

UNCLASSIFIED

AD 441146

DEFENSE DOCUMENTATION CENTER

SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION ALEXANDRIA VIRGINIA



UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

441146

NOTS TP 3109
Part 4
COPY 45

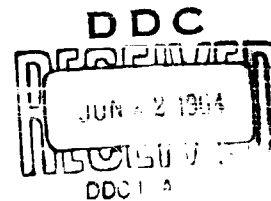
LOCALIZATION OF SOUND

Part 4. Further Developments in Sound Localization

FINAL REPORT

by

United Research, Inc.
Cambridge, Massachusetts



ABSTRACT: The theoretical considerations of the localization of sound function utilizing time delays have been extended to other areas of acoustics to provide a comprehensive theory useful in explaining speech recognition, attention, and the importance of reverberation on speech intelligibility, acoustic coloration, noise masking, and other phenomena.

The knowledge of speech recognition was used in initial work to establish inter-species communication with the porpoise.

Released to ANIA for further dissemination with out limitation beyond those imposed by security regulations.



U. S. NAVAL ORDNANCE TEST STATION

China Lake, California

April 1964

441146

U. S. NAVAL ORDNANCE TEST STATION

AN ACTIVITY OF THE BUREAU OF NAVAL WEAPONS

C. BLENNAN, JR., CAPT., USN
Commander

WM. B. MCLEAN, PH.D.
Technical Director

FOREWORD

The purpose of this report is to describe results of studies in sound localization carried out over a 3-year period. During this time the role of the external ear in sound localization was established and devices were constructed to provide means for localizing sounds in other environments.

This work was conducted by United Research, Inc., Cambridge, Mass., under Contract No. N123(60530) 32279A issued by the U.S. Naval Ordnance Test Station. The information herein covers work that completed this contract in December 1963.

This is Part 4 (Final Report) of a series of reports that were issued covering various aspects of the subject. The titles of Parts 1, 2, and 3 are as follows:

Part 1. Characteristics of Human Localization of Sound (December 1962)

Part 2. The Mechanism of Human Localization of Sounds With Applications in Remote Environments (December 1962)

Part 3. A New Theory of Human Audition (December 1963)

The text pages in this report were reproduced in facsimile from a report issued by the contractor.

WM. B. MCLEAN
Technical Director

NOTS Technical Publication 3109, Part 4

Published by Publishing Division
Technical Information Department
Collation Cover, 45 leaves, abstract cards
First printing 130 numbered copies
Security classification UNCLASSIFIED

- 1 Applied Physics Laboratory, University of Washington, Seattle
- 1 Bell Telephone Laboratories, Inc., Whippany Laboratory, Whippany, N. J.
- 1 California Institute of Technology, Pasadena (Dr. M. S. Plesset)
- 1 Cleveite Ordnance, Cleveland (Security Officer)
- 1 Davidson Laboratory, Stevens, Institute of Technology, Hoboken, N. J.
- 1 General Electric Company, Defense Electronics Division, Pittsfield, Mass. (K. I. Iglar)
- 1 General Electric Company, Johnson City, N. Y. (Light Military Electronics Department, Armament & Control Products Section, Library)
- 1 General Dynamics/Astronautics, Space Science Laboratory, San Diego (Mail Zone 596-30, S. Kaye)
- 1 Georgia Institute of Technology, Atlanta (Chief of the Physical Sciences Division)
- 1 Honeywell California Ordnance Center, West Covina, Calif.
- 1 Hudson Laboratories, Columbia University, Dobbs Ferry, N. Y.
- 1 Lincoln Laboratory, MIT, Lexington
- 1 Lockheed Aircraft Corporation, Missiles and Space Division, Palo Alto, Calif.
- 1 Ordnance Research Laboratory, Pennsylvania State University (Development Contract Administrator)
- 1 Philco Corporation, Philadelphia, via RInsMat
- 1 Servomechanisms Laboratory, MIT, Cambridge
- 1 The Rand Corporation, Santa Monica, Calif. (Astronautics Department)
- 1 United Research Inc., Cambridge, Mass.
- 1 Vitro Corporation of America, Silver Spring
- 1 University of California, San Diego, Scripps Institution of Oceanography, Marine Physical Laboratory, San Diego
- 1 Westinghouse Electric Corporation, Baltimore (Engineering Librarian)
- 1 Westinghouse Electric Corporation, Sunnyvale, Calif.
- 1 Westinghouse Research Laboratories, Pittsburgh (Arthur Nelkin)
- 1 Woods Hole Oceanographic Institution, Woods Hole, Mass.

CONTENTS

Chapter 1. Summary	1
2. Experimental Developments	3
3. Porpoise Translators	44
4. Theory	63
5. Conclusions	81
References	87

CHAPTER 1

SUMMARY

1.1 Introduction

The work reported here was done under U. S. Naval Ordnance Test Station Contract No. N123-(60530) 32279A to continue research in human localization of sound begun under earlier contracts and reported in references 1, 2 and 3.

Early attempts to reproduce the function of human localization by electronic means both in air and underwater disclosed the need for advances in the state-of-the-art in microphones, headphones, hydrophones, and a greater knowledge of the effects of mechanical structure between the ears. A study was undertaken to resolve these problem areas, which were divided into four prime tasks. (1) To study improvements in hydrophones for adaptation to underwater listening systems utilizing ear replicas; (2) to study the effects of mechanical structure between the ears on localization; (3) to study improvements in headphones for use in localization systems; (4) to continue the study of the transformation function of the pinna.

Two additional tasks were included to develop devices for use in man-to- porpoise and porpoise-to-man communication. The design concept for these translators was an outgrowth of localization research which provided new insight into the recognition of human speech.

1.2 Results

Positive results were not achieved in every task. Hydrophone improvements were not realized despite extensive efforts in this area.

Headphones, on the other hand, were made which accurately transferred the localization function. The study of the effects of structure showed the independence of interaural distance and the localization function. Significant progress was realized in the theoretical aspects of our study wherein consistent theorems for finding the inverse transformation have been formulated and the processing schematics drawn which closely parallel the known structure of neural networks

It has been established that the porpoise communicates in his environment by modulated whistles. Two translators were made, the first of which translates porpoise whistles to human-like sounds while the second translates in the other direction. Two-way communication was thus provided. In experiments at Point Mugu we succeeded in getting the porpoise to mimic certain sounds.

CHAPTER 2

EXPERIMENTAL DEVELOPMENTS

2.1 Underwater Tests

A series of tests were conducted with the underwater listening system shown in Figures 2.1-1, 2.1-2 and 2.1-3. The ears are cast from stainless steel and are 4.5 times normal size. The spherical covering is 3/8" thick Absonic A. Clevite CH-12 Oyster hydrophones were used in this system. Its response is shown in Figure 2.1-4.

The first series of tests were conducted at Morris Dam Test Facility and were designed to ascertain the navigational utility of aural coupling to water. The listening system was suspended 30 feet below a navigable barge. The subject was seated on the deck and covered so that no visual references were possible. A noise source shown in Figure 2.1-5 was lowered to various depths at distances up to 500 feet away. The clapper was struck by pulling on a line attached to it. The subject listening through the pickup indicated the direction the operator should proceed. With the sound source fixed, the barge was navigated to it every time using only aural information.

The experiment was conducted as follows. The sound source was moved to an unknown location by outboard motor boat and lowered. The subject was isolated during this time. When the sound began, the subject, now listening through the underwater pickup, indicated the direction the barge operator was to go. As the barge turned the sound was perceived to move to the front. Straight ahead direction



Figure 2. 1-1 Oyster Hydrophone Positioned
in Mounting Flange

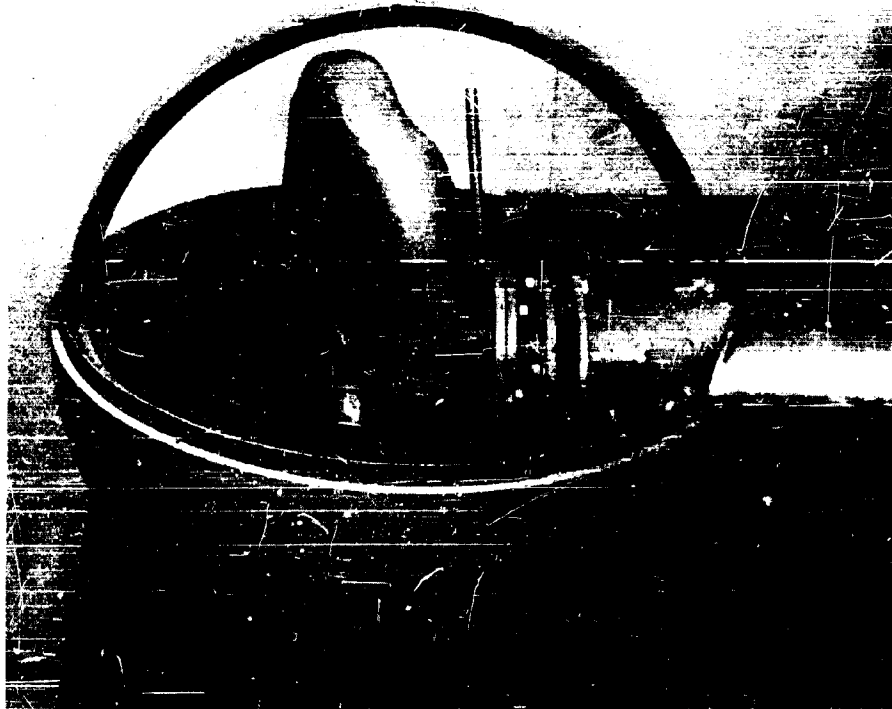


Figure 2.1-2 Enlarged Ear Replica in Underwater
Listening System

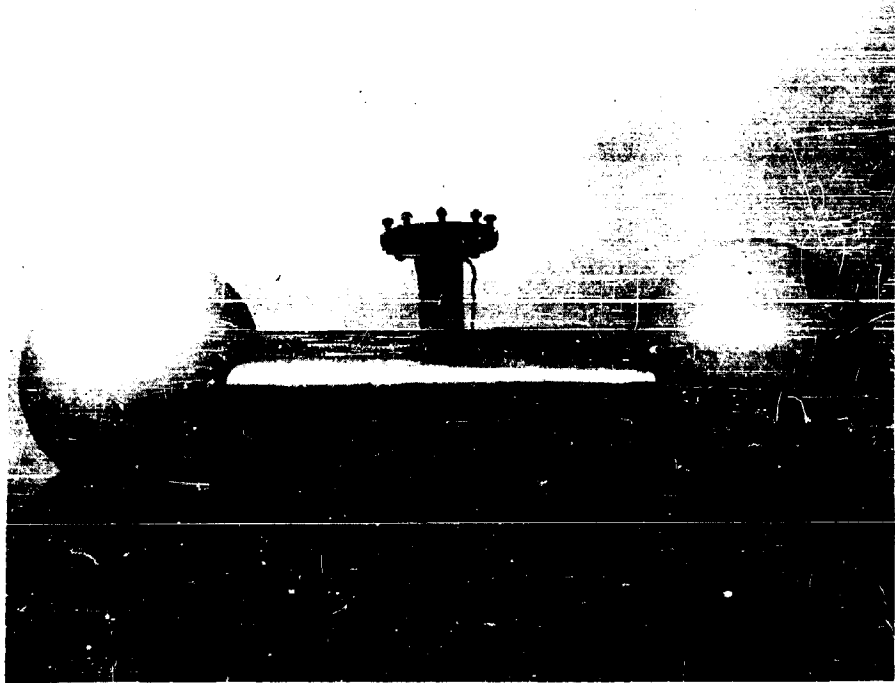


Figure 2.1-3 Experimental Underwater Listening System

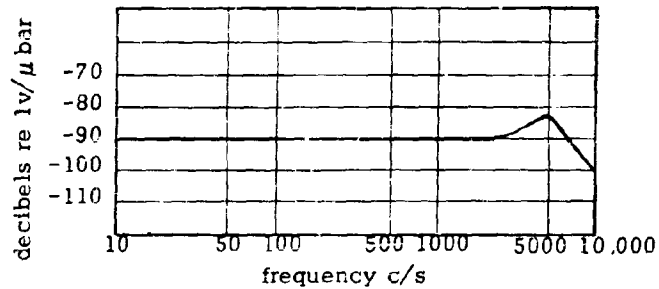


Figure 2.1-4 Response of Cyster

was then given and small corrections thereafter. It is significant to point out that no subject made gross errors in direction which he later corrected.

The second series of tests were conducted in January 1963 off San Clemente Island. These tests were designed to provide information or data on system performance in the ocean environment. The ear system was lowered over the stern of a Torpedo Recovery Boat to a depth of fifty feet. The first day was spent near shore where marine life noise was very evident, snaps, crackles, and pops apparently made by crustaceans. We also listened to small craft and noises made by divers hammering on buoys and on a barge bottom. Localization was equivalent to the expectation for the system.

Moving out to sea about two miles left the areas of crackling noise. Again, with the ears at 50 feet, the sea state (about one half) was audible, as was a sonar of about 9 kc, source unknown.



Figure 2.1-5 Sound Source Used in
Navigational Utility Tests

On the second day we set out into deep water about six miles from San Clemente. Enroute a school of migrating pilot whales was encountered and their whistles and ranging sounds recorded. Upon reaching deep water the listening system was again lowered to a depth of 50 feet. Sonar signals porpoise and whale sounds as well as a sound of arc welding were heard.

On the third day the listening system was installed in the porpoise pool at Pacific Ocean Park, Santa Monica. Unfortunately, the animals were very quiet this day.

Finally more tests were made in the Long Beach area. The ambient sea noise level was very high and although localization was satisfactory, range was very limited.

The results of these tests established the possibility and the utility of binaural coupling to the underwater environment. Effort was then directed toward improvements in the system components.

