model that mimicked peripheral filtering in humpback whales (*Megoptera noveanglea*). When the model was applied to whale song, the salient features of whale song were quite different from those one would extrapolate from a standard time-frequency analysis or a spectrogram. Because the model took into account the spectral resolution of the whale's basilar membrane, (and conventional spectrograms do not) the model representation gives a better approximation to listening from the "whales point of view." Consequently, the frequencies that the whales were most sensitive to happened to be the frequencies that propagated best in the shallow water "breeding grounds" of the Hawaiian Islands.

Descriptions of dolphin sonar abilities (detection, discrimination and matching) have traditionally focused on dimensions related to amplitude, frequency, and temporal characteristics of the echoes. Although this approach has yielded a wealth of valuable data, limitations in its ability to describe more complex tasks (such as cross-modal matching; Pack et al., 2002) suggests the need for more sophisticated biologically inspired representations. To more fully understand dolphin sonar, the representation of sound must closer approximate what the dolphin is receiving and how this received sound is processed by the act of echolocation, dolphin's auditory system and neuronal organization.

Biosonar Transformation Functions

From sound production to perception, each outgoing click will be subjected to (at least) several distinct processing stages: 1) signal production. 2) outward propagation through the dolphin head, 3) forward propagation through the water column, 4) reflection off of a target, 5) propagation back to the dolphin, 6) propagation through the lower jaw region, terminating at the tympano-periotic complex, 7) presumable amplification by the middle ear and transition to the cochlea, 8) basilar membrane displacement and transduction, 9) 8th nerve transmission to central auditory centers where the information is organized.

Stages 3-6 are considered passive processing stages because there is no active amplification by mechanical or neural mechanisms. Stages 1,7, 8 and 9 are active processes because they are under at least partial physiological control. Stage 2 is both passive and active, in that dolphins have been observed (particularly *Delphinapterus leucas*) actively distorting their melons which presumably alters the acoustic characteristics of the outgoing signal. Each stage can be represented by a transfer function that alters the acoustic characteristics of sound. Let *f(t)* be a signal in the time domain and *h(t)* a transfer function. The time-domain filter response *g(t)* can be defined by the convolution:

$$g(t) = \int_{-\infty}^{\infty} f(t) * h(t-\tau)dt$$

Convolution in the time domain is equivalent to multiplication in the spectral domain, and the later is computationally less taxing. The Fourier transform of the time domain signal *f(t)* is:

$$F(S) := \{\Im(f(t))\} := \int_{-\infty}^{\infty} e^{-ist} f(t)dt$$

Where F(S) is complex with both a real and imaginary part. It follows that *g(t)* can be rewritten as:

$$G(S) = F(S)*H(S)$$

Where *G(S)* is the filter response in the spectral domain,

F(S) is the Fourier transform of *f(t)*, and

H(S) is the Fourier transform of *h(t)*

Because the dolphin employs a binaural hearing system, stages 1-6 can be represented by the equations:

$$G_L(S) = F(S)*H_I(S)*2P(S)*T(S)*H_L(S)*N(S), \text{ and}$$

$$G_R(S) = F(S)*H_I(S)*2P(S)*T(S)*H_R(S)*N(S).$$

Where: $G_L(S)$ and $G_R(S)$ are the received signals at the tympano-periotic complex for the left and right ears respectively,

$F(S)$ is the incident signal is the spectral domain,

$H_I(S)$ is the direction dependant filtering of the incident signal by the dolphin head,

$P(S)$ is filtering due to propagation in the water column,

$T(S)$ is the transfer function of the target,

$N(S)$ is noise

and $H_L(S)$ and $H_R(S)$ are the head related transfer function of the received signal for the left and right ears respectively. Unlike the sonar equation which is concerned with detectability related to the intensity of the echo, the biosonar transfer functions (BSTs) focuses on the time-frequency content of a signal. Information derived from BSTs are better equipped to describe acoustic information necessary for discrimination and recognition capabilities A more detailed description of each processing stage is described below.

Stage 1. Sound Production

Cranford (2000) provided direct visual evidence (high-speed video endoscope), correlated with air pressure measurements (catheter) to support the hypothesis that the anatomical apparatus responsible for mechanically producing both clicks and whistles is the monkey lips dorsal bursea (MLDB). The MLDB in the bottlenose dolphin is a bilateral asymmetrical structure above the nasal plug and directly below the blow hole. Evidence suggests the apparatus produces clicks through a pneumonic mechanism. Clicks with lower peak frequencies are produced at the right MLDB, while clicks with higher peak frequencies at the left MLDB.

Stage 2. Propagation in the Dolphin Head

The signal is then projected forward in a 3-dB, 10 degree beam (Au, 1980) by reflecting from the upper portions of the asymmetric concave skull (Aroyan et al., (2000) and focusing produced by the impedance properties of the lipid melon. The melon, a lipid structure composed of triacylgycerol and wax esters (Varanasi and Malin, 1971) is located in the protruding part of the dolphin head anterior to the blowhole. The differential density profile of the melon has similar impedance properties as sea water. Sound is focused inward due to a low velocity core, and a high velocity outer-shell (Norris and Harvey, 1974), thus creating a compact beam. The incident signal is not only directional in the amplitude domain, but also directional and asymmetric in the frequency domain. Au, (1980) demonstrated that the incident signal recorded at -20, -10, 0, 10, and 20 degrees azimuth produced peak frequencies of 30, 115, 122, 38, and 40 kHz respectively. An asymmetric pattern is also produced for angles of elevation.

Aroyan et al.(2000) developed a computer model based on CT-scan density profiles from a dolphin head. The model suggested the apparent asymmetry in both the amplitude and frequency domains are caused by asymmetric anatomical features. For example, the area of the skull most responsible for producing forward reflections appears to be the asymmetric right narial depression.

The direction dependant filtering effects of the head can be represented by:

$$G(S) = F(S) * H_I(S),$$

where $F(S)$ is the Fourier transform of the incident signal produced by the MLDB, and $H_I(S)$ is the direction dependant filtering effects of the dolphin's head (skull, melon, and other anatomical features) in the frequency domain. $H_I(S)$ is not a single transfer function but a matrix of transfer functions corresponding to Cartesian coordinates. For example, each Cartesian coordinate will be associated with a unique transfer function.

Stage 3. Propagation Through the Water Channel.

Sound propagation through the water column will depend on many variables such as water temperature, salinity, depth, composition of the sea floor and proximity to the surface. Because of the dolphin's limited ability to echoically detect small targets beyond 100m (Murchison, 1980) the primary biologically relevant variable that will effect propagation is transmission loss (*TL*). Spherical transmission loss will occur when a target is at a short distance or when a target is below the dolphin where surface or floor reflections are not possible. Spherical propagation loss can be defined by:
$TL = 20\log(r)$, where r is the distance to the target. Since echolocation involves sending a signal and receiving an echo, the 2-way transmission loss will be $TL = 40\log(r)$. Because of the short duration of the dolphin's sonar signal, cylindrical spreading is unlikely. Echoes off the water surface or sea floor will not likely interfere with the initial wave front but will arrive delayed in time.

In addition to transmission loss, high frequency sounds will be converted into heat, and thus absorbed by the water column. The attenuation rate due to absorption can be defined by Fisher and Simmons (1977):

$$\alpha = A_2 f_2 f^2 / f_2^2 + f^2 \text{ dB/m}$$

Where: $A2 = (48.8 \times 10^{-8} + 65.4 \times 10^{-10}T)$ sec/m

$f2 = 1.6 \times 10^{7}(T + 273)\exp[-3052/(T + 273)]$ Hz

f = frequency in Hz

T = temperature in degrees Celsius

Transmission loss due to the combined effects of spherical spreading and absorption can be defined by:

$$TL = 20\log(r) + \alpha$$

Figure 2 demonstrates both transmission loss due to spherical spreading, and transmission loss due to a both spherical spreading and absorption. Spherical spreading is the dominant factor while absorption is almost negligible especially for lower frequencies. The values for f and T were 100 kHz and 22 degrees Celsius (the average surface temperature in Hawaii) respectively.

The transfer function for water channel; propagation can be represented by T(S) where T(S) is a distance dependant vector.

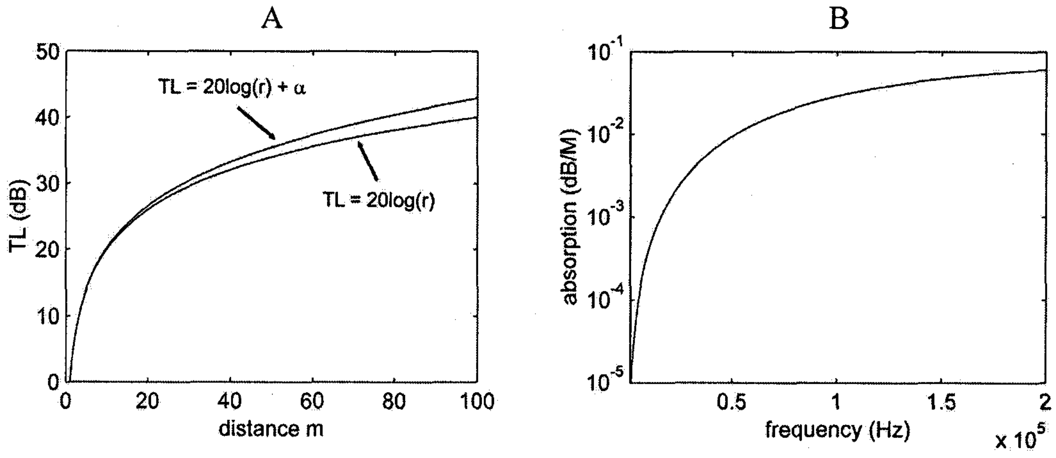

Figure 2 (A) represent both transmission loss due to spherical spreading and the combination of spherical spreading and absorption (B) illustrates absorption (dB/M) as a function of frequency.

Stage 4. Target Echoes.

Echoes are caused by sound reflecting off an object whose impedance properties are mismatched to the acoustic medium. An in depth discussion on the effects of sound interaction between two acoustic media is complex and beyond the scope of this paper. However, Neubauer (1986) provides a good overview of this topic. For simple targets (i.e., a sphere or a cylinder), target transfer function $T(S)$ can be empirically estimated by deconvolving the echo $G(S)$ if the original signal $F(S)$ is known. The equation for a simple target is: $G(S) = F(S)*T(S)$. However, $G(S)$ captures few spatial properties of the target beyond distance to target features. For more complex targets that yield aspect dependant echoes (e.g., Helweg et al., 1996; Pack et al., 2002) a position dependent transfer function, representing a spatial temporal array is a more appropriate representation. In this case, $T(S)$ will be a position dependant matrix of echoes corresponding to Cartesian coordinates of a receiver relative a stationary target. For example, a complex object with multiple features will have a different echo depending on the aspect in which it is ensonified.

Stage 5. Head Related Transfer Functions.

The lower jaw of the dolphin appears to function as a pinna analog (Ketten, 2002). Specialized lipids, with acoustic properties similar to the melon, appear to focus sound towards the tympano-periotic complex. In addition, Converging evidence from anatomical (Ketten, 2002), physiological (Supin and Popov,1993), behavioral (Brill, 2001) , as well as computer modeling (Aroyan, 2001), suggest that odontocetes lower jaws function as position dependant filters known as the head related transfer function (HRTFs). Both the left and right ears will have independent spatial filters: $H_L(S)$ and

$H_R(S)$ respectively, where each is a matrix of transfer function corresponding to sound sources originating from different Cartesian positions. A more detailed description of the HRTF is provided below in the Spectral Models of Sound Localization section.

Stage 6. Middle ear transformation

The function of the middle ear in odontocetes is debatable.. In terrestrial mammals, the middle ear functions as an amplification device to overcome impedance mismatch between the air filled sound medium and the fluid filled cochlea. Amplification is achieved by both a lever and hydraulic action. In odontocetes, the bones of the ossicular chain appear to be calcified, thus precluding analogous functioning (Ketten). In addition, the sound pathway to the middle ear has been altered and is unlikely to include the tympanic membrane (McCormick, et al.,1970). However, recent bone models of the odontocete middle ear suggest that the tympano-periotic complex may function as a double lever amplification mechanism responsible for enhancing hearing sensitivity (Hemila, Nummela, & Reuter, 1999; 2001).

Noise parameter.

Two primary sources of noise exist, internal noise and environmental noise. Environmental noise can come from many sources such as meteorological (surface disturbance, thunder), biologically noise (conspecifics, snapping shrimp), seismic (earthquakes) and human made (shipping noise, navy sonar), (Richardson et al., 1995). Internal noise can be due to both sensory and cognitive factors such spontaneous neuronal firing rates, attention and motivation.

TIME-FREQUENCY RESOLUTION

Dolphin Temporal Resolution

A three-dimensional object can reflect sound from both external and internal surfaces at variable distances. These multiple reflections within a single echo are often called echo-highlights and produce rippled spectra in the frequency domain as well as amplitude modulation in the envelope of the temporal waveform. Dolphins can potentially use envelope information for discriminating between objects (Johnson et al., 1988; Au and Pawloski, 1992). In addition, complex objects with multiple spatial features will produce echoes with a rich amplitude modulated envelope that dolphins may find crucial for extracting shape information. The envelope of the temporal waveform can be estimated using the Hilbert transform:

$$h(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{g(\tau)}{t - \tau} dt$$

where *g(t)* is the temporal waveform. The amplitude of the envelope is then calculated by:

$$ENV(t) = \sqrt{[h(t)]^2 + [g(t)]^2}$$

Figure 3 displays an echo from a metallic cylinder with an outer diameter of 37.85 mm, a wall thickness of 6.35 mm and a length of 12.7 cm (from Au and Pawloski, 1992). The first highlight or envelope peak results from a reflection off the outer surface of the cylinder. The second highlight is produced by sound propagating though the cylinder and reflecting off the back wall. Additional highlights are produced by reverberation within the cylinder and circumferential waves .

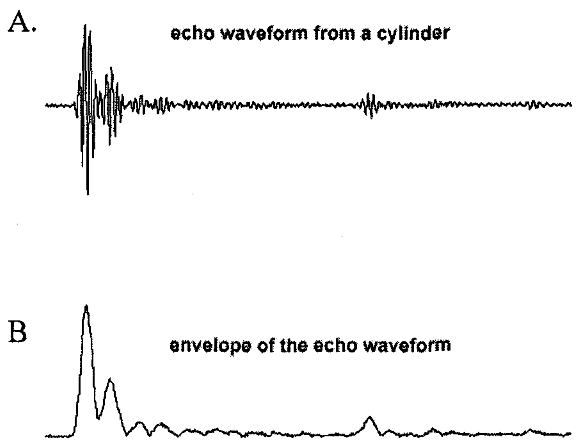

Figure 3. (a) displays the echo from a cylinder and (b) is the envelope of the echo calculated using the Hilbert transform.

The dolphin's ability to process envelope information or "echo highlights" is limited by the temporal resolution of its auditory system. At least two primary auditory mechanisms appear to limit the ability of the mammalian auditory system to resolve the temporal envelope: 1) temporal resolution at the auditory periphery (Zwicker, 1990; Goldstein, 1965) and 2) a central auditory mechanism (Viemeister, 1979). Human listening studies have shown that temporal resolution for sinusoidal amplitude modulated (SAM) tones is greater at higher carrier frequencies than lower carrier frequencies (Zwicker, 1990, Strickland and Viemeister, 1997). This is due in part to the frequency resolving characteristics of the basilar membrane. The basilar membrane response is often modeled as a series of continuously overlapping band pass filters. At low frequencies the filters are relatively narrow. However, as the center frequency of the filter increases, the filter width increases proportionally. Narrow filters produce a relatively high degree of spectral resolution with the trade-off of limited temporal resolution. The opposite holds true for wide filters: decreased frequency resolution but increased temporal resolution.

After transduction and filtering at the basilar membrane, temporal resolution will be limited by the refractory period of the 8th nerve and central auditory areas. Because neurons have refractory periods (whereas hair cells on the basilar membrane have graded potentials limited only by the speed of ionic displacement) the 8th nerve is relatively sluggish.

Several studies have examined temporal resolution in dolphins using broadband stimuli (Moore et. al,1984; Au & Pawloski, 1989; Vel'min & Dubrovskiy, 1975, 1976, 1978) resulting in a critical interval of approximately 264 μs. Dolphin (1995) used sinusoidal carriers to measure the modulation rate transfer function resulting in a typical low-pass filter characteristic similar to humans (Viemeister, 1979).

Dolphin Frequency Resolution

When detecting a sinusoidal signal in the presence of masking broadband Gaussian noise, only the noise within a narrow channel will contribute to the masking of the tone (Fletcher, 1940). The intensity of the signal at threshold (I_{th}) will be:

$$I_{th} = N_o \, x \, \Delta f$$

Where: N_o = the noise spectral density in $\mu Pa^2/Hz$

Δf = the bandwidth of the peripheral filter centered on the signal frequency

Thus, the frequency resolution of peripheral filter (measured in bandwidths) can be estimated my determining the threshold of a signal being masked by broadband noise. The critical ratio in dB is then defined as:

$$CR=10\log(\Delta f)$$

Although the critical ratio is a measure of the auditory filter bandwidth, nothing can be inferred about the shape of the auditory filter employing this method. It is assumed that the filters are rectangular. A more accurate method to obtained direct filter bandwidths is the critical bandwidth measurement. In a band widening experiment, I_{th} is estimated as a function of the bandwidth of noise centered on the signal frequency. As the bandwidth of noise widens, I_{th} increases, but only to a "critical bandwidth." Any noise beyond this critical bandwidth does not effect I_{th}. The auditory filter bank can be described as a bank of constant Q-filters where the bandwidths of the filters are proportional to their center frequencies by the equation:

$$Q = f_o / \Delta f$$

where: f_o is the center frequency of the frequency channel and Δf is the filter bandwidth measured by either the critical ratio or the critical band method. Critical ratio and critical bandwidth estimates for *Tursiops truncatus* have been measured at 12.3 (Johnson, 1968b) and 2.2 (Au and Moore, 1990) respectively.

Precise auditory peripheral filter shapes were measured for *Tursiops truncatus* employing a tone detection in notched-noise paradigm (Lemonds, 1999). A two parameter roex filter model was used to estimate the filter shapes from masking data. The 2 parameter roex filter can be defined by:

$$W(g) = (1-r)(1+pg)\exp(-pg) + r$$

Where g is the deviation from the center frequency of the filter, divided by the center frequency (normalization), and p and r determine the slope and shape of the filter. Because the roex filter is defined in the frequency domain, the impulse response is not well defined (Irino and Patterson, 1997). Therefore, models of auditory filters have often employed time domain models (e.g., gammatone auditory filters) whose frequency responses are very similar to roex filters (see next section).

MODEL OF THE DOLPHIN AUDITORY PERIPHERY

A model was developed to more closely represent sound at various levels of the dolphin auditory periphery. The frequency resolving characteristics of the basilar membrane were modeled employing a gammatone filter bank. Gammatone filters were employed because several models with flexible parameters appear to model a variety of human psychophysical and physiological data well. The current model was adapted from Slaney (1998). Modifications to the gammatone filter bank in the current model include the incorporation of broadband ultrasonic listening of the dolphin and constant Q-ratios from dolphin masking experiments. The impulse response of the gammatone filter is:

$$g_t(t)=at^{(n-i)}\exp(-2\pi bt)\cos(2\pi f_c t-\phi)$$

$$(t > 0),$$

where: f_c is the center frequency of a channel in Hz, ϕ is the starting phase, and a, b, n, are parameters determining ramping and duration of the impulse response and thus filter shape and width in the spectral domain (for more details see Patterson, 1994; Slaney, 1993). The only parameter in the current model taking on a new value is b where $b =$ k**ERB*(f_c). k is a constant fit to human data and is equal to 1.019 and *ERB*(f_c) is the equivalent rectangular bandwidth (Glasberg and Moore, 1990). For current modeling purposes, the general form for the *ERB* was used:

$$ERB(f_c)=\left(\left(\frac{f_c}{Q}\right)^p+(MinBW)^p\right)^{\frac{1}{p}}$$

where: f_c is the center frequency of the filter, Q = 12.3 (Au and Moore, 1990), MinBW is the minimum bandwidth for low frequency channels and in humans is estimated at 24.7 (Glasberg and Moore, 1990), and p=1. Thus, for Tursiops truncatus, ERB(f_c) = 0.0813f_c + 24.7. A total of N overlapping filters are logarithmically spaced between the frequencies of 11 kHz and 156 kHz to reflect the broad range of dolphin frequency sensitivity (Johnson, 1967).

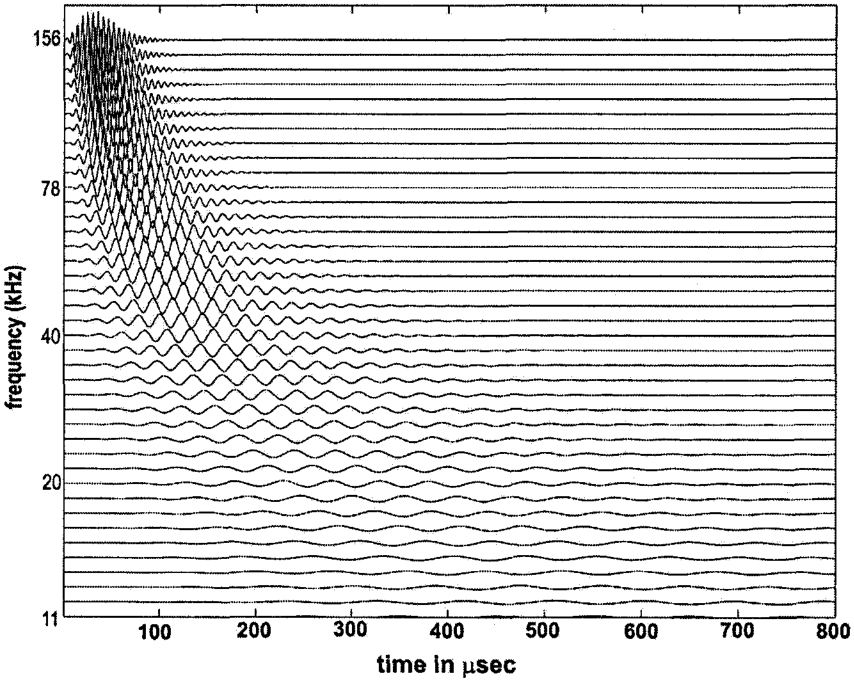

Figure 4. Impulse response of the gammatone filter bank

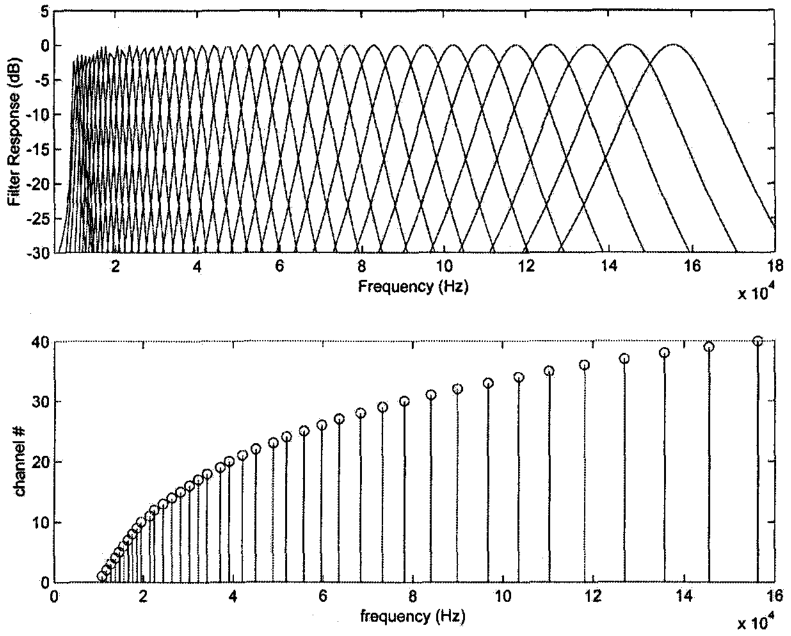

Figure 5. (A) frequency response of the gammatone filter bank. (B) Peak frequency for each channel.