2.2 Hydrophones

The problem of selecting a hydrophone for adaptation to underwater listening systems prompted original efforts in hydrophone design to incorporate the features known to be necessary for adequate aural coupling. These requirements are (1) Frequency response in free field from 40 cycles per second to 70,000 cycles per second within ± 2 db; (2) detection threshold 5 db below sea state zero to 10,000 cycles; (this implies high element sensitivity, approximately -85 to -90 db re $1\text{v}/\mu\text{bar}$ and a correspondingly low preamplifier noise level, -140 to -145 db); (3) configuration and size so as to span the entrance of the ear canal; (4) capable of operation to 5000' depth.

In a survey of available commercial hydrophones no one unit was found to meet all these requirements. A compromise was made to use Clevite CH-12, Type O hydrophones in the initial underwater listening system primarily because of its physical size and cost. With allowance for limited hydrophone characteristics, the underwater listening system worked as expected, as reported above. However, improvements were desirable. To this end, a program to develop a hydrophone particularly for mounting behind an ear was undertaken. It was decided to use capacitance change techniques.

It seemed possible to design a hydrophone along the lines of a condenser microphone by using plastic dielectric material. In order to obtain an effective high potential while using a low supply voltage, a scheme of parallel charging, series discharging was designed. This is illustrated in Figure 2.2-1.

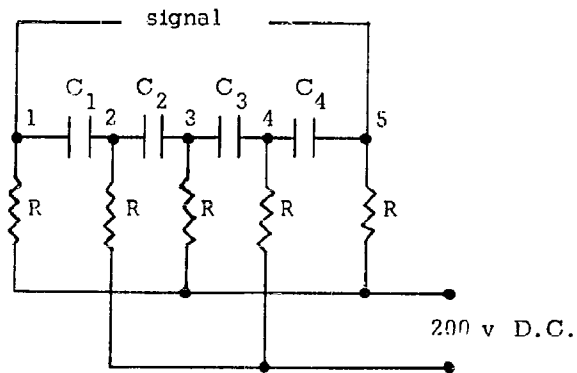


Figure 2.2-1 Parallel Charge — Series Discharge

The signal voltage, e_s is given

$$\begin{aligned} e_s &= e_{12} + e_{23} + e_{34} + e_{45} \\ &= q_{12} C_1^{-1} + q_{23} C_2^{-1} + q_{34} C_3^{-1} + q_{45} C_4^{-1} \end{aligned} \quad (2.2-1)$$

It is assumed that all q_{ij} are equal in magnitude and of alternate sign, so

$$e_s = q (C_1^{-1} - C_2^{-1} + C_3^{-1} - C_4^{-1}) \quad (2.2-2)$$

It is also assumed that static capacitance are all equal, so that the d.c. signal is zero. The change in signal is derived from changes in capacitance due to pressure

$$\frac{de_s}{dt} = q \sum_{i=1}^n (-1)^i \frac{d}{dt} (C_i^{-1}) \quad (2.2-3)$$

The basic capacitance is given by

$$C_o = \frac{kA}{d_o} \quad (2.2-4)$$

k = dielectric constant

A = area

d_o = static separation

Under pressure, the true spacing, d , is given by

$$d = d_o \left(1 - \frac{P}{\sigma} \right) \quad (2.2-5)$$

P = pressure

σ = modulus of elasticity

Thus the change in inverse capacitance with time can be written

$$\begin{aligned} \frac{dC^{-1}}{dt} &= \frac{d}{dt} \frac{d_o}{kA} \left(1 - \frac{P}{\sigma}\right) \\ &= \frac{d_o}{\sigma kA} \frac{dP}{dt} \end{aligned} \quad (2.2-6)$$

Choosing different σ materials for alternate capacitors permits us to write

$$\frac{de_s}{dt} = \frac{nq d_o}{2kA} \left(\frac{1}{\sigma_1} - \frac{1}{\sigma_2}\right) \frac{dP}{dt} \quad (2.2-7)$$

Thus gaining a larger signal voltage by increasing the number of elements.
Experimental materials chosen were mylar and polypropylene for which

$$\begin{aligned} \sigma_{\text{mylar}} &= 300 \times 10^3 \text{ psi} \\ \sigma_{\text{polypropylene}} &= 1.5 \times 10^3 \text{ psi} \end{aligned}$$

Equation (2.2-7) may be rewritten in terms of the polarization voltage:

$$\frac{de_s}{dt} = \frac{n}{2} V_o \left[\frac{1}{\sigma_1} - \frac{1}{\sigma_2}\right] \frac{dP}{dt} \quad (2.2-8)$$

since

$$V_o = \frac{q_o}{C_o} = \frac{q_o}{kA/d_o}$$

Sensitivity is defined as

$$S \triangleq \frac{de_s}{dP} \quad (2.2-9)$$

Hence

$$S = \frac{n}{2} V_o \left[\frac{1}{\sigma_1} - \frac{1}{\sigma_2}\right]$$

For a 100 element sensor and $V_o = 200$ volts

$$S = 6 \text{ volts/psi} = 100 \mu \text{ volts}/\mu \text{ bar} = -80 \text{ db re } 1 \text{ v}/\mu \text{ bar}$$

Certain difficulties are inherent in making this hydrophone. The charging resistance R in Figure 2.2-1 must be approximately 1000 megohm and be achievable in a film. Figure 2.2-2 shows a typical element and the techniques to be used in providing electrical continuity.

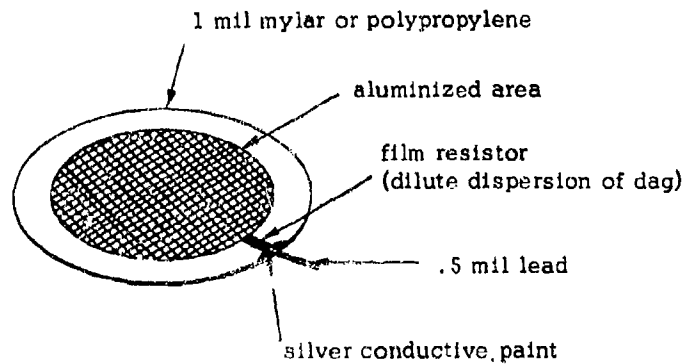


Figure 2.2-2 Typical Element

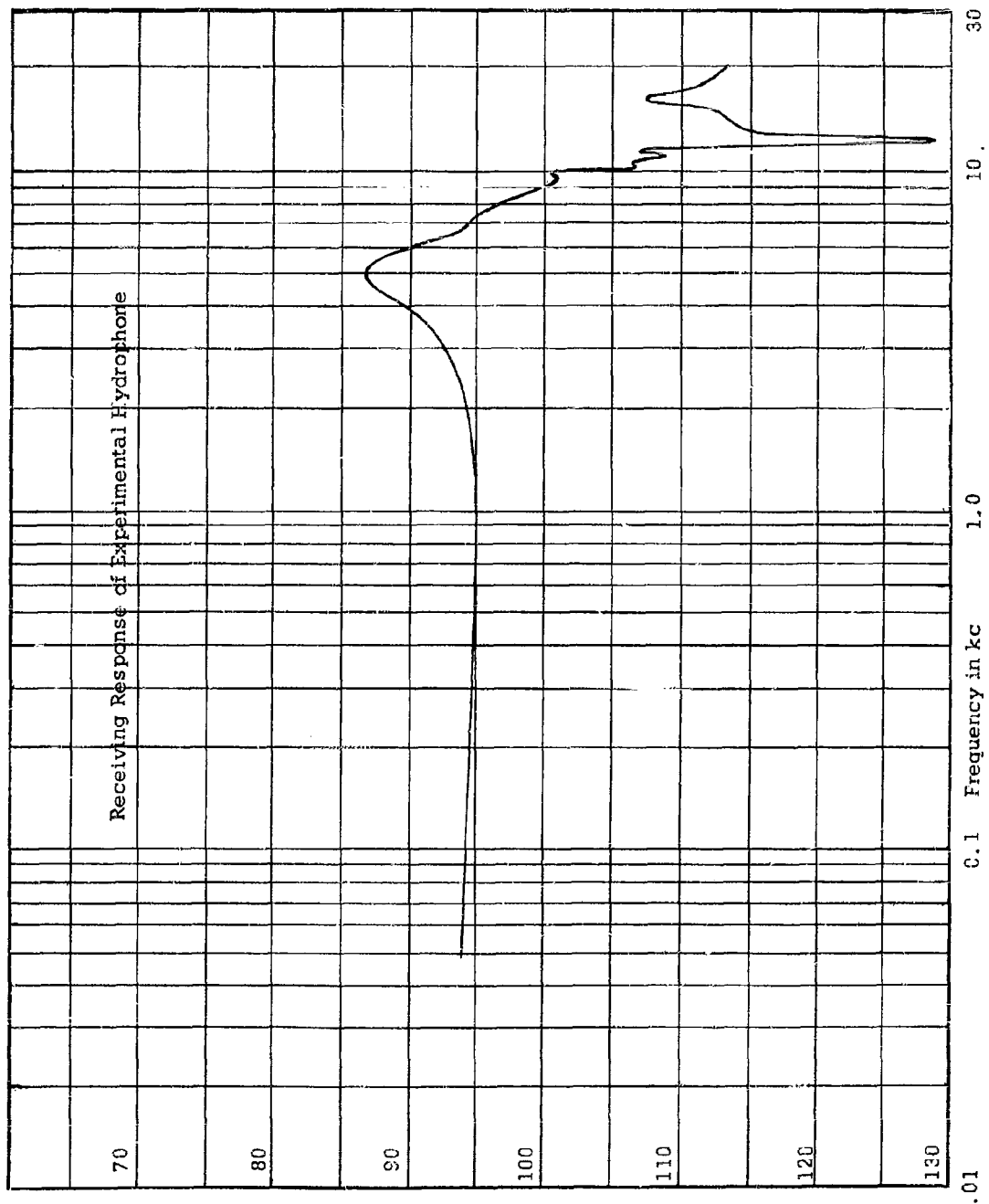
The complete sensor is made by alternately stacking mylar and polypropylene.

In spite of the promising features of this design, construction had to be abandoned in favor of a more conventional technique when details of fabrication and assembly proved too difficult to overcome.

An experimental hydrophone shown in Figure 2.2-3 was tested by U. S. Navy Electronics Laboratory. Figure 2.2-4 is the frequency response characteristic. The element used in this hydrophone was



Figure 2.2-3 Experimental Hydrophone



db below 1 volt/ μ bar
Figure 2.2-4

constructed as shown in Figure 2.2-5. From the response curve it was concluded that the mass of the fiber glass to a large extent limited the high frequency end. A second element was made as shown in Figure 2.2-6. It was installed in the steel flange to which the ear was fastened, Figure 2.2-7. Unfortunately, no adequate test data was acquired due to breakdown in test facilities. Arrangements were made to use facilities at U. S. Sonics Corporation, Cambridge but it was found that the techniques and equipment used there were inadequate for broad band testing required. A binaural listening system was nevertheless constructed using these hydrophones and tested in a nearby lake. The performance of these hydrophones was unsatisfactory due in part to the rapid decrease in sensitivity with depth. At fifty feet no audible signal was heard despite the presence of a scuba diver in the vicinity at the same depth knocking rocks together. Further development of condenser hydrophones was subsequently abandoned.

Later hydrophone work with the Massa TR-14A, the Clevite CH-13 and Chesapeake TP-120 revealed an anomaly in observations made in earlier work. It was found and reported earlier that the output of the broad band, secondary standard TP-120 hydrophone, when brought close to polyurethane foam, decrease appreciably. Similarly, when it was enclosed behind the ear canal opening, the behavior was the same. These effects were not observed with the CH-13 or the Massa TR-14A. A known difference between these hydrophones was the type of piezoelectric crystal used. The TP-120 utilizes lithium sulfate, sensitive to hydrostatic pressure fields while the others use barium titanate, sensitive to directed acoustic pressure. However, after repair and recalibration of the TP-120 necessitated by damage incurred in transit, it was found that it no longer behaved differently than the