The output of each gammatone filter can then, in parallel, be half wave rectified:

$$f_{rect}(t) = f(t)\left[[f(t)/\sqrt{f(t)^2}+1]/2 \right]$$

Half-wave rectification mimics the afferent activity of an inner hair cell on the basilar membrane. Transduction occurs only when the stereo cilia of inner hair cells are displaced toward the stria vascularis (unidirectional response). Inner hair cells, therefore, produce a graded potential per cycle in a direct coupling with the physical stimulus.

The sluggishness of the 8th nerve and higher auditory areas can be modeled by a low-pass filter defined in the time domain by the exponential decay window:

$$h(t)=ke^{-t/\tau}$$

where k is a constant scaling factor, t is time and $\tau = 264$ μs, the dolphin critical interval suggested by Moore et al. (1984). The output of the low pass filter is the convolution:

$$g(t) = \int_{0}^{T} f_{rect}(t) * h(t-\tau)dt$$

The output $g(t)$ is often referred to as a "leaky integrator" and resembles the envelope of the temporal waveform calculated by the Hilbert transform. However, unlike the Hilbert transform, the leaky integrator output captures the phase locking capability and the envelope following response of the hair cell / auditory nerve complex. Furthermore, leaky integrator output incorporates the temporal resolving characteristics of the dolphin auditory system. In the frequency domain, $G(S)$ is calculated by the equation $G(S)=F_{rect}(S)*H(S)$, where $F_{rect}(S)$ and $H(S)$ are the Fourier transforms of $f_{rect}(t)$ and $h(t)$ respectively. Because multiplication is more mathematically economical than convolution, all computations were conducted in the spectral domain.

Dolphin cochleogram and tursiogram

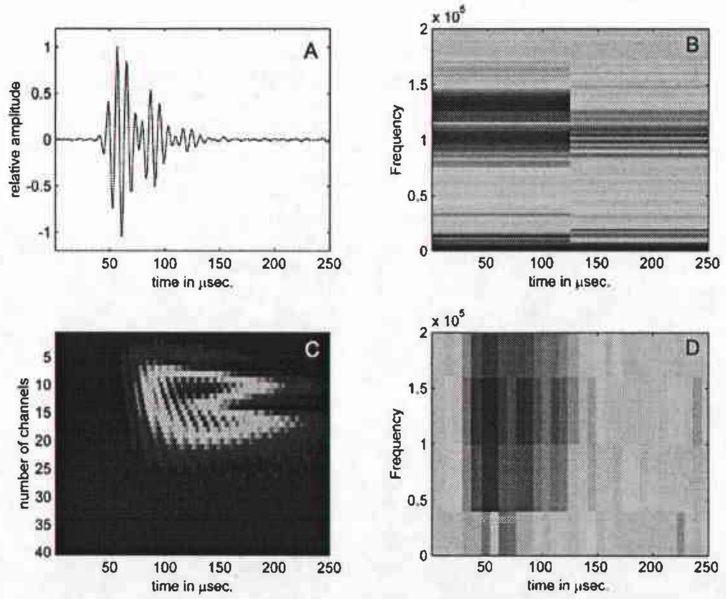

Figure 6. Four different representations of the same echo. (A) echo in the time domain. (B) spectrogram with 512 point hanning window. (C). cochleogram: output of gammatone filter bank. Higher frequencies are at the top. (D) Spectrogram with 16 point hanning window.

Time-frequency representations (spectrograms) of echoes are seldom employed due to an echo's short duration. Figure 6 displays four different representation of an echo sampled at 1 MHz. Figure 6A is the time domain representation. Figure 6B is a spectrogram with a 512 point hanning window with 50% overlap. Figure 6C is a "cochleogram" produced from the output of the gammatone filter bank described above with 40 channels. Figure 6D is a spectrogram with a 16 point hanning window with 50% overlap. The spectrograms in figures 6B and 6D demonstrate problems associated with the spectral-temporal tradeoff inherent with conventional spectrograms. Figure 6B has good spectral resolution with he cost of poor temporal resolution. The opposite hold true for figure 6D, which has good temporal resolution but poor spectral resolution. The cochleogram not only has superior temporal and spectral resolution compared to the spectrograms, but also incorporates the dolphin's spectral and temporal limitations at the level of the basilar membrane, making this representation more similar to what a dolphin actually hears.

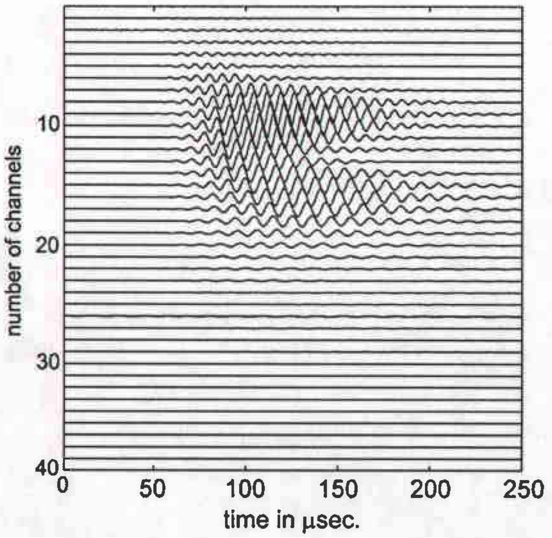

Figure 7. Response of a gammatone filter bank to a recorded echo. The input to the filter is the same echo used in Figure 5.

Figure 7 was produced by the same output of the gammatone filter bank in figure 6c, however, in the current figure, amplitude is represented by displacement from a null at each channel rather than a color gradient as in the previous figure. The current figure is rich in detail and the response of each channel can be evaluated.

Figure 8 is a tursiogram of the same echo from figures 6 and 7. The tursiogram incorporates spectral and temporal resolution of the basilar membrane, hair cells, and 8th nerve. The representation can be considered a more complete model of the dolphin auditory periphery.

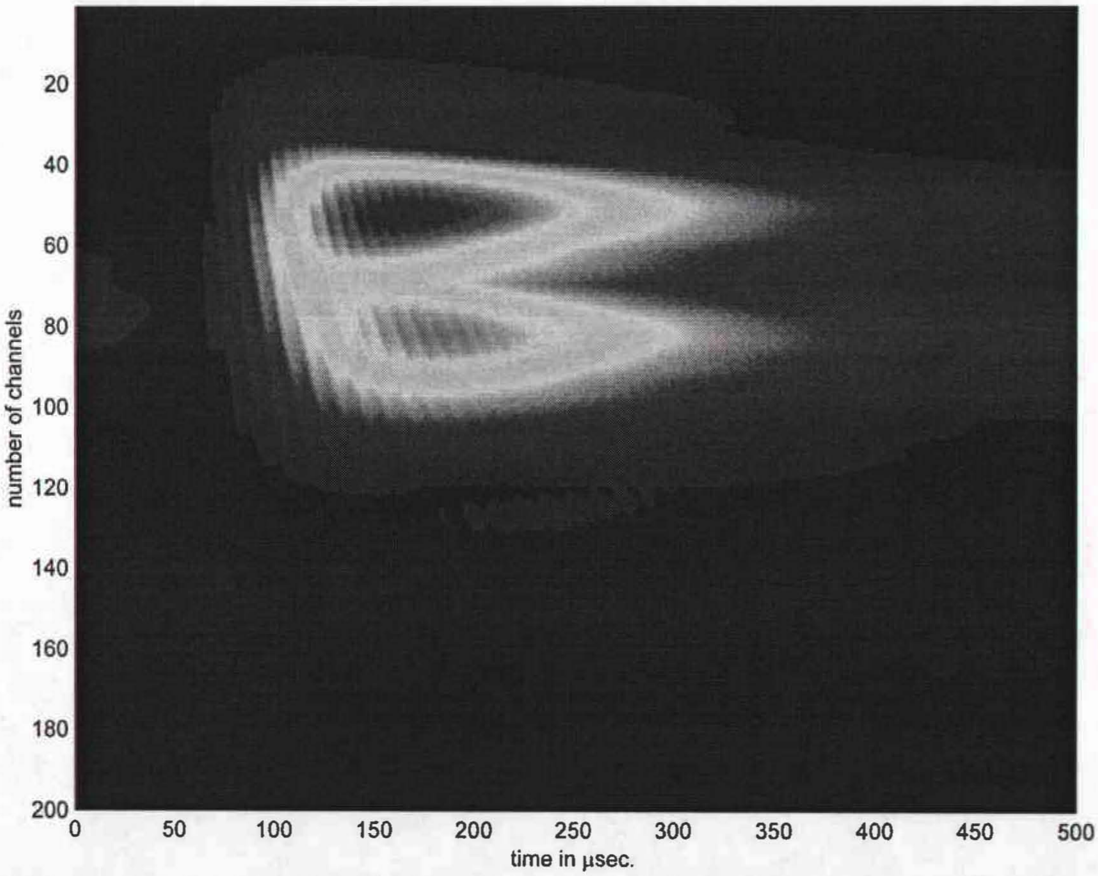

Figure 8. Tursiogram of an echo (same echo as Figures 6 and 7). The representations was created by a bank of 200 gammatone filters spaced between 11kHz and 156 kHz (Q=12.3), half-wave rectification and low-pass filtering. Higher frequencies are at the top.

Representing Multiple Discrimination Cues.

Human listeners can discriminate short duration sounds based on perceptual differences related to pitch (spectral profile), loudness (intensity), and roughness or timbre (temporal envelope). Furthermore, multiple cues can be available simultaneously and different listeners can employ different preferred cues (Southworth and Berg, 1995). The ability to attend to different simultaneously available cues is further supported by the presence of auditory cortical field with differential sensitivities to different stimulus attributes *(*Merzenich & Schreiner 1992*).* Evidence suggests dolphins can also discriminate between sounds differing in cues related to frequency (Herman and Arbeit, 1972), intensity (Evans, 1973) and the temporal envelope (Helweg et al., 1996).

Three biomimetic models were developed that organize received sounds into the three different discrimination cues described above (spectral profile, intensity, and temporal envelope). The models were then tested against each other in simulations of two discrimination tasks previously performed by dolphins to determine which cue was most likely employed by the dolphin.

Envelope detection model

The temporal envelope can be represented by an envelope detection model (EDM). The current dolphin EDM has four stages: gammatone filtration, half-wave rectification, lowpass filtering, and across channel summation (see figure 9). Across channel summation is used to produce a single representation of the "perceived" envelope *ENV(t)* and is calculated by:

$$ENV(t) = \sum_{i=1}^{N} g_i(t)$$

Where *g(t)* is the output from the lowpass filter.

Viemeister and Plack (1993) suggested that information from independent peripheral filters may be pooled to form a "reasonably faithful recovery of the envelope." Evidence from several experiments suggest that information from separated peripheral filters can be pooled together or compared. Such experiments include comodulation masking release (Hall, 1984), modulation detection interference (Yost and Sheft, 1989), as well as several other empirical and theoretical findings (Berg, 1996; Viemeister, 1979).