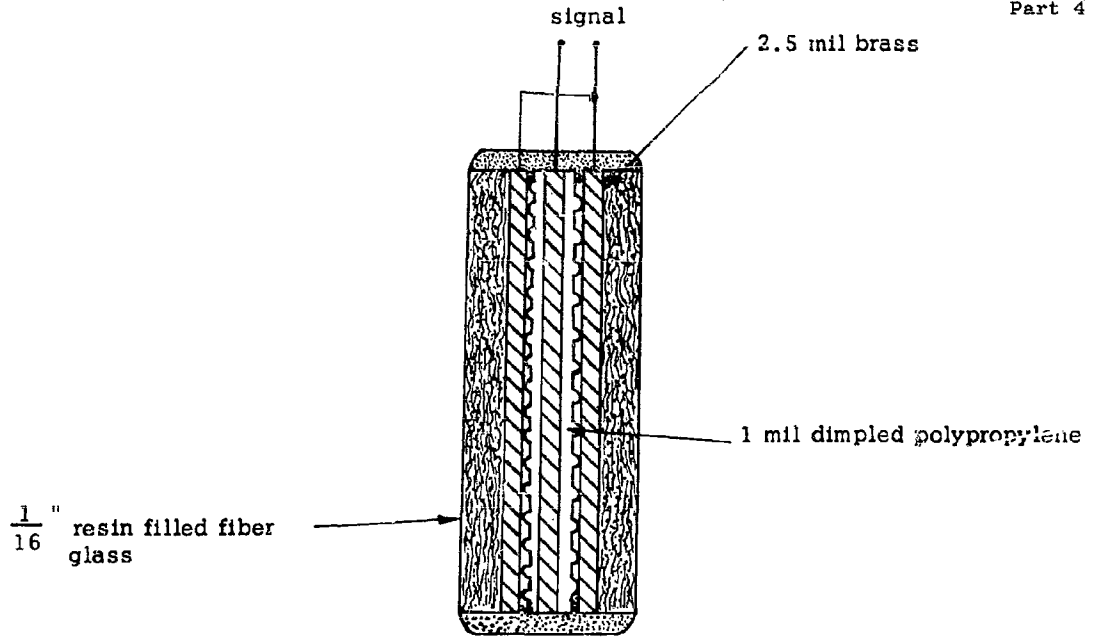


Figure 2.2-5 First Experimental Hydrophone Element

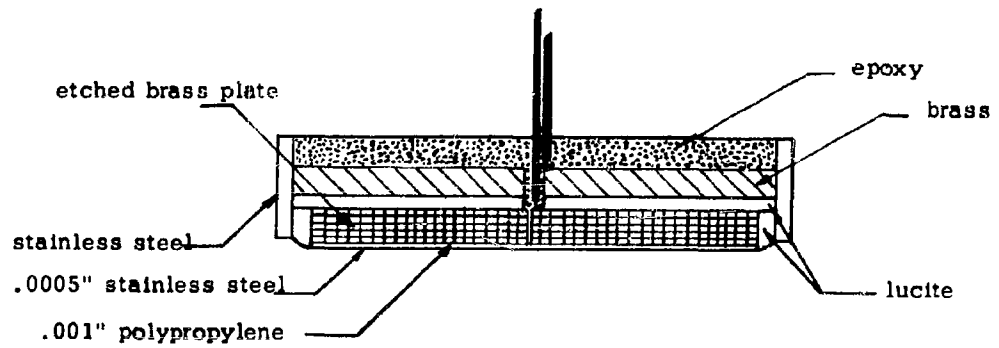


Figure 2.2-6 Can-Type Hydrophone Element

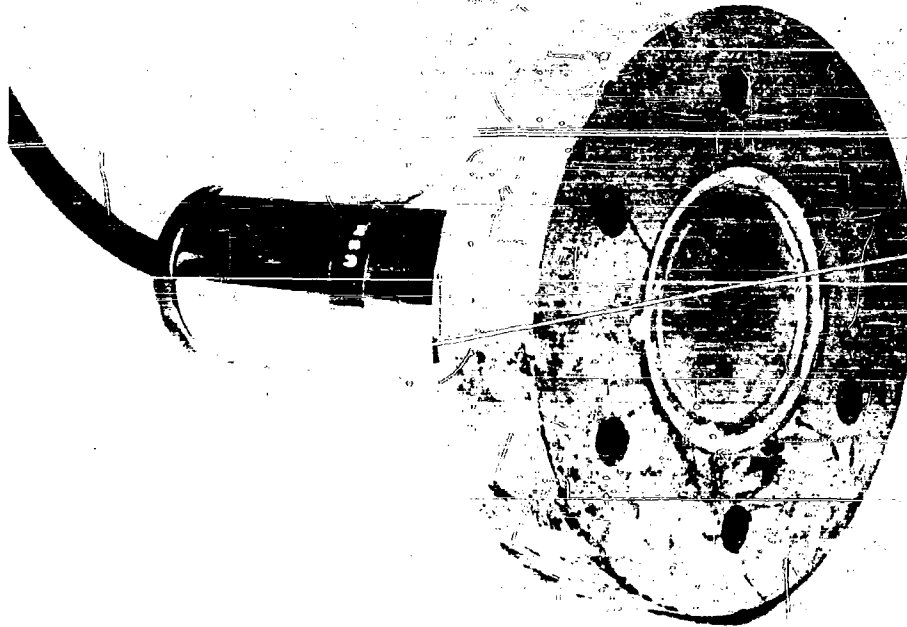


Figure 2.2-7 Capacitance Hydrophone Installed in
Mounting Flange

other hydrophones. In addition, it was further observed that the responses of all three hydrophones were now similar whether the ear behind which they were placed was made of polyurethane foam or solid stainless steel. This disconcerting behavior is as yet unexplained. However, it has emphatically indicated that the ease of adapting microphones to ears for air localization systems is not paralleled in the water environment. Despite these negative results, the data now on hand will allow greater progress to be made in providing effective aural coupling to water.

2.3 Structures

The diffraction and reflection effects of any structure within the interaural distance has frequently been assessed a high value of importance in localization. Although it has been stated that the mental function utilizes all possible information in localization, we were interested to know if some of these clues might be eliminated without impairing localization ability. To this end three binaural pickups were made as shown in Figures 2.3-1, 2.3-2 and 2.3-3. In order, John II maintains the features of a human head, although lacking hair-like material, the significance of which is unknown; Coco, a featureless head, and Peter, no head.

In each case the ears were separated by the equivalent acoustic distance, 10-3/4 inches. Subjective responses indicated that in all cases the ability to localize was neither impaired nor aided by the lack or presence of structure between the ears! This was a gratifying result and was pursued to certain extremes. The ears on Peter (no head at all!) were moved close together until separated by four inches then moved out until separated by 34 inches.



Figure 2.3-1 John II

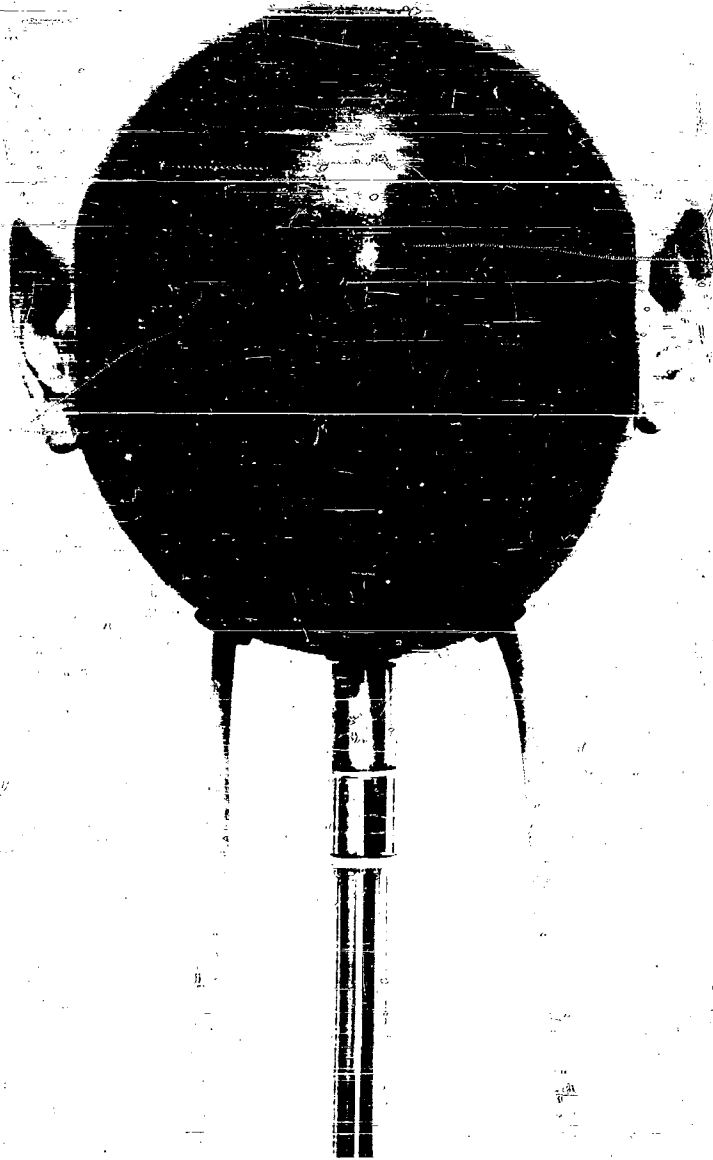


Figure 2.3-2 Coco

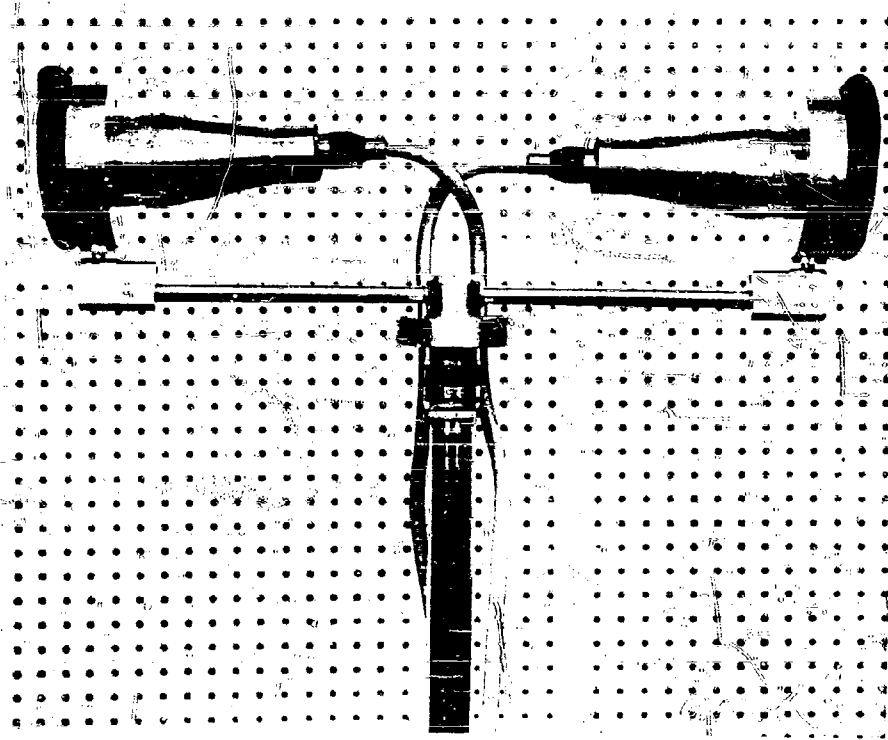


Figure 2.3-3 Peter

In both cases localization was unaffected if the sound source remained outside the spherical space whose diameter is defined by the interaural distance. Thus it appears that localization depends more on autocorrelation of the pinna transformed sound than on the cross correlation between ears, which has been found to be independent of interaural distance.

Effective aural coupling with binaural pickups has been improved by the awareness of the importance of acoustic isolation and by the choice of proper material out of which ears are cast.

Figure 2.3-4 is a section drawing of the ear-microphone assembly and the means used for acoustic isolation.

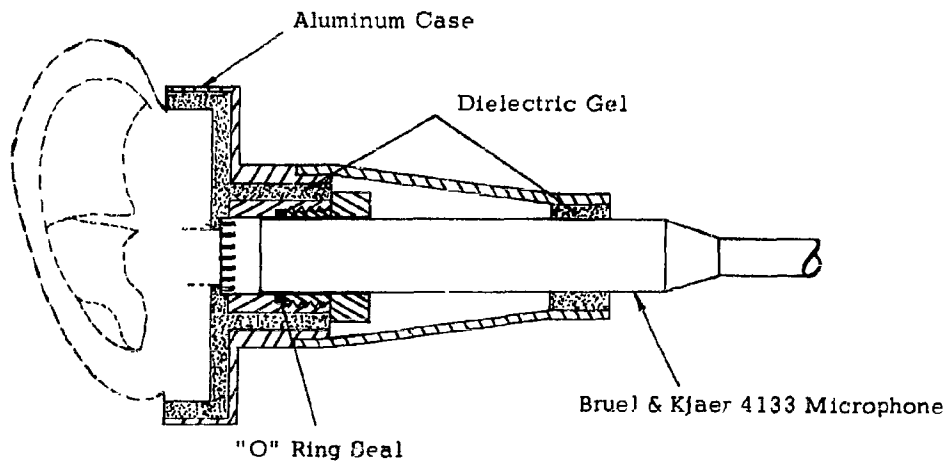


Figure 2.3-4 Ear-Microphone Assembly

From previous examination of the ultrasonic range 20 KC to 40 KC it appeared that the acoustic coloration of the material from which the pinnae are constructed could be of importance. This has been found to be true and the coloration of silicone rubber from which the first ear replicas were made is much too red for best use. Pinnae were cast from different epoxies and the coloration measured. Figures 2.3-5 and 2.3-6 illustrate the different colorations of each material and permits a comparison with a human ear, the data for which was measured by inserting a Bruel & Kjaer Model 4135 (1/4" diameter) microphone cartridge into the subject's ear canal and locating his ear in the same position as the plastic ears by means of a jig made for this purpose. From a comparison of the results, steel filled epoxy approaches a human ear best. More work is necessary in this area before the optimum material is found. It will be shown in a later section of this report that the transformation provided by the pinna may be expressed as the Laplace transform of the delay-dependent reflection coefficients. The significance of material selection is now known.

2.4 Headphones

We have long been aware of the limitations of available headphones for use in localization work. The requirements of flat frequency response over a broad bandwidth in an insertion type headphone cannot be met by any available product. The extent of the limitations were not known however. Therefore, the frequency response of a headphone-microphone system using the mannequin head, John II, as part of the equipment was determined. Various headphones were placed on the head and driven by an oscillator. The output of the microphone behind the ear was recorded. The results are shown in Figures 2.4-1 and 2.4-2. The poor headphone-microphone response substantiated the need for

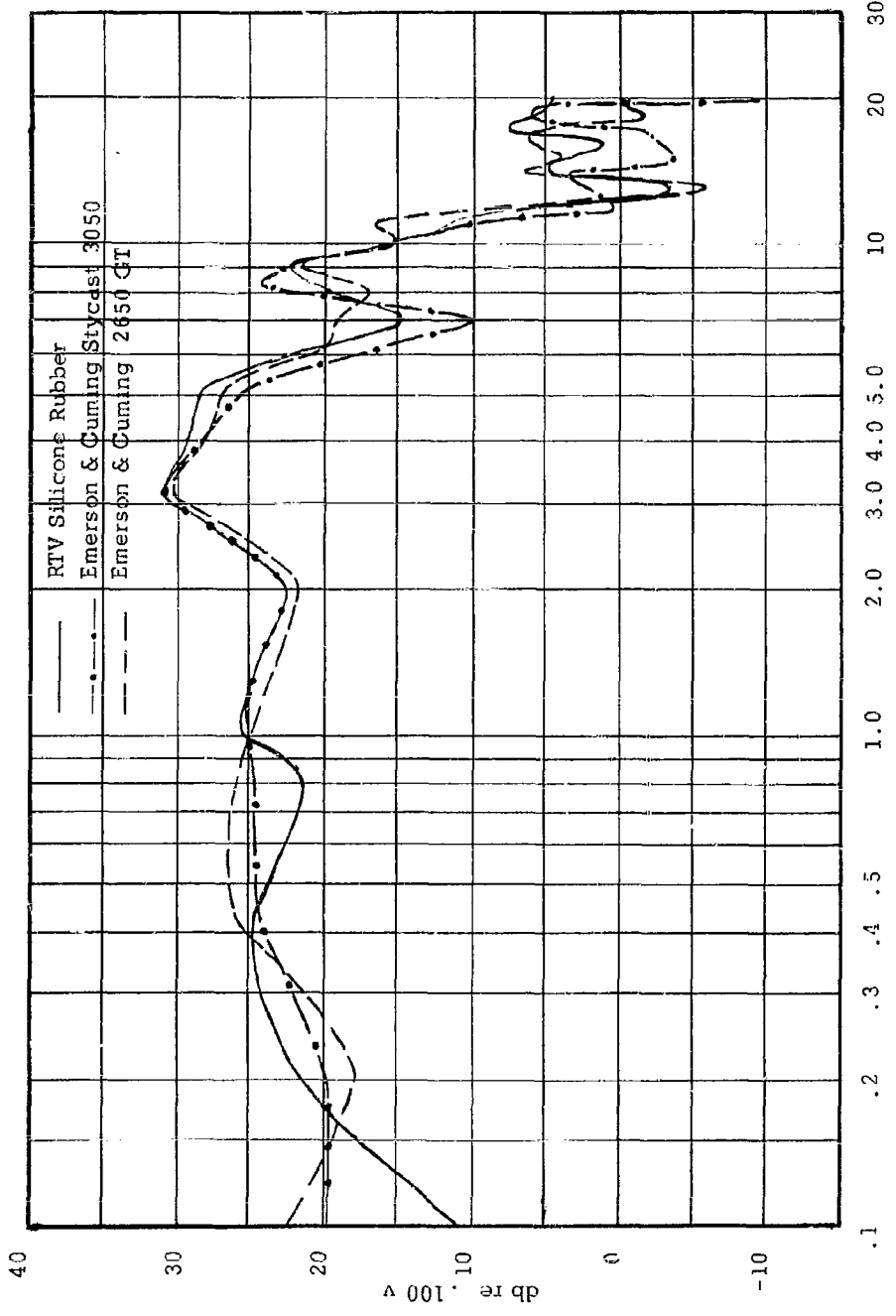


Figure 2.3-5 Acoustic Coloration

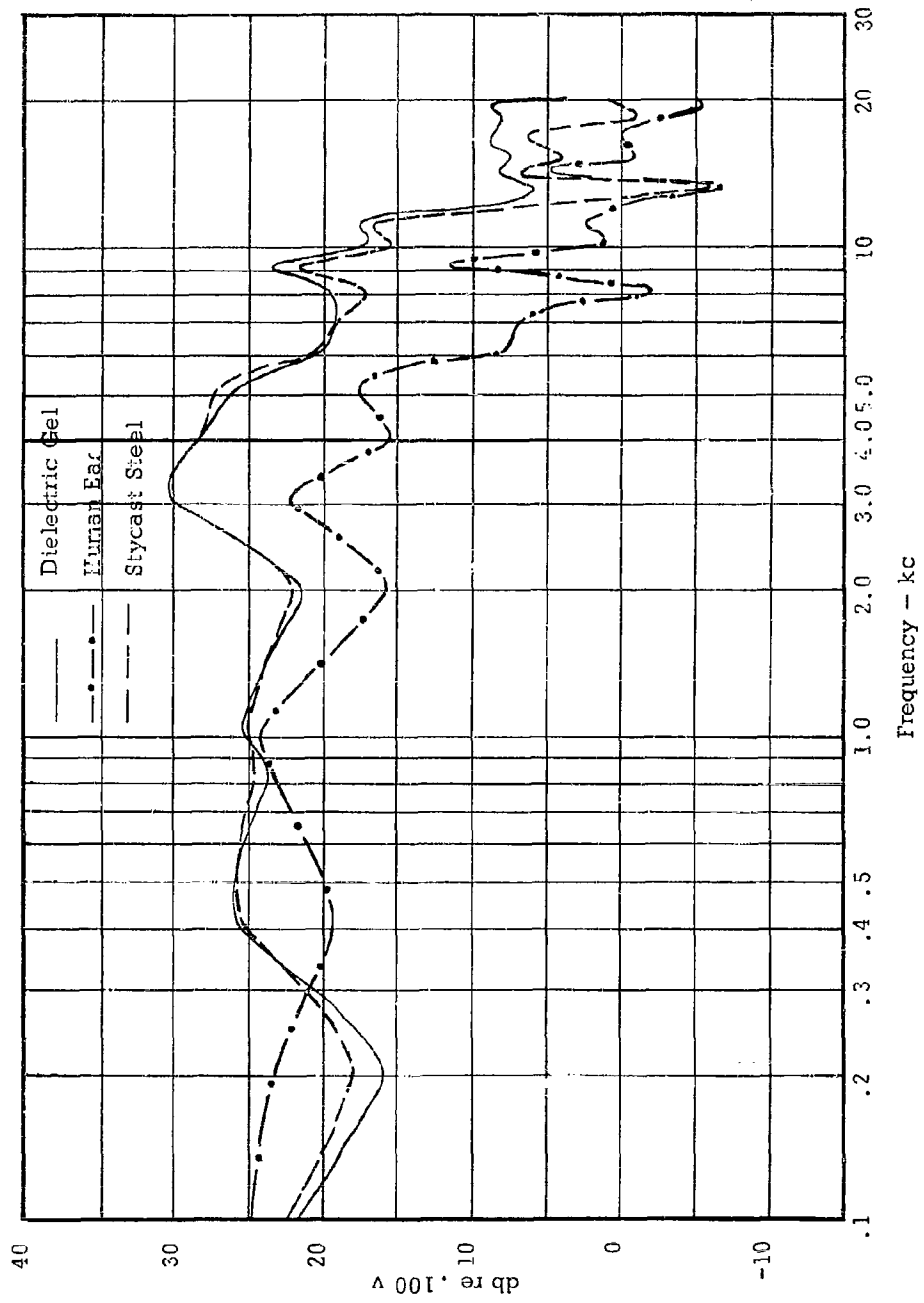


Figure 2.3-6 Acoustic Coloration

Beyer DT-508 semi-insertion headphone measured
on mannequin head equipped with AKG CK-26
cartridge and CK-60 follower

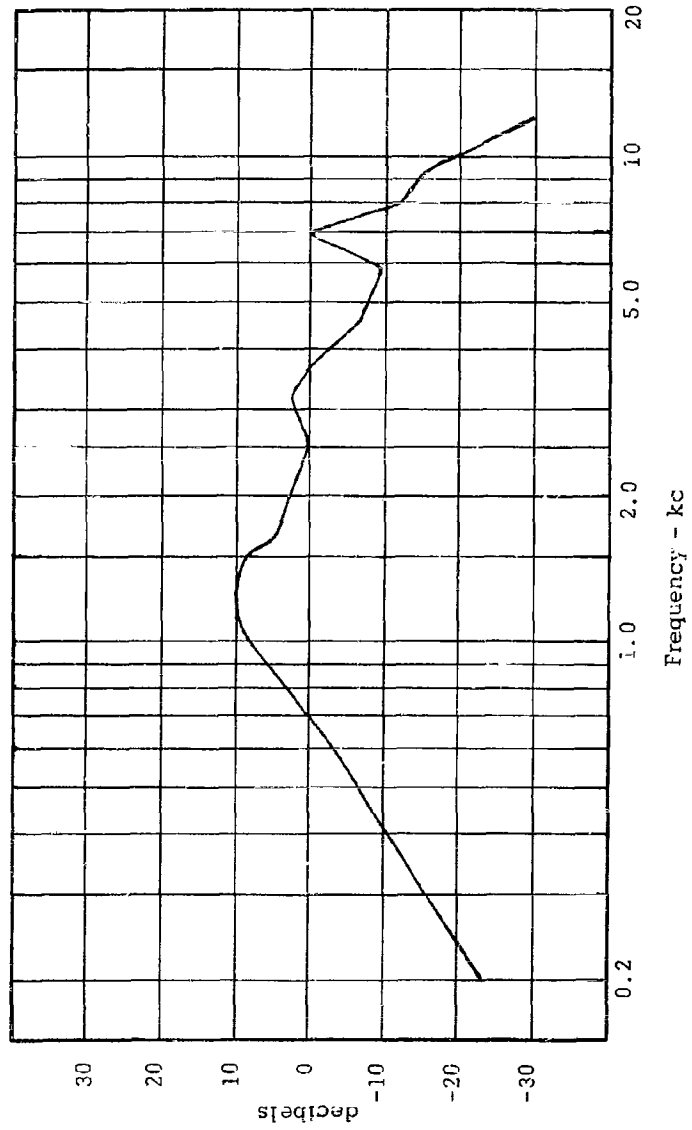


Figure 2.4-1 Insertion Headphone Response

AKG K-50 headphone response measured
on mannequin head equipped with
AKG CK-26 cartridge and CK-60 follower

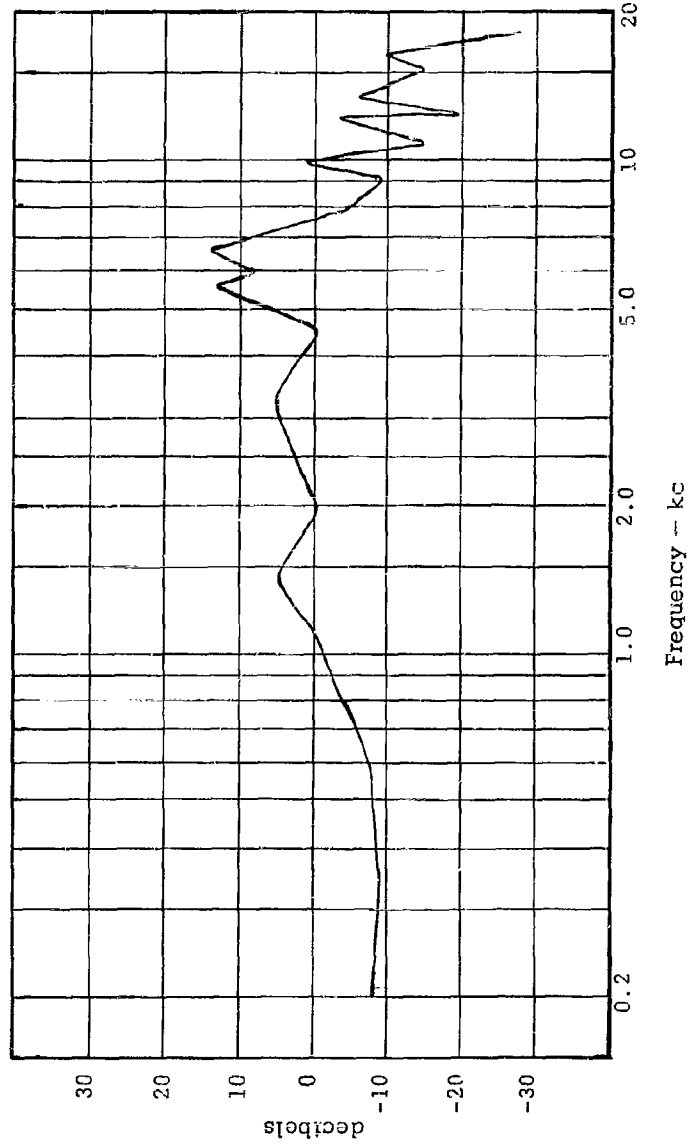


Figure 2.4-2 External Headphone Response

extensive work in this area if established requirements were to be fulfilled.

The block diagram of Figure 2.4-3 illustrates a single channel of a binaural localization system.

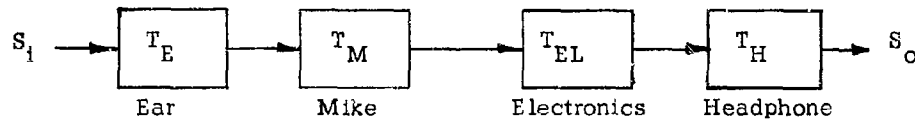


Figure 2.4-3 Single Channel of Localization System

The output signal S_o is the transformed input signal and may be written:

$$[T_E T_M T_{EL} T_H] S_i = S_o$$

- T_E - transformation function of the ear
- T_M - transformation function of microphone
- T_{EL} - transformation function of electronics
- T_H - transformation function of headphone

For localization

$$[T_E] S_i = S_o$$

that is, the transformation introduced by the pinna should be faithfully reproduced by the intervening components. Therefore

$$[T_M T_{EL} T_H] S_i = S_i$$

or

$$T_M T_{EL} T_H = 1 \quad \text{over the bandwidth from} \\ 40 - 40,000 \text{ cycles per second}$$

The fidelity of electronic equipment is readily achieved. Therefore T_{EL} can be made unity. However, the combined effects of the headphone and microphone create certain difficulties. Fortunately, the state of the art of microphone design and manufacture has progressed rapidly within the past few years. There are currently available microphones with a bandwidth extending from 20 cycles per second to 80,000 cycles per second within ± 3 manufactured by Bruel & Kjaer of Denmark. However, the state of the art in headphone development is not as far advanced. Headphones specified from 15 cycles per second to 15,000 cycles per second are found to have many sharp resonances within this bandwidth. These resonances effectively distort the delays introduced by the ear. Two major directions of headphone development were therefore pursued: (1) diaphragm magnetic and (2) capacitor.