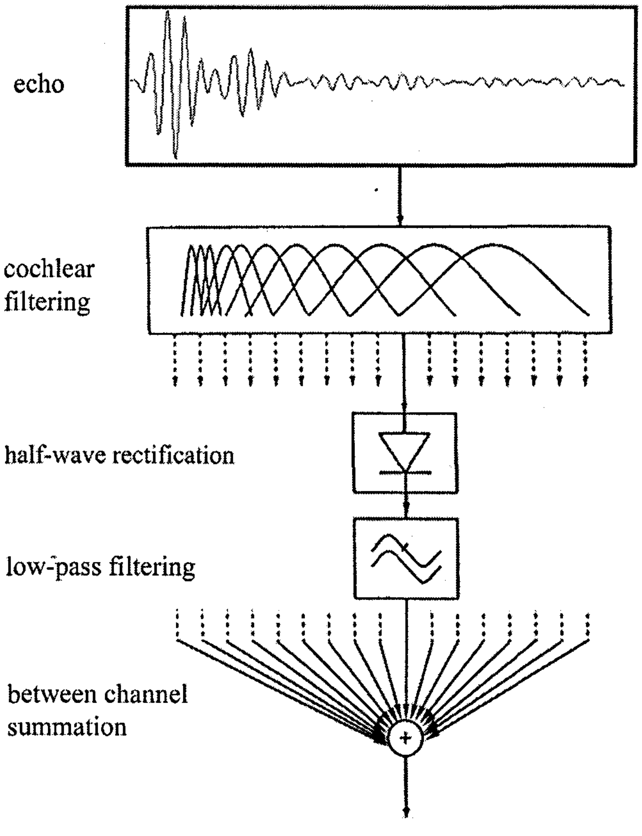

Figure 9. Processing stages of the envelope detection model.

Spectral profile model

The majority of dolphin biomimetic sonar models that have been published base decisions on the spectral profile of the echo (e.g., Au et al., 1995; Au, 1994; Roitblat et al., 1996; Roitblat et al., 1993). The spectral profile can be represented with a four stage model: gammatone filtering, half-wave rectification, lowpass filtering, and within channel summation (see Figure 10). Within channel summation assumes the energy within each channel is summed. The output is a spectral profile resembling the power spectral density function of the original signal with the dolphin temporal and spectral resolution accounted for.

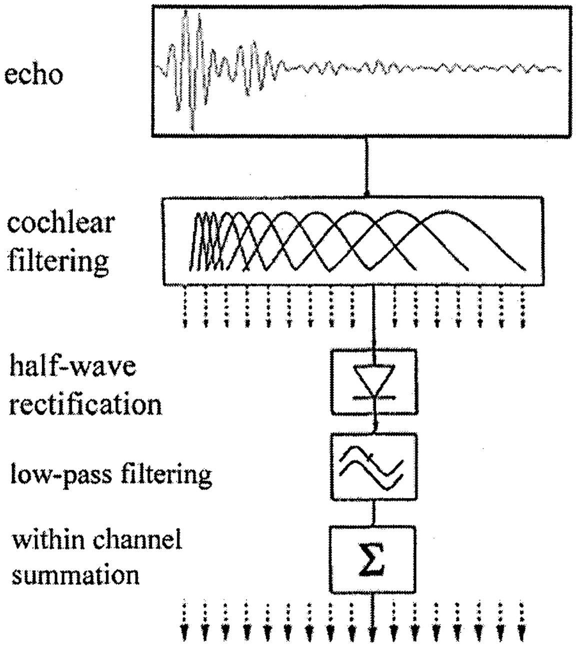

Figure 10. Processing stages of the spectral model

Intensity model

Level cues due to the intensity of a stimulus can be represented by employing the spectral profile model, with the addition of the across channels summation stage from the EDM. The model sums energy within, and between channels to produce a single value representing the perceived intensity of the stimulus (Figure 11).

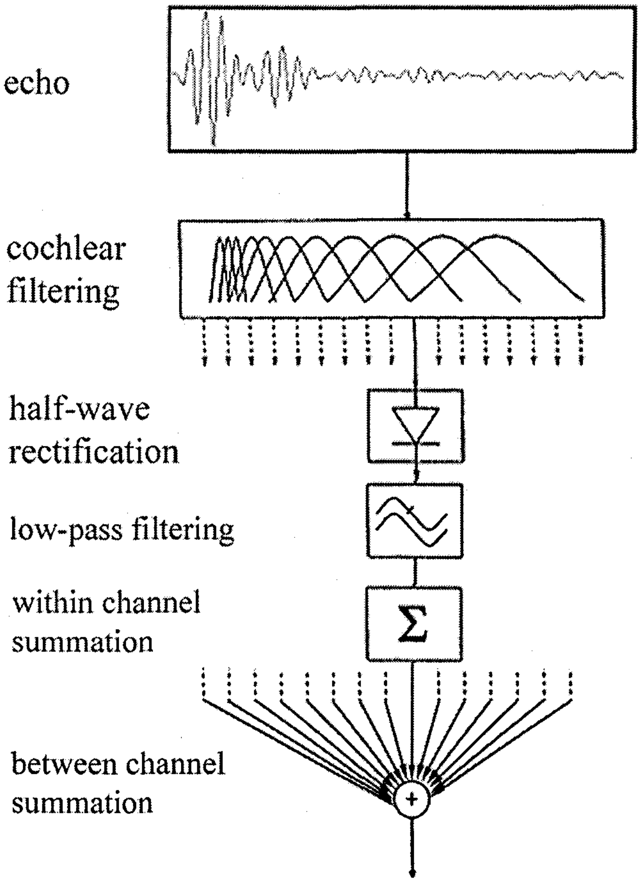

Figure 11. Processing stages of the intensity model

Simulation of temporal order discrimination

In a phase order discrimination experiment (Johnson et al., 1988), a dolphin was trained to discriminate between two click train stimuli: 1) large click preceded a small click (figure $12a_1$) and 2) small click preceded a large click (figure $12b_1$). Each click was a single cycle from 60 kHz sinusoid and the time separation between the clicks was 200 μs. Each click train was composed of 30 echoes. A dolphin was able to discriminate between the two stimuli with 75% accuracy despite the fact that the stimuli have identical spectra, and the time separation between the clicks is smaller the dolphin critical interval. The authors could only speculate how the dolphin was able to perform this discrimination.

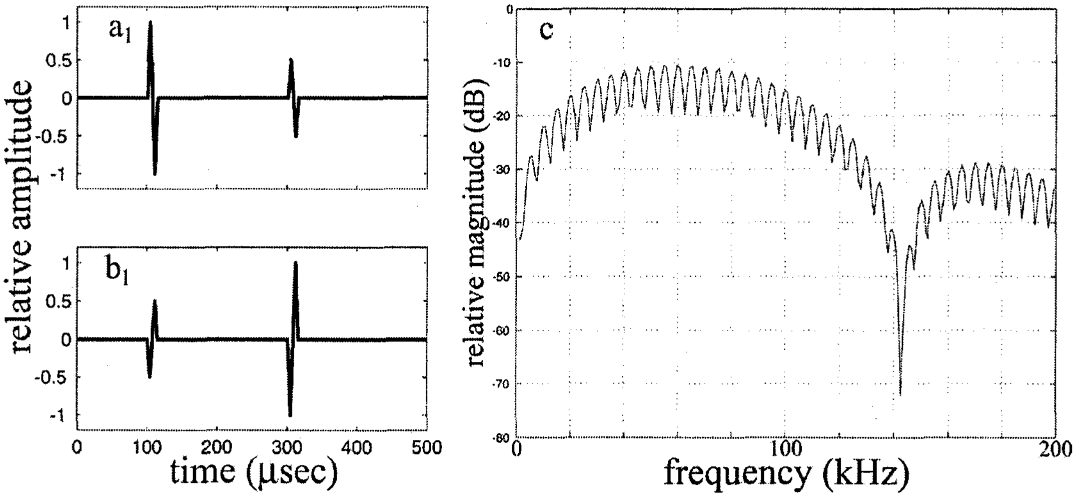

Figure. 12. Click stimuli (a_1 & a_2) and their identical spectra (c) super imposed. Note the two clicks have identical spectra.

The phase order discrimination experiment was simulated with the three models described above within the Matlab 6.0 programming environment. The goal was to pit the models against each other to determine which model would perform best for this specific task. The gammatone filter bank was composed of 40 channels whose center frequencies were spaced between 11 kHz and 156 kHz. Each model was required to perform a 2-alternative match-to sample, with the sample and the target stimuli always being the stimulus with the large amplitude click preceding the small amplitude click (S+) and the non target stimuli was the stimulus with the small amplitude click preceding the large amplitude click (S-). A total of 400 comparisons were made for each model, 200 between the sample and the S+ stimulus, and 200 between the sample and the S-stimulus. All model decisions were based by comparing the Euclidean distance between the sample, and the two alternatives. The Euclidean distance measure can be defined as:

$$d = |x - y| = \sqrt{\sum_{i-1}^{n} | x_i - y_i |^2}$$

where x is the sample vector and y is the alternative vector. The alternative that resulted in the smallest Euclidean distance was selected as the match and the percentage of correct selections were calculated. Noise was added to the stimuli to degrade model performance so the best model would perform at 75%. Figure 13 displays model outputs for the spectral and EDM. Because the intensity model results in a single value, its output was not displayed. Even before noise is added, casual visual inspection of the model outputs suggest that the EDM (figures 13, A3 and B3) will produce the greatest Euclidean distances.

Results

The EDM performance was superior to the spectral and intensity models. Internal noise was added and adjusted to degrade EDM model performance to approximately 75%. This internal noise value was then used for the spectral model and the intensity model, both resulting in chance performance (52% and 46% respectively).

Conclusion

The simulations suggest that the most salient information available to the dolphin within this task would originate from the envelope of the temporal waveform. Changes in the temporal envelope for humans are often perceived as changes in "roughness" or

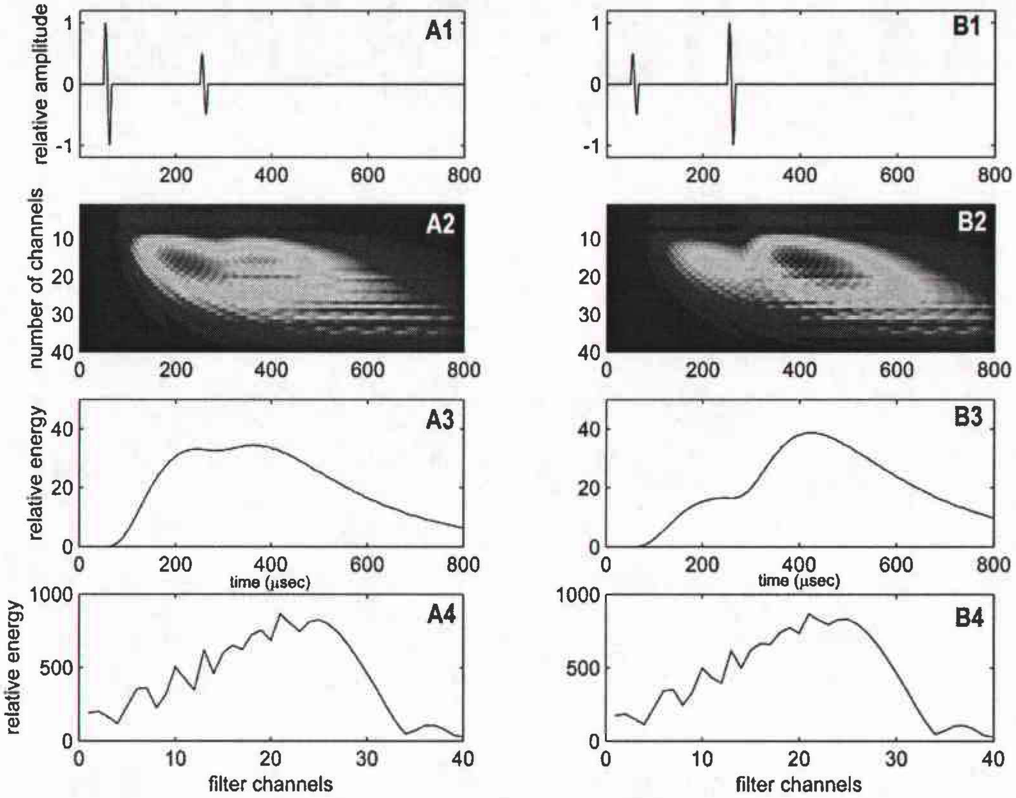

Figure 13. A1 and B1 are the S+ and S- stimuli respectively. A2 and B2 are Turiograms from A1 and B1. A3 and B3 are EDM outputs of A1 and B1. A4 and B4 are the corresponding outputs of the spectral model.