The diaphragm magnetic headphone with the conductors integral with the diaphragm is shown in Figure 2.4-4. A typical experimental

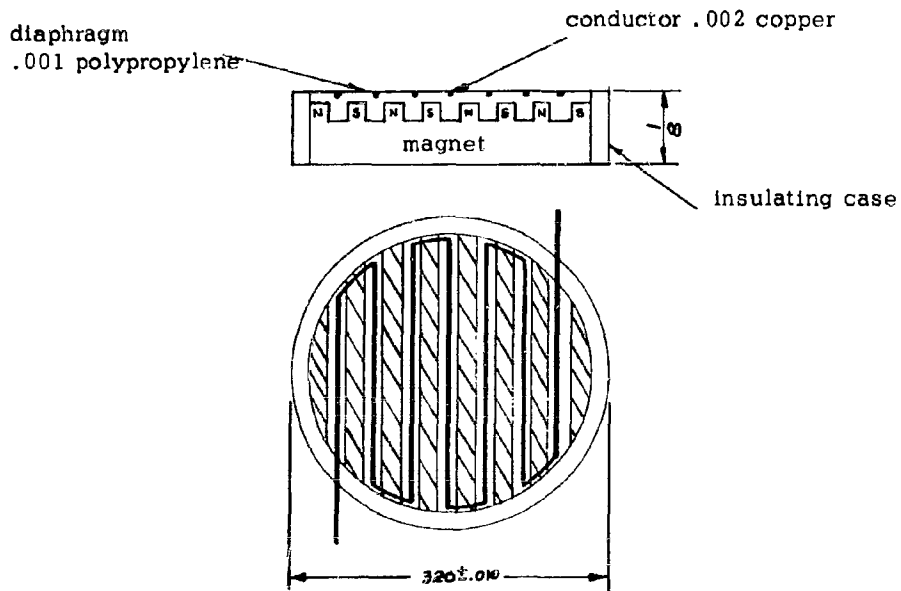


Figure 2.4-4 Magnetic Headphone

model had output sound pressure level of 85 db with 250 mw power input and a frequency response between 100 cps and 35 kcps measured against a Bruel & Kjaer 4133 microphone. DC impedance was 1.5 ohms. However, the 2 mil thick bakelite diaphragm made by Baldwin Lima Hamilton Co. showed many resonances between 5 kc and 25 kc. Figure 2.4-5 shows the test results.

The difficulties offered by this design included (1) choice of diaphragm material, (2) attachment of conductors to the diaphragm, (3) mounting of the diaphragm on the magnet to improve magnetic coupling and increase sensitivity. While none of these appears insurmountable, practically they are difficult to attain without affecting sensitivity frequency response or size.

Headphone elements made by NOTS personnel in which the conductors were printed on the diaphragm were tested and showed the same features as earlier models. Further development of this principle was terminated in favor of an insertion condenser element. Excepting for the low sensitivity of the magnetic headphone, it was eminently suitable to experimental needs.

The first condenser headphones employed Bruel & Kjaer Model 4135 1/4" diameter microphone cartridges. These had limited dynamic range and low sensitivity. However, the response characteristic was better than any headphone element tried and provided the first unequivocal transfer of the localization function by aural coupling. Since these cartridges are expensive and limited, as noted, a capacitor headphone using .15 mil aluminized mylar was developed which provides greater sensitivity and wider dynamic range at considerably less cost. Figure 2.4-6 shows a binaural condenser headset.

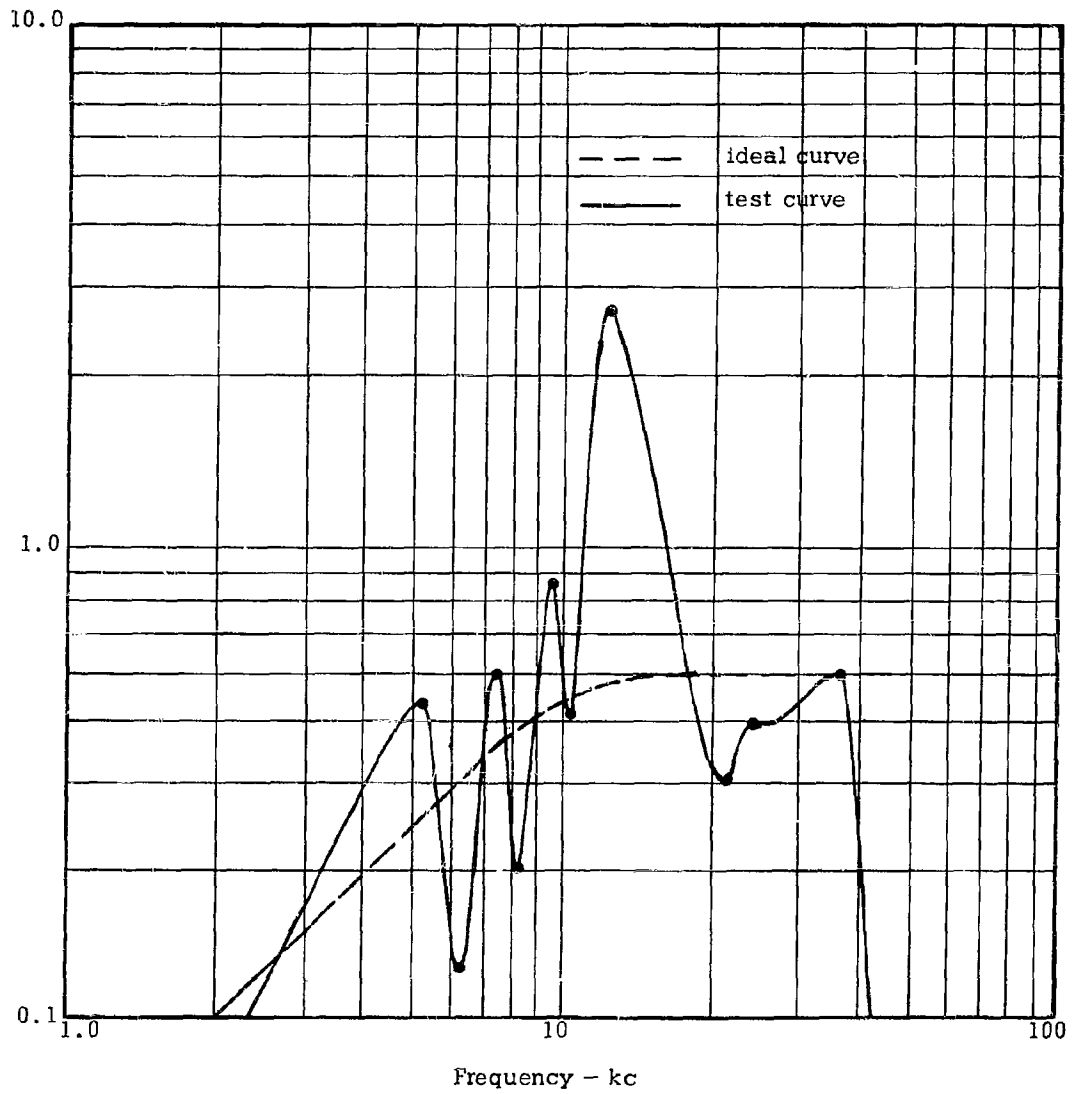


Figure 2.4-5 Experimental Results from BLH Diaphragm

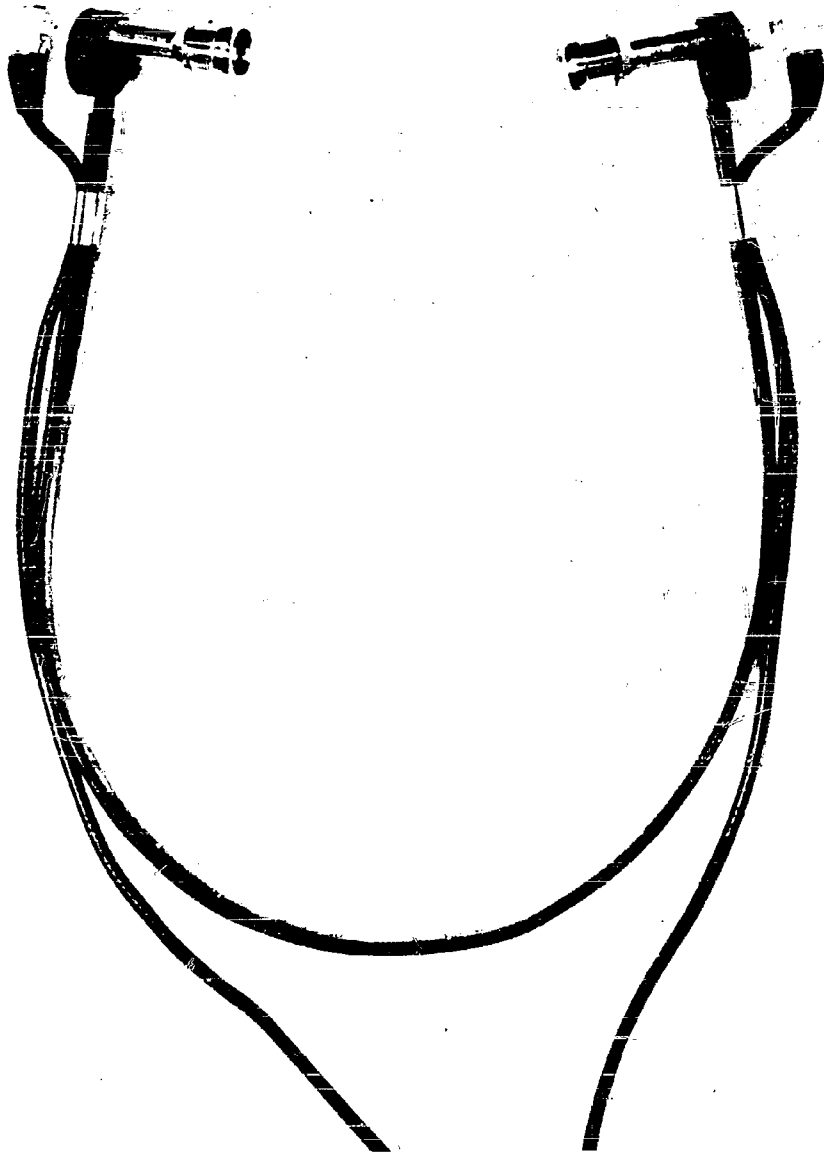


Figure 2. 4-6 Binaural Condenser Headset

In order to drive the condenser headset a high gain binaural amplifier, Figure 2.4-7, was designed to provide +300 volts DC polarization and a maximum undistorted signal of 84 volts RMS. The specifications for this unit are:

Gain:	54,000:1
Freq. Response:	12 cps - 74 kcps \pm 1 db
Noise:	Less than 1 μ volt RMS referred to input
Hum:	16 μ volts RMS referred to input
Max. Undistorted Output:	84 volts RMS
Harmonic Distortion:	40 volts RMS - 3% 2nd harmonic 20 volts RMS - 1.5% 2nd harmonic 10 volts RMS - 1% 2nd harmonic
Polarization Voltage:	+ 300 volts DC

Acoustic evaluation of headphones in closed cavities is complicated by cavity resonances for frequencies above about 14 kcps. As a result, a simple device was constructed to measure diaphragm displacement of a condenser headphone in order to infer the acoustic pressure response and the effects of loading as when coupled into the ear canal. Operation was based on varying the capacity of a circuit tuned to 21.4 mc. The output is a voltage linearly proportional to the displacement of the diaphragm over the range of interest. Figure 2.4-8 is the response curve of a condenser element found in this manner.

2.5 Evaluation of Aural Coupling

2.5.1 Localization Tests

Earlier work established the accuracy of human localization of sounds in a real environment, both binaurally and monaurally (Ref. 2).