"time separation pitch." What the dolphin perceives remains speculative.

Simulation of wall thickness discrimination

Au & Pawloski (1992) investigated an echolocating dolphin's ability to discriminate wall thickness differences between hollow cylinders. The dolphin was trained to respond only to a standard cylinder of 6.35mm wall thickness while the comparison cylinders varied by ±0.2, ±0.3, ±0.4, and ± 0.8mm. All cylinders had identical length and outer diameter. Results of the experiment are plotted in figure 16. Because the echoes from the cylinders vary in both the spectra (Figure 14c) and the temporal envelopes (Figure 14a and 14b), the dolphin could presumably use either cue to perform this discrimination. Au (1994) simulated the task using a neural network whose inputs were the pattern of energy from a bank of band pass filters mimicking the frequency selectivity of the auditory periphery. This model was very similar in features as the spectral model proposed above and provided good fits to dolphin performance. Au et al., (1995) used a time-frequency model where spectral energy from each channel was evaluated in time increments of $1/\Delta f$ where Δf was the 3dB bandwidth of each peripheral filter. An additional model, limited information within a 264 μsec rectangular window was evaluated. The procedure increased the signal to noise ratio (and thus model performance) by eliminating any noise beyond the windows.

Although all three models are important advancements in biomimetic modeling, the temporal windows make at least 3 assumptions that can be improved upon, 1) the shape of the temporal windows are rectangular, 2) no temporal information within the windows are resolvable, and 3) information within and outside of the windows do not effect each other. Biological auditory systems appear to "smear" information in the time domain rather than partition information into discrete time units (Viemeister, 1979). The exponential decay function (leaky integrator) used in the new models presented above capture this smearing process. Because, the dolphin in the cylinder wall thickness discrimination tasks could potentially use either temporal envelope or spectral information to base decisions, this experiment was simulated with the new models.

The major temporal difference between the echoes of each cylinder was the time separation between the first and second highlights (Au and Pawloski, 1992). For modeling purposes, the stimuli were therefore synthesized by the method from Au &

Pawloski (1992) using the incident signal from Au (1994). Each synthetic echo contained only the first two highlights. The time separation between the first and second highlight ($\Delta\tau$) can be expressed as:

$$\Delta\tau = \frac{2th}{c_1} + \frac{2(O.D. - 2th)}{c_o}$$

where *th* is the cylinder wall thickness, *O.D.* is the outer diameter equal to 37.85 mm, c_o and c_l are the sound velocities in water and the cylinder wall and are equal to 1530 m/s and 5150 m/s respectively (Kinser et al, 1982). Once $\Delta\tau$ was calculated for each cylinder, the relative amplitude of the first and second highlights were estimated by averaging 10 clicks from the recorded standard cylinder from Au (1994). The attenuation of the second highlight relative to the first highlight was calculated by dividing the average amplitude of the second highlight by the average amplitude of the first highlight. A two highlight simulated echo can then be synthesized by the equation:

$$e(t) = s(t) + k[s(t - \Delta\tau)]$$

where: *s(t)* was the synthetic incident signal from Au (1994),

k is the attenuation of the second highlight and

$\Delta\tau$ is the time separation between the first and second highlights.

Echoes from the standard cylinder (sample) were compared with echoes from 2 alternatives: the standard (matching alternative) vs. a comparison (non-matching alternative) in a 2-alternative, matching-to-sample procedure. Each comparison was performed 200 times. The same gammatone filter configuration used in the previous simulation was also employed here.

Results

The spectral model outperformed the EDM and intensity model. When noise was added to produce dolphin like performance, the EDM and intensity model resulted in chance performance (Figure 16). The output of the EDM produced envelopes that are almost identical (Figure 15, A3& B3). All highlight information was smeared and unresolvable. However, clear differences in the spectral models output (Figures 15, A4 & B4) are visible.

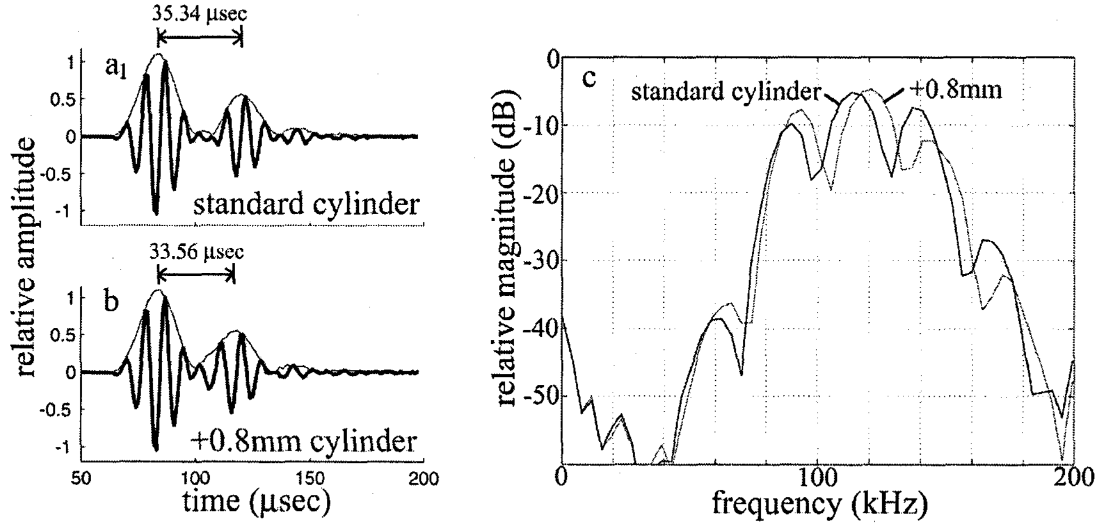

Figure 14. (a1) echo from standard cylinder. (b) echo from +0.8mm cylinder. (c) Spectral content of both cylinders

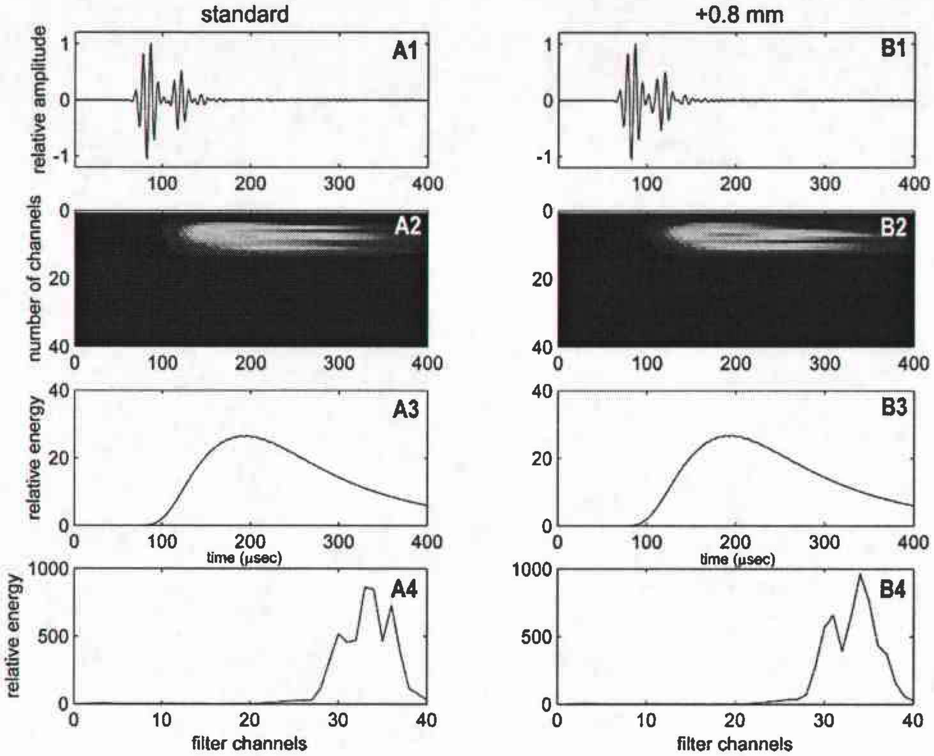

Figure 15. A1 and A2 are echoes from the standard and +0.8mm cylinders respectively. A2 and B2 are their corresponding Tursiograms. A3 and B3 are the correspond EDM outputs. A4 and B4 are the corresponding outputs of the spectral model

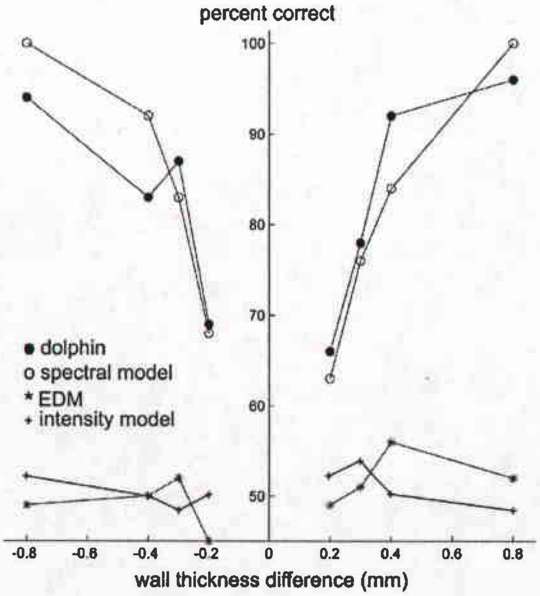

Figure 16. Results from cylinder wall thickness discrimination simulation.

Conclusion

The spectral model provides the most plausible representation of the information dolphins use within this task. The EDM, intensity model and the time-frequency model (Au et al., 1995) are less plausible assuming temporal information within the dolphin critical interval (264μs) is "smeared." These results agree well with simulations from both Au (1994) and Au et al. (1995) where spectral information provides the basis for the dolphin's discrimination capabilities.

General Conclusion

The simulations suggest that different auditory cues are employed for different tasks. Which auditory cues are employed depends on how well the dolphin's auditory system can resolve the temporal and spectral characteristics of the sound. The model could be improved by adding an additional stage that mimics the spectral sensitivity of the dolphin's audiogram.

SPECTRAL MODEL OF SOUND LOCALIZATION

Localization cues

Mammals can localize sound sources by comparing sound differences between the ears (binaural) or by analyzing systematic changes to a sound source at one ear (monaural). The most influential binaural cues are interaural temporal differences (ITDs) and interaural intensity differences (IIDs). Both ITDs and IIDs can be used for localizing sounds in the horizontal plane. Any sound source off the midsagital plane will produce time of arrival differences between the ears (ITDs). The time differences will be a direct function of the speed of sound and the distance between the auditory receivers. If received sounds span a broad range of frequencies, ITDs can be determined from envelope characteristics. But, if the sounds cover a narrow band of frequencies, with little amplitude modulation (e.g., sinusoidal tones), then source location must be determined by comparing phase differences between the ears (called interaural phase differences, or IPDs). Because the wavelengths of sinusoids become progressively smaller as frequency increases, phase difference cues degrade as frequency increases. This is due to the auditory system's limited ability to faithfully track phase information at higher frequencies. In humans, IPDs provide ambiguous information about source location for frequencies above 5 kHz. Because dolphin sonar tends to be composed of high frequencies, it is unlikely IPDs are a primary localization mechanism.

When a sound source is off the median plane, the ears are at different distances R_1 and R_2, from the sound source, resulting in an intensity difference between them equal to:

$$\mathrm{IID} = 20[\log(R_1/R_2)]$$

This type of IID is the result of the difference in distance from each ear to the sound source. In terrestrial mammals, the head will reflect sound, casting a significant sound shadow due to impedance differences between air and a listener's head. The sound shadow is the primary factor in producing IIDs for terrestrial mammals.