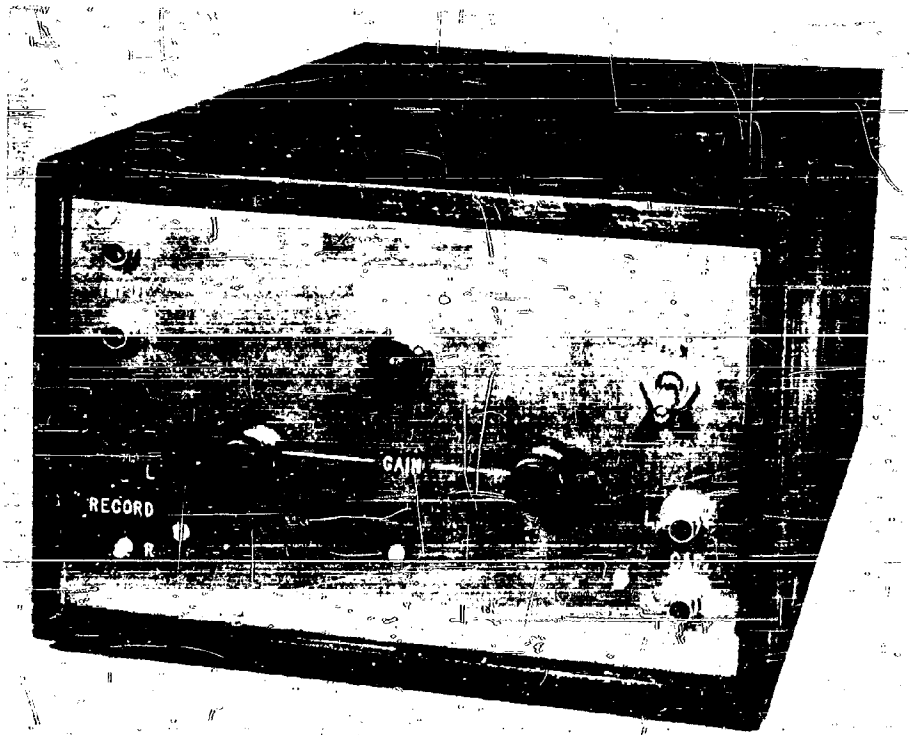


Figure 2.4-7 Condenser Headphone Driver

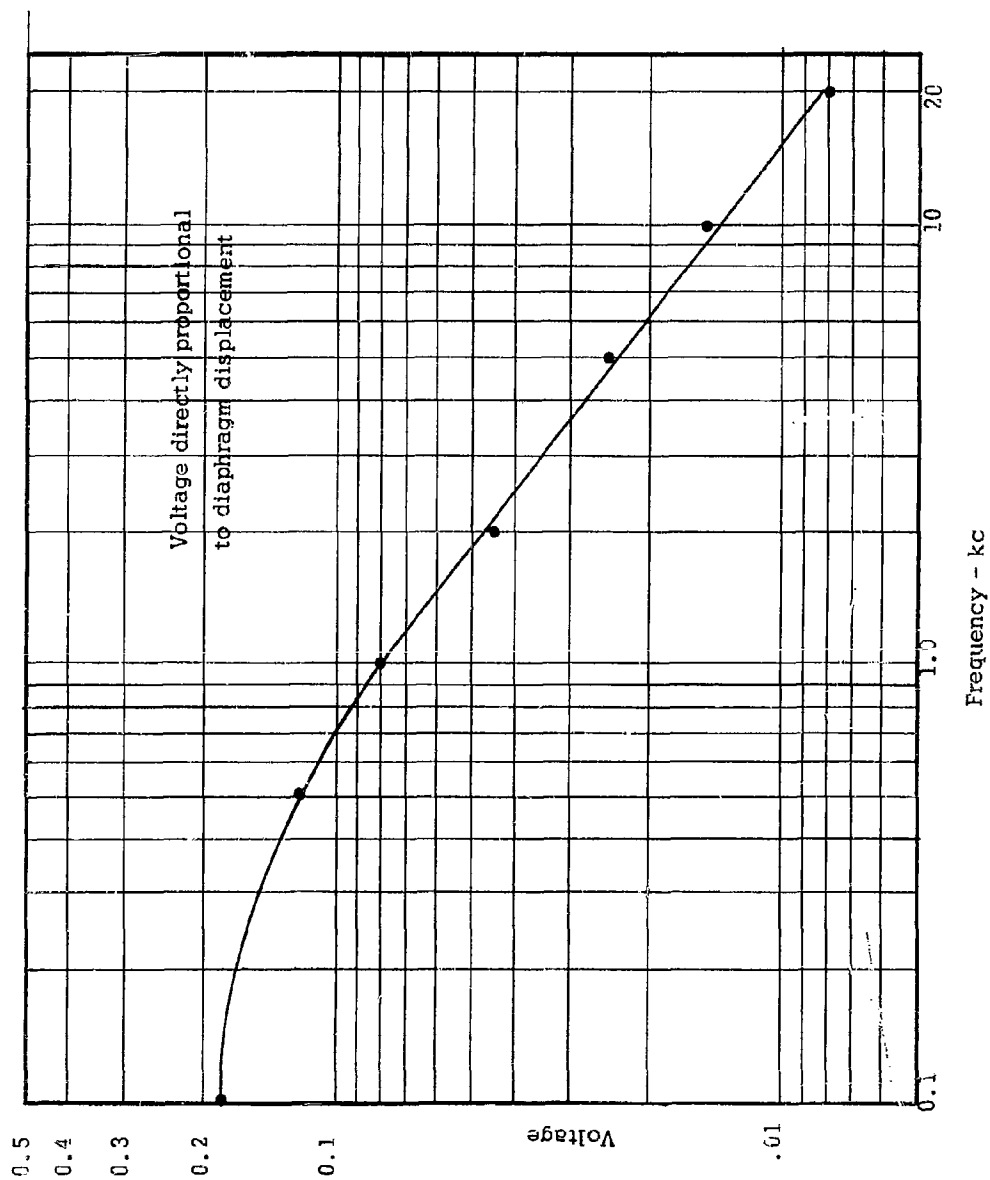


Figure 2.4-8 Condenser Headphone Diaphragm Response

In order to adequately evaluate the performance of any underwater sound localization system, it is necessary to know the accuracy achievable when a listener is coupled to another environment by means of the John II, or air system. From the standpoint of operational utility it is also desirable to know the learning time required before optimum accuracy is realized. A test was designed to determine these factors.

Eight 3" speakers were equally spaced on a 12-foot diameter ring and the mannequin head located at the center atop a tripod facing one of the speakers, Figure 2.5-1. The head was six feet above the floor with the speaker ring positioned at ear level. The subsequent tests involved azimuth only. A control box with eight switches was provided to permit selection of any one speaker by the testor. The subject who was in a separate room had a similar panel on which eight switches were placed corresponding to the speaker - ring layout, Figure 2.5-2. When a speaker was energized by the testor, the subject pressed the button corresponding to the position from which the sound appeared to originate. For a correct selection, a green light flashed on both the subject's panel and the testors panel. If the initial choice was incorrect, a red light flashed and another selection was made. This was repeated until the correct choice was made. The data recorded were:

1. position of speaker selected relative to head: no. 1 - front; no. 2 - 45° front right; no. 3 - right side
2. the number of incorrect choices made before the correct position was indicated
3. the total time between selection of the speaker and the subject's correct answer. The time was measured by producing a start pulse for a Beckman Universal Eput and Timer, Model 8370, when a speaker was energized and a stop pulse when the correct position was indicated.

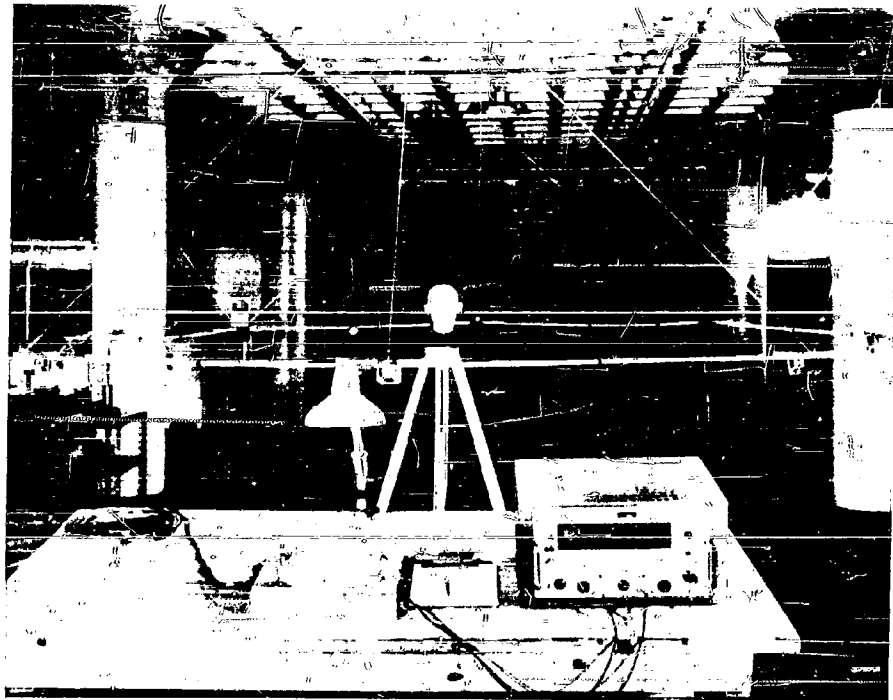


Figure 2.5-1 Localization Test Set-Up



Figure 2.5-2 Localization Test Subject

The actual positions of each incorrect answer were not recorded. However, observation of subjects revealed that any incorrect choice was seldom more than one position off. The exception was in front-back positions where the ambiguity was not so readily resolved by most subjects.

A random noise generator was used to excite the speakers. A single test consisted of selecting each speaker six times at random. This test was repeated six times to complete one subject's series. After the third test the head was reoriented within the ring in order to minimize locating certain speakers by identifiable characteristics caused by reverberations. The button switches on the subject panel were similarly turned by relative rotation of the two halves of an eight pin connector so that each button remained in the same relative position with respect to the head throughout the series. The subject was unaware of the change. An initial calibration test in which the subject stood at the center of the ring was also conducted to determine any localization peculiarities and to permit an evaluation of the coupling system. A total of eleven subjects were tested.

2.5.2 Test Results

1. Figure 2.5-3 is a graphical plot of the number of mistakes made in any one test (48 trials) and the number of tests.
2. Figure 2.5-4 is a polar diagram of localization accuracy of various individuals. Figure 2.5-5 is the average accuracy of six subjects.

A learning time would be indicated by a reduction in the number of incorrect choices and a corresponding decrease in the average time to select the correct position. However, such trends were not found. The conclusions therefore are that no learning period is necessary or the

tests were not conducted over a sufficient length of time to observe any improvement. The fact that certain individuals who were directly involved in this research had high scores suggests other reasons for the lower scores other than a lack of learning. These might include mental disposition vis-a-vis the tests, fatigue, or disinterest.

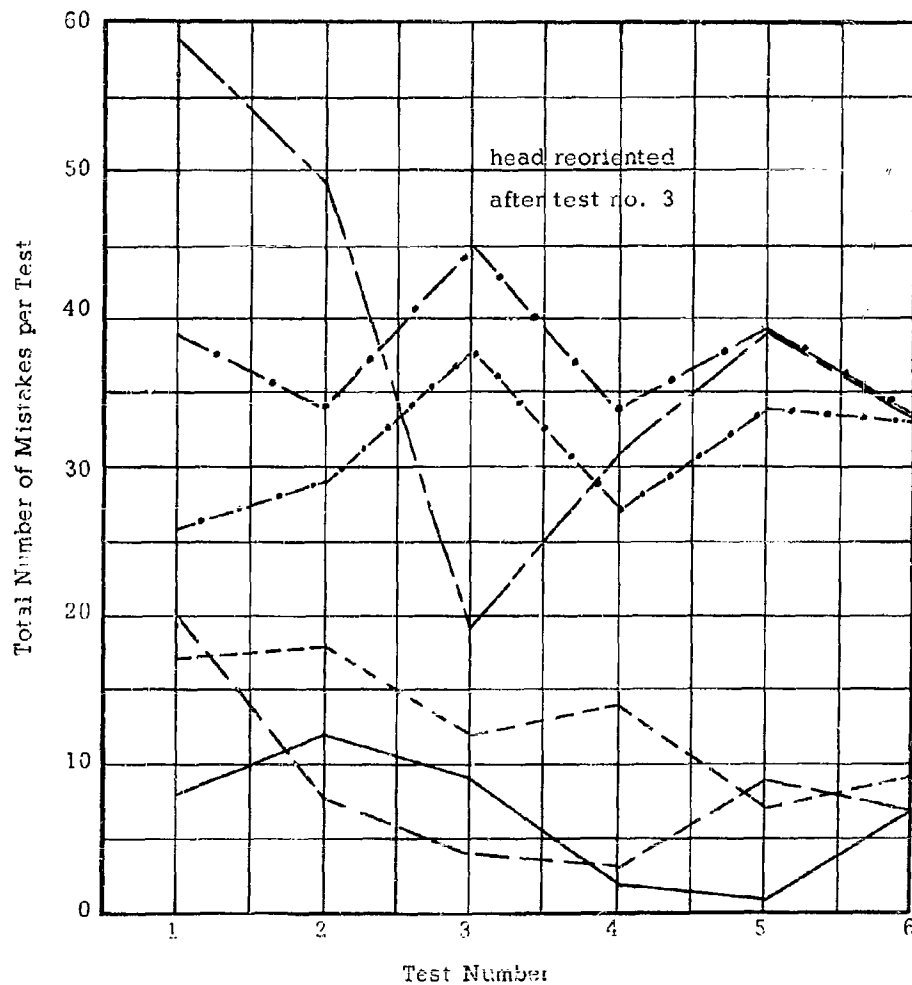


Figure 2.5-3 Learning Curves of Six Subjects

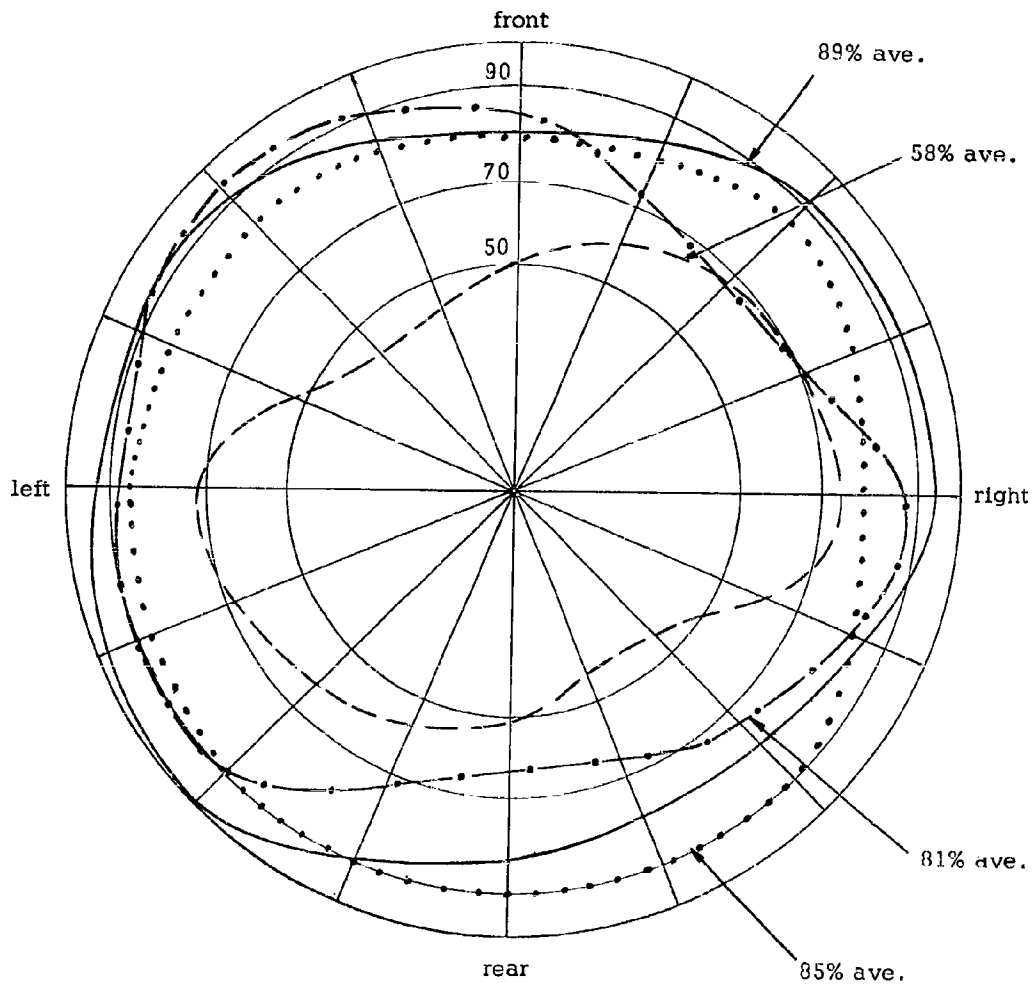


Figure 2.5-4 Azimuth Localization Ability as a Function of Direction
Using Aural Coupling - Results for Four Subjects

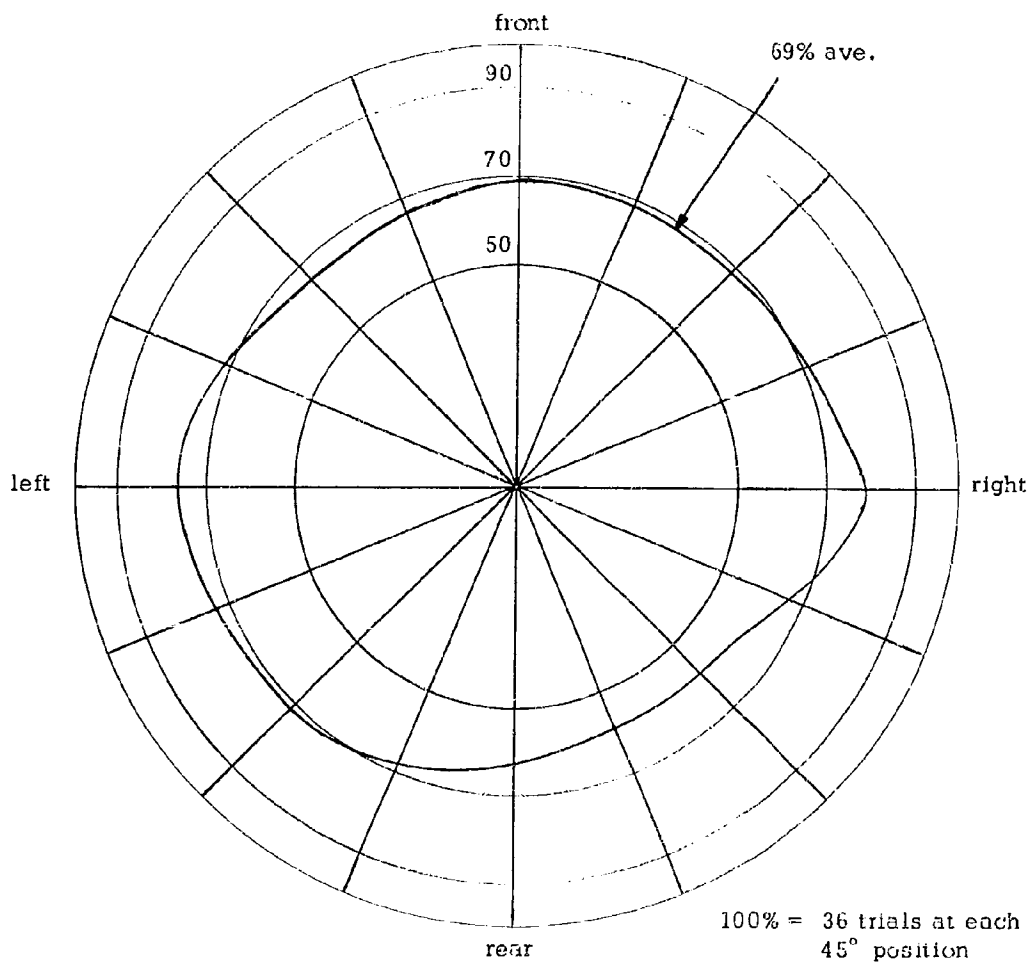


Figure 2.5-6 Azimuth Localization Ability as a Function of Direction Using Aural Coupling - Average for Six Subjects

CHAPTER 3

PORPOISE TRANSLATORS

3.1 Equipment

3.1.1 Bases for Design

The theoretical aspects of man-to-porpoise communication are discussed in another section of this report. Evolving from the theory and from experimental work are the bases for the design of equipment to facilitate man-porpoise communication experiments:

1. Porpoise whistles (not echo locating sounds) range from 4 to 18 kc for the most part.
2. The characterizing time interval for human vowel sounds ranges from 400 to 2600 microseconds.

It was decided to limit the initial work to translation of vowel sounds only. The intention was to translate the porpoise whistles into vowel sounds which when repeated by an operator would generate the original porpoise whistle. Two instruments were designed and constructed for this purpose. The first is termed the porpoise-to-man translator, the second the man-to-porpoise translator. Model I designs were fabricated during the summer; these instruments were tested (see other section) and on the basis of these experiments, Model II designs were prepared and the instruments fabricated later in the year. Improvements in frequency stability, reduction of long-tailed transients, gating and matching of the two units were made. Further details are contained in later paragraphs.

These instruments are based on preserving a relationship between the porpoise whistle frequency and the characteristic time interval of the vowel sounds; 4 to 18 kc in the porpoise whistle corresponding to 2600 to 400 microseconds in the vowel sounds.

3.1.2 Operating Principles of Model I Designs

1. Porpoise-to-Man Translator

The internal functioning of the Model I porpoise-to-man circuits is shown in block diagram form in Figure 3.1-1. Basically, the porpoise

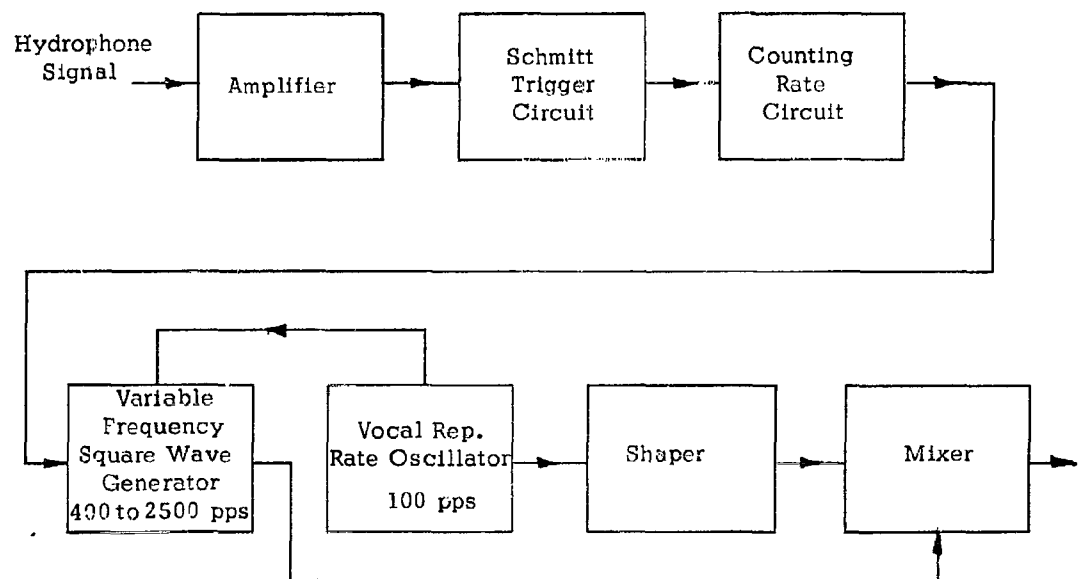
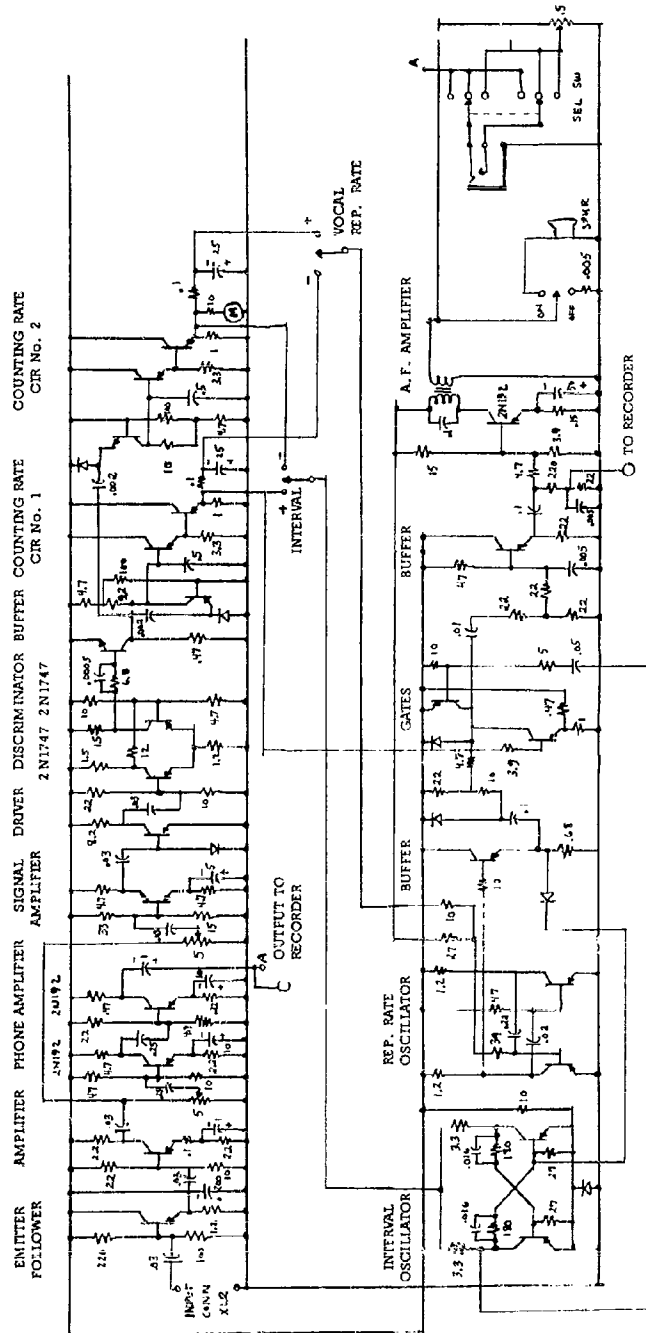


Figure 3.1-1 Functional Diagram of Model I Porpoise-to-Man Translator

whistle signal is amplified, then fed to a Schmitt trigger circuit. This trigger circuit produces an output voltage of one of two values depending on whether the signal is plus or minus. In this way the porpoise signal is changed to a square wave. This square wave feeds a counting rate circuit which produces a voltage proportional to the frequency of the incoming square wave. This voltage in turn controls the frequency of a multivibrator. This multivibrator is synchronized by the vocal repetition rate oscillator. The two waveforms produced by these oscillators are combined in a mixer to produce the synthetic voice output of the system. In the actual circuits the direction of oscillator frequency change with porpoise whistle frequency is controllable and the vocal repetition rate is varied slightly in accordance with porpoise whistle frequency. The output signal is gated off when the porpoise whistle frequency is less than 3 kc. Figure 3.1-2 is the circuit diagram. The instrument is housed in a case similar to that of the Model II. Details of performance are omitted, there being detailed information on the Model II instrument.