Whereas IID, ITD, and IPD cues provide information about the azimuthal position of a sound source, these cues typically provide no information about source elevation or distance. In addition, IIDs and ITDs cannot explain a listeners ability to localize sound monaurally (Butler, Humanski, & Musicant, 1990). For vertical and monaural localization, the spectral filtering properties of the pinna must be employed. The pinna behaves as a position-dependent spectral filter that produces subtle differences in the received sound depending on the location of the sound source. The direction dependent transfer function created by the pinna (and to a lesser degree, the head and torso) is known as the head related transfer function (HRTF). Interestingly, auditory predators often display elaborately convoluted pinnae (e.g., members of the order *Chiroptera*) and asymmetry in external auditory apparatus (e.g., the barn owl, *Tyto alba*). Because each individual has variation in pinna shape, there can be substantial differences in the HRTF between listeners. Statistical models from two independent human studies suggests HRTF generated cues are greatest for frequencies above 5 kHz (Kistler & Wightman, 1992; Middlebrooks & Green, 1992). Furthermore, cues related to HRTFs appear more salient when sound sources are broadband, because level information from independent peripheral filters can be compared across frequency channels. A direct correlation between bandwidth and sound localization accuracy supports this hypothesis (Wightman & Kistler, 1995, pp 181).

If apriori knowledge of the sound source spectrum is known, any monaural change in the sound spectrum will be due to the HRTF. Thus, the position of the sound source can be determined by deconvolution allowing monaural localization. Let $f(t)$ and $h(t)$ represent the source signal and the HRTF in the time domain, respectively. Then $F(S)$ and $H(S)$ are their frequency domain Fourier transformations. The received signal $G(S)$ is calculated by: $G(S) = F(S)*H(S)$. IF $F(S)$ is known from prior experience, and $G(S)$ is known because it is received, $H(S)$ can be calculated by: $H(S) = G(S) / F(S)$. $H(S)$ serves as a direct indicator of sound source position.

For novel sound sources, comparisons between the ears must be made. Let the received sound spectrum at the right ear $G_R(S)$ be:

$$G_R(S) = F(S) * H_R(S)$$

where $H_R(S)$ is the HRTF of the right ear and $F(S)$ the spectrum of the sound source. The received signal at the left ear $G_L(S)$ is then:

$$G_L(S) = F(S) * H_L(S).$$

Let the binaural spectral ratio (BSR) equal the ratio between received spectra between the ears:

$$BSR = G_R(S) / G_L(S).$$

The equation can then be expanded:

$$BSR = F(S) * H_R(S) / F(S)\ H_L(S)$$

and simplified:

$$BSR = H_R(S) / H_L(S).$$

The pattern of the BSR will be unique for every given sound source position if $H_L(S)$ and $H_R(S)$ are complex (broadband) and display response asymmetry with respect to Cartesian position. Because the BSR is unique and the $F(S)$ term cancels out, exact stimulus location can be determined even if $F(S)$ is novel.

Dolphin Localization

The dolphin's ability to localize sound sources has been tested both passively and through echolocation. Renaud and Popper (1975) estimated the dolphin's minimum audible angle (MAA) in both the horizontal and vertical plains for a variety of stimuli. The minimum audible angle is the smallest angle between two sound sources where the sound sources are still perceived as being discrete. Any angle less than the MAA, the sound sources are perceived as an acoustic singularity. For broadband click stimuli with a peak frequency of about 65 kHz, the MAA was 0.9 and 0.7 degrees in the horizontal and vertical planes respectively. For sinusoidal stimuli, the MAA was typically between 2-3 degrees for both the vertical and horizontal planes. Despite the absence of binaural stimulus differences in the vertical plane, the dolphin's localization ability was excellent. In fact, vertical localization was as good for sinusoidal stimuli and slightly superior for click stimuli.

Branstetter et al., (2003) investigated a dolphin's ability to resolve angles between multiple targets in the horizontal plane through echolocation. The dolphin was required to echoically discriminate horizontal angular differences between two arrays of vertical, air-filled, PVC rods. The blindfolded dolphin swam and stationed in a vertically oriented hoop, 2 radial meters from the stimuli, and indicated whether an array with four rods (S+) was to the left or the right of an array with two rods (S-), by pressing a corresponding paddle. The angular separation between the rods within each array (θ_w) was maintained at 2 degrees but the angular separation between the two arrays (θ_b) was varied to produce angular differences ($\Delta\theta = \theta_b - \theta_w$) ranging between 0.25 degrees to 4 degrees. A modified method of constant stimuli was used to test for angular discrimination ability, and yielded a psychometric function having a 75% correct threshold of 1.6 degrees. The results are in good agreement with both MAA measurements from Renaud and Popper (1975) and echoic angular discrimination data with bats (Simmons et al., 1983).

Dolphin Localization Cues

The same localization cues available to terrestrial mammals are potentially available to dolphins. Of these localization cues, ITDs are the simplest to evaluate. Using a simple two-receiver model, Branstetter et al.,(2003) calculated the angular discrimination threshold of 1.6 degrees corresponds to an ITD of 2.6 μs. The smallest reported ITD threshold for a bottlenose dolphin is 7 μs, for click stimuli having a peak frequency of 30 kHz (Moore et al., 1995). For click stimuli with peak frequencies between 60-90 kHz, ITD thresholds increased to between 17-18 μs. The ITD threshold that would be required to support the level of angular discrimination observed in Branstetter et al. is thus several times smaller than those reported by Moore et al. (1995). In addition, Moore et al. calculated that an ITD for the dolphin MAA of 1 degree, reported by Renaud and Popper (1975), would be about 1.3 μs, again a value several times smaller than the empirical measurements reported by Moore et al. There are at least two possibilities for these discrepant ITD values. First, the simple geometric models used for calculating ITDs did not take into account reflective or refractive properties of the dolphin's head, such as those produced by the skull or the lipid channels in the lower jaw, which possibly could increase ITDs. However, it is unlikely that the ITDs would be doubled or even increased by a factor of six to be consistent with ITDs measured by Moore et al. (1995). Second, the dolphin may not have used ITDs for its angular discriminations. Instead, interaural intensity differences (IIDs), binaural spectral differences, or monaural spectral cues may have been employed. IIDs are more salient at higher frequencies for terrestrial mammals and also for dolphins (Supin & Popov, 1993). The high degree of sound shadowing produced by the dolphin head (> 20dB; Supin & Popov, 1993) and the high level of IID sensitivity (< 1dB) measured by Moore et al. (1995) suggest IIDs may play a significant and possibly dominant role in sound source localization for fine horizontal angular discrimination. Exploiting high frequency IIDs may have been a governing factor selecting for high frequency hearing in the suborder Odontoceti. Furthermore, the dolphin head appears to behave as a position dependent spectral filter in much the same manner as the terrestrial mammalian pinna (Aroyan, 2001; Brill et al., 2001; Supin & Popov, 1993).

For example, Supin and Popov (1993) used the auditory nerve response (ANR) to investigate direction-dependant spectral sensitivity of an Amazon river dolphin (Inia geoffrensis). The ANR is the short latency, first wave of the auditory brain-stem-evoked response (ABR). Because the ANR has monaural properties, the ANR is assumed to originate from the auditory nerve. Spectral sensitivity was measured for several frequencies at different azimuthal positions. The results indicated that direction depend sensitivity was greater for higher frequencies. In addition, each angular position produced a unique pattern of spectral sensitivity indicative of a HRTF.

Similar behavioral; results were obtained by Brill (2001). A bottlenose dolphin's behavioral auditory sensitivity for 4 frequencies at different positions along its head.. Rather than presenting the tones from a far field underwater speaker, a small transducer was placed at 41 positions on the dolphin's head by a suction cup. Thresholds for each frequency at each position were then individually estimated. The results indicate an asymmetric pattern of sensitivity, again consistent with a HRTF.

Aroyan (2001) employed a 3-D model of sound propagation based on tissue density and velocity profiles of a common dolphin (Delphinus delphis). Results indicated receptivity patterns differed between the ears for different elevation and azimuth locations. These differences were apparent across frequencies and were asymmetric. The results not only indicated the receptivity pattern was complex but provided evidence that the pan bones and the fat channels in the lower jaw focus sound towards the typanoperiotic complex. Thus, the lower jaw appeared to function as the analog of the terrestrial pinna, focusing sound and functioning as a position dependant spectral filter.

Indeed, if the results from the vertical MAA experiment (Renaud and Popper, 1975) are accurate, position dependant spectral filtering seems the most plausible explanation since simple time and intensity differences should not exist for sound sources along the mid-sagital plane. The particular anatomical structures responsible for spectral filtering remain speculative; however, converging evidence from the above studies suggest that odontocetes have well developed HRTFs that may play a significant role in sound localization.

Neuronal map of auditory space.

Almost nothing is known about the neural computations underlying spatial perception in dolphins to develop biologically realistic computational models of these processes. Nevertheless, it is clear that dolphin auditory systems must perform some calculation that maps cochlear responses onto representations of space. Computational models of auditory localization can also give some sense of the relative complexity of computations required to map biologically relevant acoustic events to spatial positions.

Although data on dolphin auditory spatial maps is lacking, comparative data on another auditory specialist, the barn owl (Tyto alba), provide a model for how such maps are organized. In the midbrain of the owl, the mesencephalicus lateralis dorsalis (MLD) is composed of neurons with receptive fields that respond to specific spatial positions of sound sources relative the owl's head (Knudsen and Konishi, 1978; Knudsen, 1983; 1984;). The receptive fields are systematically arranged according to azimuth and elevation creating a topographical map of auditory space in MLD. The response of each receptive field also appears to be unaffected by changes in sound pressure level or sound type (e.g., white noise, broadband clicks, tonal pulses). Thus, the only parameter these receptive fields are influenced by is the position of the sound source Similar spatial maps have also been found in the optic tectum of the barn owl (Knudsen, 1984). However, these receptive fields not only respond to position dependant auditory events, but also to position dependant visual stimuli. Each receptive field in the optic tectum maps spatial events independent of the sensory modality that registered the event. A similar multimodal integration center that maps spatial positions is likely to exist in dolphins. Recall the various cross modal experiments (Azzali, Manzini, & Buracchi, 1995; Harley, Putman, & Roitblat, 2003; Harley, Roitblat, & Nachtigall, 1996; Pack & Herman, 1995; Pack, Herman, & Hoffmann-Kuhnt, 2004; Pack, Herman, Hoffmann-Kuhnt, & Branstetter, 2002) that provide evidence dolphins can match objects across the senses of echolocation and vision. Because sound (echolocation) and light (vision) are not physical correlates, the spatial structure of the object (its shape) is the only physical parameter by which a match could be achieved. A high resolution neuronal spatial map, similar in properties to the barn owl's optic tectum, is a probable candidate to explain the cross-modal results.

Simulation of dolphin auditory spatial maps

The model presented here attempts to demonstrate the functional characteristics of a dolphin auditory spatial map that responds to HRTF derived information. HRTFs were chosen because their position dependant nature provides both azimuth and elevation information, whereas other localization cues (IIDs, ITDs and IPDs) provide only azimuth information. To the extent possible, the model incorporates the known anatomical, physiological, and behavioral constraints of the dolphin's auditory system. To simulate neuronal receptive fields, an artificial neural network (ANN) was employed. The t binaural ANN model has three main stages simulating passive filtering by the head of the dolphin, active filtering at the dolphin auditory periphery, and experience-based mapping of auditory peripheral responses to positions in two-dimensional space. All computations were performed within the Matlab 6.0 programming environment.

Stage 1: Filtering by the dolphin's head and auditory system.

Because dolphin HRTFs are only available in rudimentary form (Brill, et al 2001; Aroyan, 2001; Supin and Popov, 1993), we have chosen to employ arbitrary HRTFs. Like actual mammalian HRTFs, the arbitrary HRTFs systematically alter the spectrum of incoming sound as a function of sound source position (see figure 17). Two sets of HRTFs were created, one for the right ear, and one for the left ear. No optimization of the HRTFs was attempted. Consequently, model results are considered conservative compared to real HRTFs that have the distinct advantage of long-term evolutionary fine tuning. The incident signal presented to the HRTFs was a dolphin click employed in previous modeling efforts (Au, 1994; Au et al., 1995). The click was sampled at 1 MHz, with a peak frequency between 100 kHz and 140 kHz, and a duration of approximately 40 μsec.

After head related transformation, the signal was binaurally presented the spectral model described earlier (gammatone filter bank, half-wave rectified, lowpass filtered, and summed within channel). This ensures information presented to the neural network will be subjected to similar filtering limitations as the dolphin auditory system.

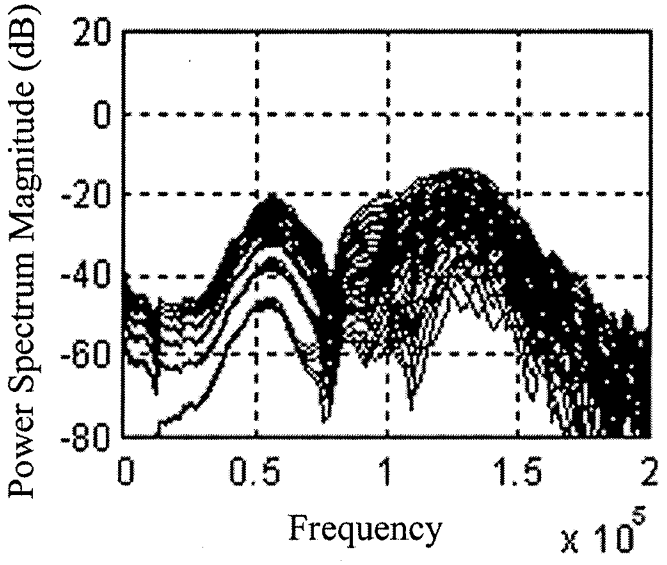

Figure 17. Power spectrum of clicks that have been filtered by 49 different HRTFs that correspond to 49 different spatial positions. The original unfiltered click is the same plotted in figure 1.

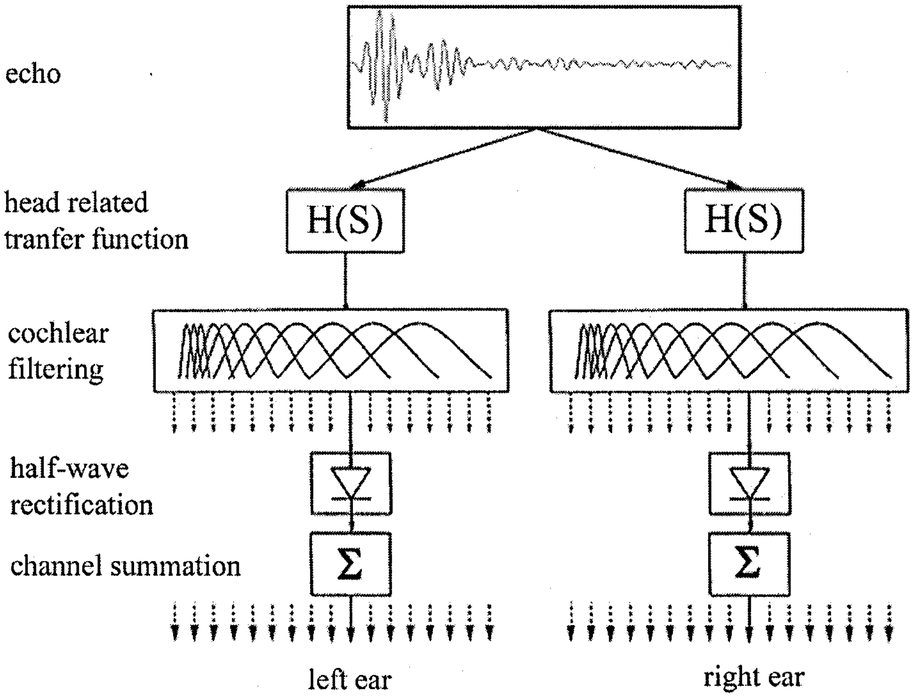

Figure 18. Diagram of the model processing stages prior to neural network presentation

Spectral response profiles calculated in stage one were presented to a three-layer backpropagation neural network that was trained to associate spectral profiles of a sound source with the corresponding Cartesian coordinates of the sound source. The behavior of each neuron in the network was governed by the equation:

$$a = f(Wp+b)$$

where a is the output of the neuron, p is the input vector, W is the weight matrix, b is a bias vector and f is a transform function. The first layer consisted of inputs (p) from the dolphin peripheral filtering stage followed by a hidden layer. Tan-sigmoid transfer functions (f) were used to limit outputs between -1 and 1. The tan-sigmoid transfer function used by Matlab can be defined as:

$$n = \frac{2}{(1+e^{-2n})} - 1$$

The network output space was arranged on a two-dimensional grid to simulate Cartesian coordinates much in the same way neurons are topographically organized in the owls MLD and optic tectum.

Learning proceeded by incrementally adjusting the weights of the individual cells in the network (backpropagation) over a series 5000 iterations or until the rms output error fell below a preset parameter value. A single iteration of the backpropagation algorithm can be expressed as:

$$\mathbf{x}_{k+1} = \mathrm{x}_k - \alpha_k \mathrm{g}_k$$

where x_k is the vector of the current weights and biases, g_k is the current gradient and α_k is the learning rate.

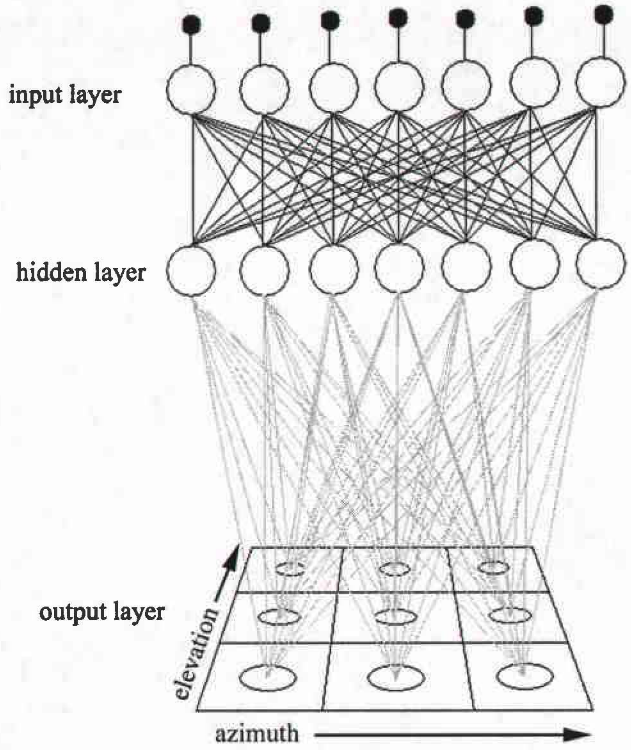

Figure 19. Neural network architecture which simulates mapping of auditory features onto a topographic map of auditory space.

Simulation Results

The connectionist model was able to localize single sound sources in both the horizontal and vertical planes. Figure 20 demonstrates how an echo from two different spatial positions is processed by the model. For simplicity, two examples of monaural processing are presented. Figure 20, A is the echo measured at the sound source. Figure 20, A1 and 20, A2 illustrate the echoes after filtering from two different HRTFs. This is equivalent to position-dependant filtering from two different sound source locations. Figure 20, C1 and 20, C2 are the outputs from the gammatone filters. Figure 20,D1 and 20,D2 show the result of channel summation which is equivalent to the perceived spectral profile. High frequencies are to the left. Figure 20, E1 and 15, E2 are responses in the output layer of the neural network to the processed echoes. Spikes indicate the spatial position of the sound source . Although the current example only illustrates how single sound sources can be mapped to the network, work is currently in progress modeling how a dolphin may localize multiple sound sources simultaneously and ultimately "image" a complex object.

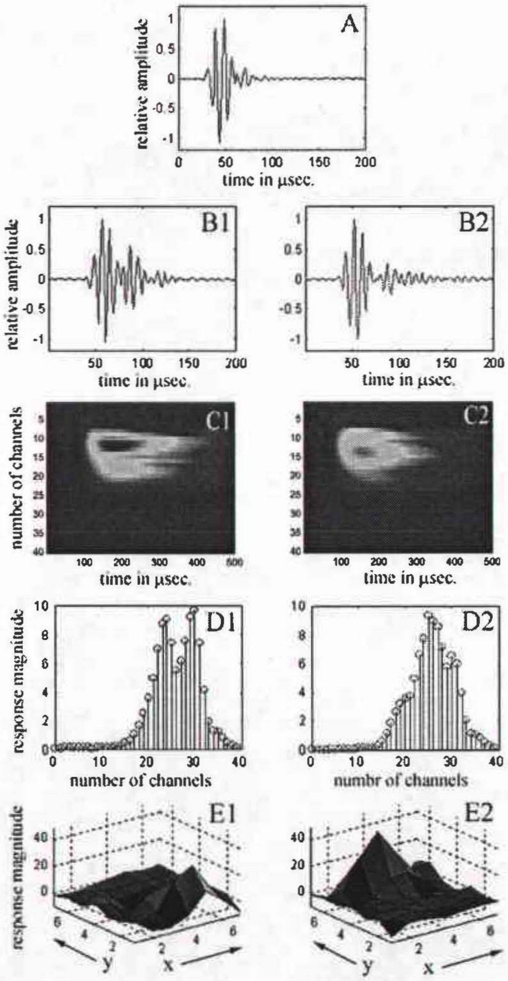

Figure 20. Simulated localization of an (A) echo at sound source; (B1 and B2) echoes from different positions in space have undergone filtration by HRTFs; (C1 and C2) output of gammatone filter bank; (D1 and D2) channel summation; (E1 and E2) neural net response to the echoes identifying the position (elevation and azimuth) from which they originated.

GENERAL DISCUSSION

Living in a visually restricted environment has placed pressure on dolphins to exploit the sound medium. Because sound propagates through water differently than it does through air, cetaceans have revamped the archetypal terrestrial mammalian auditory system with unique aquatic adaptations for producing, receiving, and localizing these sounds. Dolphins have also developed unique neural computational approaches to processing and organizing auditory information, demonstrated by the ability of bottlenose dolphins to immediately visually recognize objects that they previously have only experienced acoustically (Harley et al., 2003; Pack & Herman, 1995; Pack et al., 2002).

Dolphins likely exploit any localization cue available. However, evidence suggest that sensitivity to ITDs may not be sufficient for fine sound source localization. Results from Renaud and Popper (1975) suggest dolphins can localize sounds equally well in both the horizontal and vertical planes despite the fact that there are no binaural stimulus differences in the latter. This result suggests that dolphins may employ HRTF based spectral cues for localization. To date, there are no detailed HRTFs for any odontocete. Acquiring a HRTF would be beneficial in understanding how dolphins can localize sound sources to such an extent that they perceive the spatial structure of complex objects.

Empirical studies with dolphin are often difficult to perform due to the high costs of maintaining and training an animal in captivity. Furthermore, invasive physiological studies are not possible due to legal and ethical concerns. Computational models provide an alternative method for exploring and generating testable hypotheses about dolphin auditory abilities, and of determining whether specific acoustic cues are necessary or sufficient to account for their hearing abilities.

References

Aroyan, J. L. (2001). Three-dimensional modeling of hearing in *Delphinus delphis*. *J Acoust Soc Am, 110*(6), 3305-3318.

Aroyan JL, McDonald MA, Webb SC, Hildebrand JA, Clark D, Laitman JT, Reidenberg JS (2000) Acoustic Models of Sound Production and Propagation. In: Au WWL, Popper AN, Fay RR (eds), ***Hearing by Whales and Dolphins***. New York: Springer-Verlag, pp. 409-469.

Au, W. W. L. (1980). Echolocation signals of the Atlantic bottlenose dolphin (*Tursiops truncatus*). In: R.G. Busnel and J. F. Fish, eds., *Animal Sonar Systems*. New York: Plenum Press, pp.251-282.

Au, W. W. (1994). Comparison of sonar discrimination: dolphin and an artificial neural network. *J Acoust Soc Am, 95*(5 Pt 1), 2728-2735.

Au, W. W. (1996). Acoustic reflectivity of a dolphin. *J Acoust Soc Am, 99*(6), 3844-3848.

Au, W. W., Andersen, L. N., Rasmussen, A. R., Roitblat, H. L., & Nachtigall, P. E. (1995). Neural network modeling of a dolphin's sonar discrimination capabilities. *J Acoust Soc Am, 98*(1), 43-50.

Au, W. W., & Moore, P. W. (1990). Critical ratio and critical bandwidth for the Atlantic bottlenose dolphin. *J Acoust Soc Am, 88*(3), 1635-1638.

Au, W. W. L., & Pawloski, J. L. (1989). Detection of rippled noise by an Atlantic Bottlenose dolphin. *J Acoust Soc Am, 83*, 662-668.

Au, W. W. L., & Pawloski, D. A.. (1989). Cylinder wall thickness discrimination by an echolocating dolphin. Journal of Comparative Physiology. A 172, 41-47.

Au, W. W. L. (1980). Echolocation signals of the Atlantic bottlenose dolphin (*Tursiops truncatus*) in open waters. In R. G. Busnel & J. F. Fish (Eds.), *Animal sonar systems* (pp. 251-282). New York: Plenum Press.

Au, W. W. L. (1993). *The sonar of dolphins*. New York: Springer.

Azzali, M., Manzini, A., & Buracchi, G. (1995). Acoustic recognition by a dolphin of shapes. In R. A. Kastelein & J. A. Thomas & P. E. Nachtigall (Eds.), *Sensory systems of aquatic mammals* (pp. 137-156). Woerden, The Netherlands: De Spil.

Berg B. G. (1996). On the relationship between comodulation masking release and temporal modulation transfer functions. J. Acoust. Soc. Am. 100,1013-1023

Branstetter, B. K. (1999). *Envelope detection model for cylinder wall thickness discrimination by an echolocating dolphin (Tursiops truncatus).* Paper presented at the 114th Biennial Conference on the Biology of Marine Mammals, Maui, HI.

Branstetter, B. K., Mevissen, S. J., Herman, L. M., Pack, A. A., & Roberts, S. P. (2003). Horizontal angular discrimination by an echolocating bottlenose dolphin *Tursiops truncatus*. *Bioacoustics, 14,* 15-34.

Brill, R. L., Moore, P. W., & Dankiewicz, L. A. (2001). Assessment of dolphin (*Tursiops truncatus*) auditory sensitivity and hearing loss using jaw phones. *J Acoust Soc Am, 109*(4), 1717-1722.

Brill, R. L., Moore, P. W. B., & Helweg, D. A. (2001). *Investigating the dolphin's peripheral hearing system: Acoustic sensitivity about the head and lower jaw.* San Diego: SSC San Deigo.

Butler, R. A., Humanski, R. A., & Musicant, A. D. (1990). Binaural and monaural localization of sound in two-dimensional space. *Perception, 19*(2), 241-256.

Cranford T.W. (2000). In search of impulse sound sources in odontocetes in: W.W.L Au, A.N. Popper & R.R. Fay (eds): Hearing by whales and dolphins, pp.109-155 Springer-Verlag, New York, Berlin

Evans, W. E. (1973). Echolocation by marine delphinids and one species of fresh water dolphin. J. Acoust. Soc. Am. 54, 191-199.

Evans, W. E., Sutherland, W., & Beil, R. (1964). The directional characteristics of delphinid sounds. In W. N. Tavolga (Ed.), *Marine bioacoustics, vol 2* (pp. 353-372). New York: Pergamon Press.

Fisher, F. H., and Simmons, V. P. (1973). Sound absorption in sea water. J. Acoust. Soc. Am., 62. 558-564.

Fletcher, H. (1940). Auditory patterns. *Review of Modern Physics, 12*, 67-72.

Glasberg, B. R., & Moore, B. C. (1990). Derivation of auditory filter shapes from notched-noise data. *Hear Res, 47*(1-2), 103-138.

Harley, H. E., Putman, E. A., & Roitblat, H. L. (2003). Bottlenose dolphins perceive object features through echolocation. *Nature, 424*(6949), 667-669.

Harley, H. E., Roitblat, H. L., & Nachtigall, P. E. (1996). Object representation in the bottlenose dolphin (*Tursiops truncatus*): integration of visual and echoic information. *J Exp Psychol Anim Behav Process, 22*(2), 164-174.

Hemila, S., Nummela, S., & Reuter, T. (1999). A model of the odontocete middle ear. *Hear Res, 133*(1-2), 82-97.

Hemila, S., Nummela, S., & Reuter, T. (2001). Modeling whale audiograms: effects of bone mass on high-frequency hearing. *Hear Res, 151*(1-2), 221-226.

Herman, L. M., & Arbeit, W. R. (1972). Frequency difference limens in the bottlenose dolphin: 1-70 kHz. *Journal of Auditory Research, 12*, 109-120.

Herman, L. M., Pack, A. A., & Hoffmann-Kuhnt, M. (1998). Seeing through sound: dolphins (*Tursiops truncatus*) perceive the spatial structure of objects through echolocation. *J Comp Psychol, 112*(3), 292-305.

Ketten, D. R. (2000). Cetacean ears. In W. W. L. Au & A. N. Popper & R. R. Fay (Eds.), *Hearing by whales and dolphins* (pp. 43-108). New York: Springer.

Kistler, D. J., & Wightman, F. L. (1992). A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *J Acoust Soc Am, 91*(3), 1637-1647.

Knudsen, E. I. (1983). Early auditory experience aligns the auditory map of space in the optic tectum of the barn owl. *Science, 222*(4626), 939-942.

Knudsen, E. I. (1984). Auditory properties of space-tuned units in owl's optic tectum. *J Neurophysiol, 52*(4), 709-723.

Knudsen, E. I., & Konishi, M. (1978). A neural map of auditory space in the barn owl. *Science, 200*, 759-797.

Knudsen, E. I., & Konishi, M. (1979). Mechanisms of sound localization in the barn owl (*Tyto alba*). *Journal of Comparative Physiology, 133*, 13-21.

Lammers, M. O., & Au, W. W. L. (2003). Directionality in the whistles of Hawaiian spinner dolphins (*Stenella longirostris*): A signal feature to cue direction of movement? *Marine Mammal Science, 19*, 249-264.

Lemmons, D. W., (1999). Auditory filter shapes in an Atlantic bottlenose dolphin (Tursiops truncatus). Unpublished doctoral dissertation. University of Hawaii.

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

Mercado, E., III. (1998). Humpback Whale Bioacoustics. Unpublished doctoral dissertation. University of Hawaii

Merzenich, M. M., & Schreiner, C. E. (1992). Mammalian auditory cortex - some comparative observations. In D. B. Webster & R. R. Fay & A. N. Popper (Eds.), *The evolutionary biology of hearing* (pp. 673-689). New York: Sprinfer-Verlag.

Middlebrooks, J. C., & Green, D. M. (1992). Observations on a principal components analysis of head-related transfer functions. *J Acoust Soc Am, 92*(1), 597-599.

Moore, P. W., Hall, R. W., Friedl, W. A., & Nachtigall, P. E. (1984). The critical interval in dolphin echolocation: what is it? *J Acoust Soc Am, 76*(1), 314-317.

Moore, P. W. B., Pawloski, D. A., & Dankiewicz, L. A. (1995). Interaural time and intensity difference thresholds in the bottlenose dolphin (*Tursiops truncatus*). In R. A. Kastelein & J. A. Thomas & P. E. Nachtigall (Eds.), *Sensory systems of aquatic mammals* (pp. 11-23). Woerden, Netherlands: De Spil.

Murchison, A. E. (1980). Detection range and range resolution of echolocating bottlenose porpoise (*Tursiops truncatus*). In R. G. Busnel & J. F. Fish (Eds.), *Animal sonar systems* (pp. 43-70). New York: Plenum Press.

Nachtigall, P. E. (1980). Odontocete echolocation performance on object size, shape, and material. In R. G. Busnel & J. F. Fish (Eds.), *Animal sonar systems* (pp. 71-95). New York: Plenum Press.

Norris, K. S., & Harvey, G. W. (1974). Sound transmission in the porpoise head. *J Acoust Soc Am, 56*(2), 659-664.

Pack, A. A., & Herman, L. M. (1995). Sensory integration in the bottlenosed dolphin: immediate recognition of complex shapes across the senses of echolocation and vision. *J Acoust Soc Am, 98*(2 Pt 1), 722-733.

Pack, A. A., Herman, L. M., & Hoffmann-Kuhnt, M. (2004). Dolphin echolocation shape perception: from sound to object. In J. A. Thomas & C. F. Moss & M. Vater (Eds.), *Echolocation in bats and dolphins* (pp. 288-298). Chicago: University of Chicago Press.

Pack, A. A., Herman, L. M., Hoffmann-Kuhnt, M., & Branstetter, B. K. (2002). The object behind the echo: dolphins (*Tursiops truncatus*) perceive object shape globally through echolocation. *Behav Processes, 58*(1-2), 1-26.

Renaud, D. L., & Popper, A. N. (1975). Sound localization by the bottlenose porpoise *Tursiops truncatus*. *J Exp Biol, 63*(3), 569-585.

Roitblat, H. L., Ketten, D. R., Au, W. W. L., & Nachtigall, P. E. (1996). A computational model of early stages of dolphin hearing. *Journal of the Acoustical Society of America, 100*, 2643.

Roitblat, H. L., Moore, P. W. B., Helweg, D. A., & Nachtigall, P. E. (1993). Representation and processing of acoustic information in a biomimetic neural network. In J. A. Meyer & S. W. Wilson & H. L. Roitblat (Eds.), *From animals to* *animats 2: Simulations of adaptive behavior* (pp. 90-99). Cambridge, MA: MIT Press.

Simmons, J. A., Kick, S. A., Lawrence, B. D., Hale, C., Bard, C., & Escudie, B. (1983). Acuity of horizontal angular discrimination by the echolocating bat, *Eptesicus fuscus*. *Journal of Comparative Physiology*(153), 321-330.

Slaney, M. (1998). *Auditory toolbox (ver. 2.0)*: Apple Computer Technical Report #45.

Slaney, M. (**1993**). "An efficient implementation of the Patterson-Holdsworth auditory filter bank." Apple Computer Technical Report #35. Perception Group-Advanced Technology Group.

Southworth, C. & Berg, B. G. (1995). Multiple cues for the discrimination of narrow-band sounds. J. Acoust. Soc. Am., 98, 2486-2492.

Stickland, E. A., Viemeister, N. F. (1997). The effects of frequency region and bandwidth on the temporal modulation transfer function. J. Acoust. Soc. Am., 102, 1799-1810.

Supin, A., & Popov, V. V. (1993). Direction-dependent spectral sensitivity and interaural spectral difference in a dolphin: evoked potential study. *J Acoust Soc Am, 93*(6), 3490-3495.

Varanasi, U., and Malin, D. C. (1971). Unique lipids of the porpoise (Tursiops gilli): differences in tricylglycerols and wax esters of acoustic (mandibular canal en melon) and blubber tissues. Biochem. Biophys. Acta. 231, 415-418.

Viemeister, N. F. **(1979)**. "Temporal modulation transfer functions based on modulation thresholds," ," J. Acoust. Soc. Am. **66**, 1364-1380.

Viemeister, N. F., and Plack, C. J. **(1993)**. "Time Analysis," In: *Human Psychophysics*, edited by W. A. Yost, A. N. Popper, and R. R., Fay (Springer-Verlag, New York), pp. 116-192.

Wightman, F. L., & Kistler, D. (1995). Sound localization. In W. A. Yost & A. N. Popper & R. R. Fay (Eds.), *Human psychophysics*. New York: Springer-Verlag.

Yost, W. A. and Sheft, S. **(1989)**. "Across-critical-band processing of amplitude-modulated tones," J. Acoust. Soc. Am. **85**, 848-857.

Zwicker, E., and Fastl, H. **(1990)**. Psychoacoustics. Facts and Models (Springer-Verlag. Berlin).