2. Man-to-Porpoise Translator

The internal functioning of the man-to-porpoise circuitry is illustrated in the block diagram of Figure 3.1-3. Voiced sounds are picked up by the microphone then amplified and fed to a Schmitt trigger circuit. The trigger circuit changes the signal waveform to a square wave. The output transitions occur at the zero crossings of the input. The period to voltage circuit produces an output voltage linearly related to the period between positive and negative going transitions of the square wave. The value obtained is clamped and held until the next value is obtained. A measure is made for each plus to minus transition. The voltage obtained by this process is filtered slightly and then used



NOTES

CAPACITORS: VALUES IN MFD
RESISTORS: VALUES IN KΩ-OHMS
TRANSISTORS: PNP'S ARE 2N591'S UNLESS MARKED
NPN'S ARE 2N438A'S

DIODES: IN3193
METER: 0-500 μA

UNITED RESEARCH INC.
PORPOISE-TO-MAN TRANSLATOR
MODEL I

Figure 3.1-2 Circuit Diagram for Porpoise-To-Man Translator Mod. I

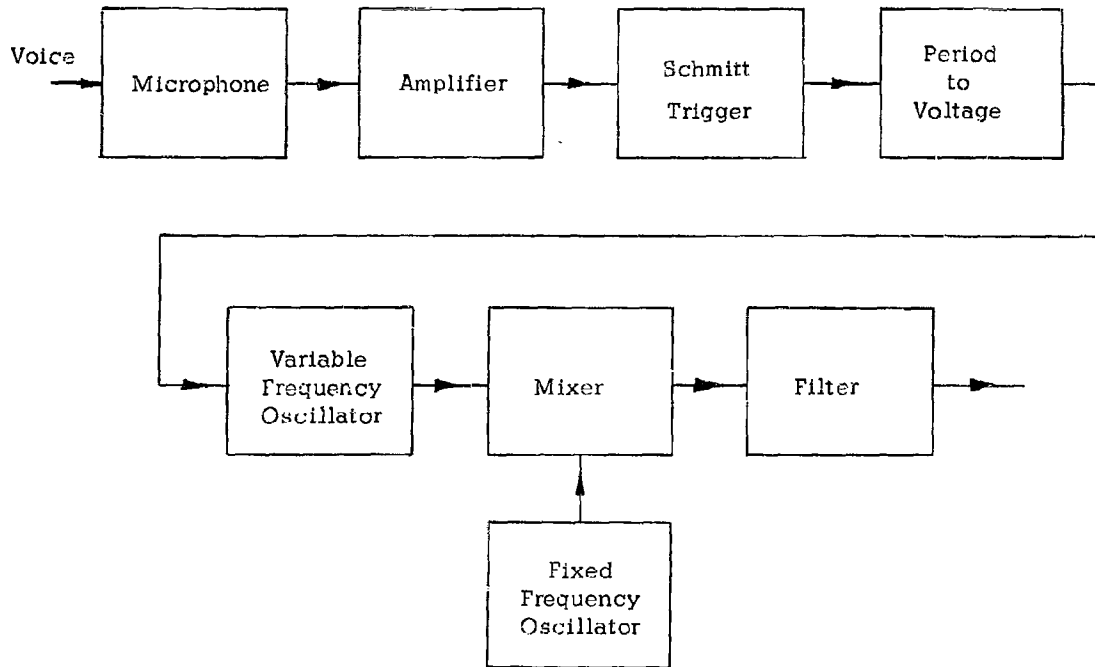


Figure 3.1-3 Functional Diagram of Model I Man-to-Porpoise Translator

to control the frequency of a variable frequency oscillator. Output of this oscillator mixes with the output of a fixed frequency oscillator producing a beat frequency which varies from 4 to 18 kc. This beat frequency is filtered out from the composite signal obtained from the mixer and is the output of the system. A gate serves to shut off the output in the absence of a voiced signal. Figure 3.1-4 is the circuit diagram. This instrument is portable, being contained in an instrument case similar to the Model II instrument. Details of performance are omitted so that full information may be presented on the Model II instruments.

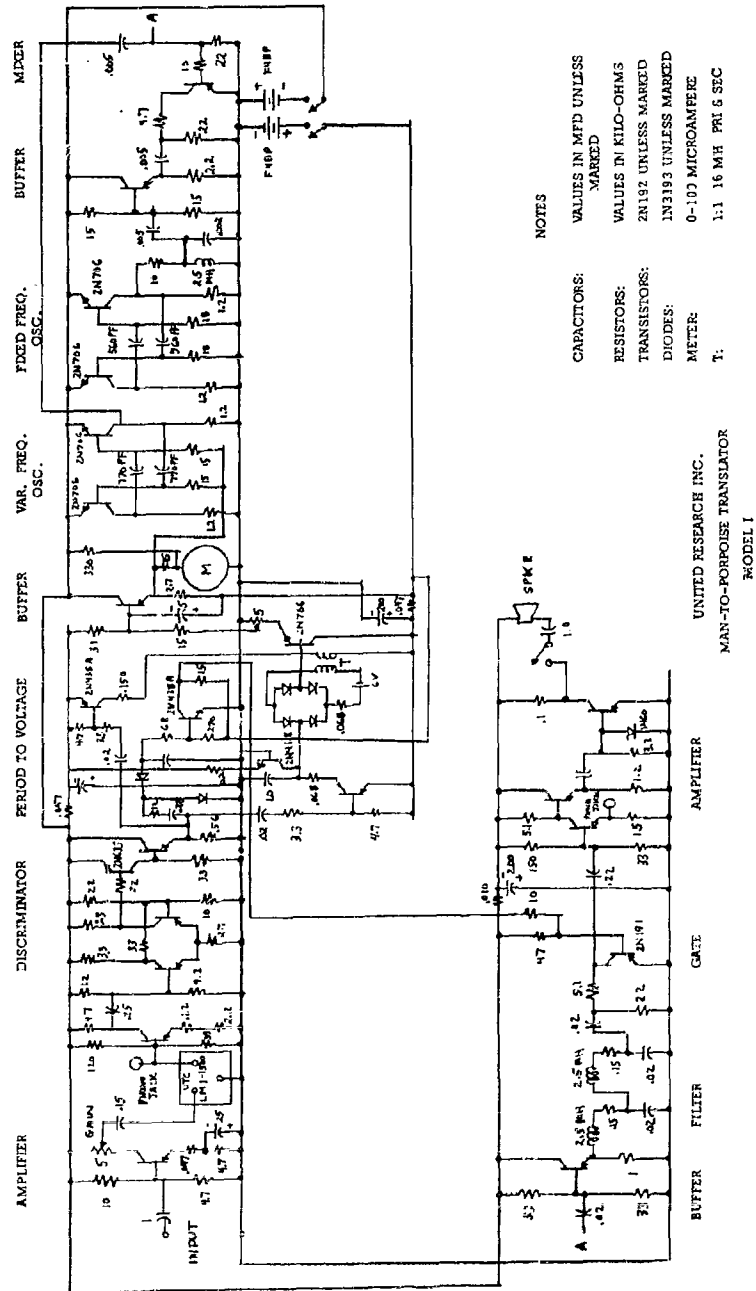


Figure 3.1-4 Circuit Diagram for Man-To-Porpoise Translator Mod. I

3.1-3 Model II Instruments (Figures 3.1-6, 3.1-7)

1. Porpoise-to-Man Translator

The internal functioning of the Model II porpoise-to-man translator is illustrated in Figure 3.1-5. The difference between this and the Model I instrument is in the method of deriving a voltage proportional to input frequency. In the Model II circuit the number of cycles of signal in a 12.5 millisecond period are counted, a voltage proportional

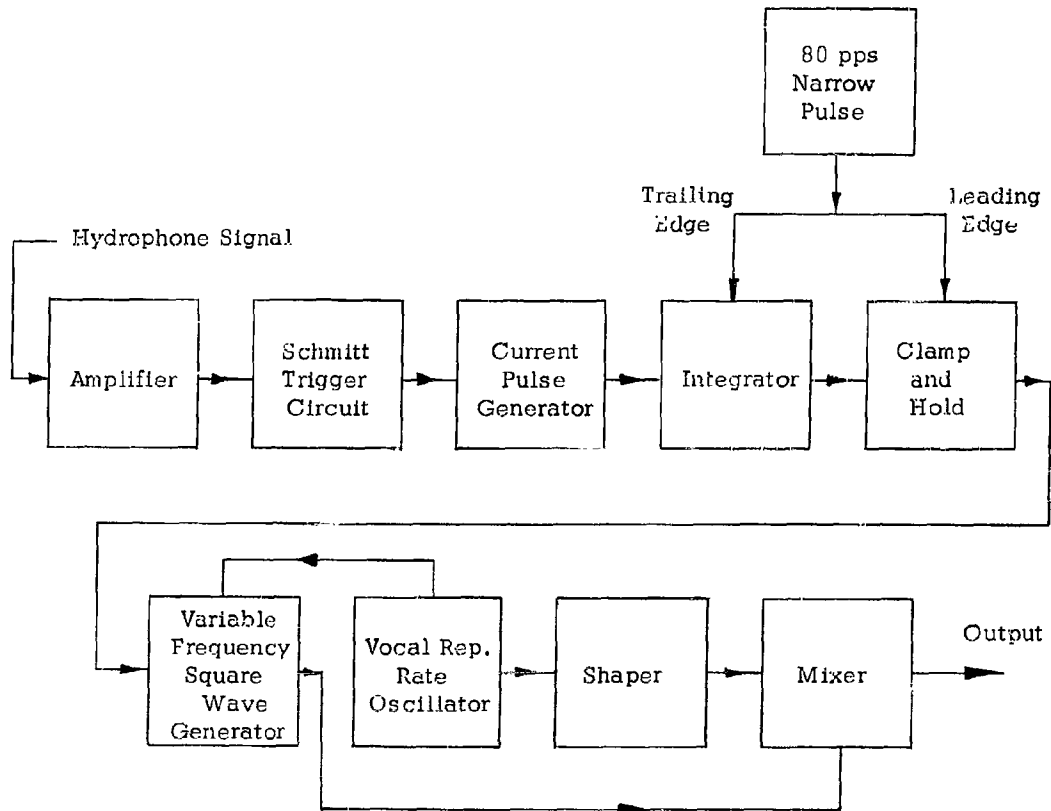


Figure 3.1-5 Functional Diagram of Model II Porpoise-to-Man Translator



Porpoise-To-Man
Translator

Man-To-Porpoise
Translator

Figure 3.1-6



Porpoise-To-Man
Translator

Man-To-Porpoise
Translator

Figure 3.1-7 Internal Construction

to this count is obtained, clamped and held until the next measure is obtained. This method of obtaining a voltage proportional to frequency eliminates transients due to the averaging circuits required in the standard counting rate circuits. For sudden increases or decreases of input frequency, steady state response occurs within 25 milliseconds. In other respects, the block diagram is the same as for the Model I circuit. Figure 3.1-8 is a circuit diagram. Here several significant differences are apparent: (1) the frequency to voltage circuit is much more complex than that previously used. (2) the interval oscillator is of a different and improved design. (3) the mixer and gate transistor is a special purpose one and (4) gate control is by a Schmitt circuit to provide a positive action.

Performance characteristics are presented following the discussion of the man-to-porpoise translator.

2. Man-to-Porpoise Translator

The Model II man-to-porpoise translator has the same internal functioning as the Model I circuit shown in Figure 3.1-4. Circuit improvements have been made. These may be seen in the diagram of Figure 3.1-9 to be: (1) The period to voltage circuit has been simplified and includes an improved gate circuit. (2) A special design of buffer amplifier is used. (3) Special purpose gate transistors are used. (4) Output gating is controlled by a Schmitt trigger which assures positive action.

Performance characteristics are presented in the next section.

3.1-4 Circuit Behavior

1. Porpoise-to-Man Translator

Figure 3.1-10 illustrates the translation of 4 to 18 kc sine waves into synthetic speech. Figure 3.1-11 illustrates the reverse process with sine wave input to the man-to-porpoise translator. Figure 3.1-12 illustrates the behavior when the output of the porpoise-to-man translator was used to feed the man-to-porpoise translator. This test translates a whistle to a synthetic speech and then from the synthetic speech back to whistle. Figure 3.1-13 illustrates the reverse process.

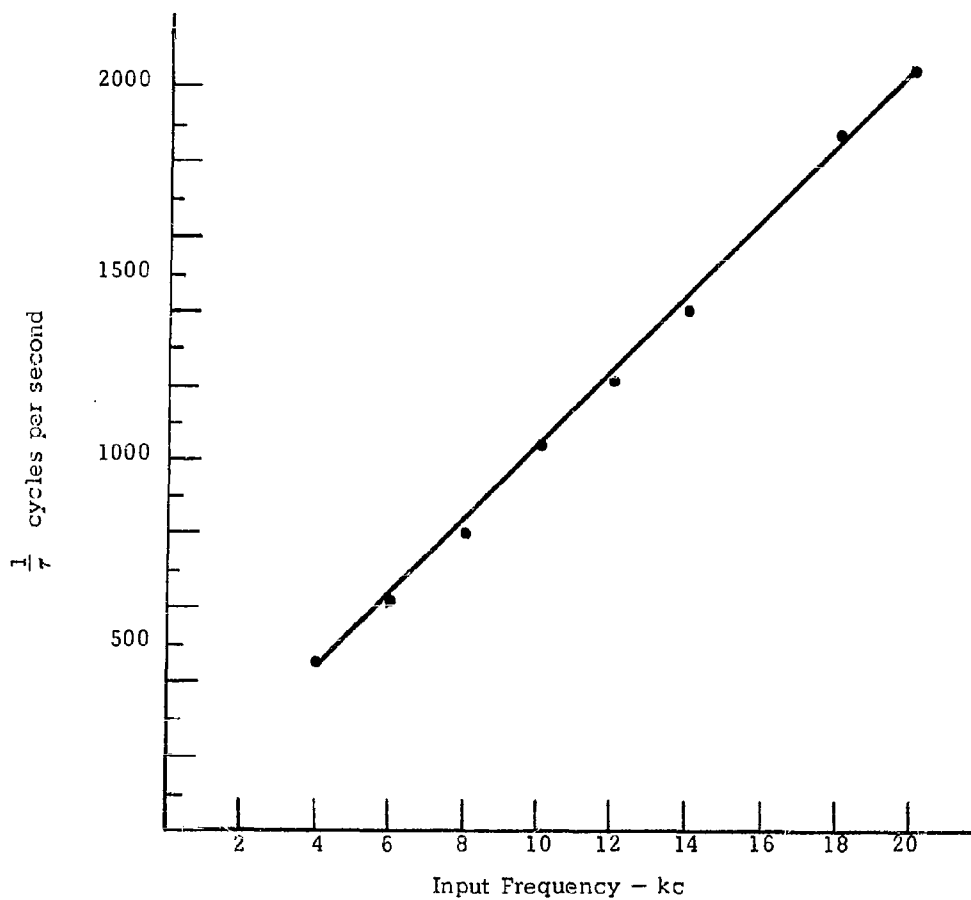
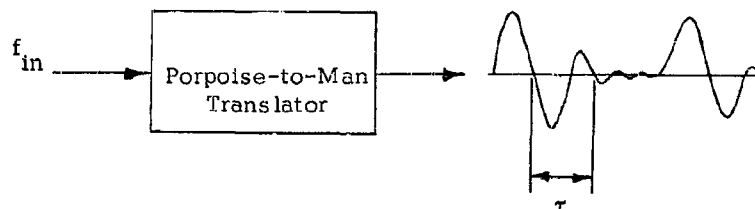


Figure 3.1-10 Characteristic of Porpoise-to-Man Translator

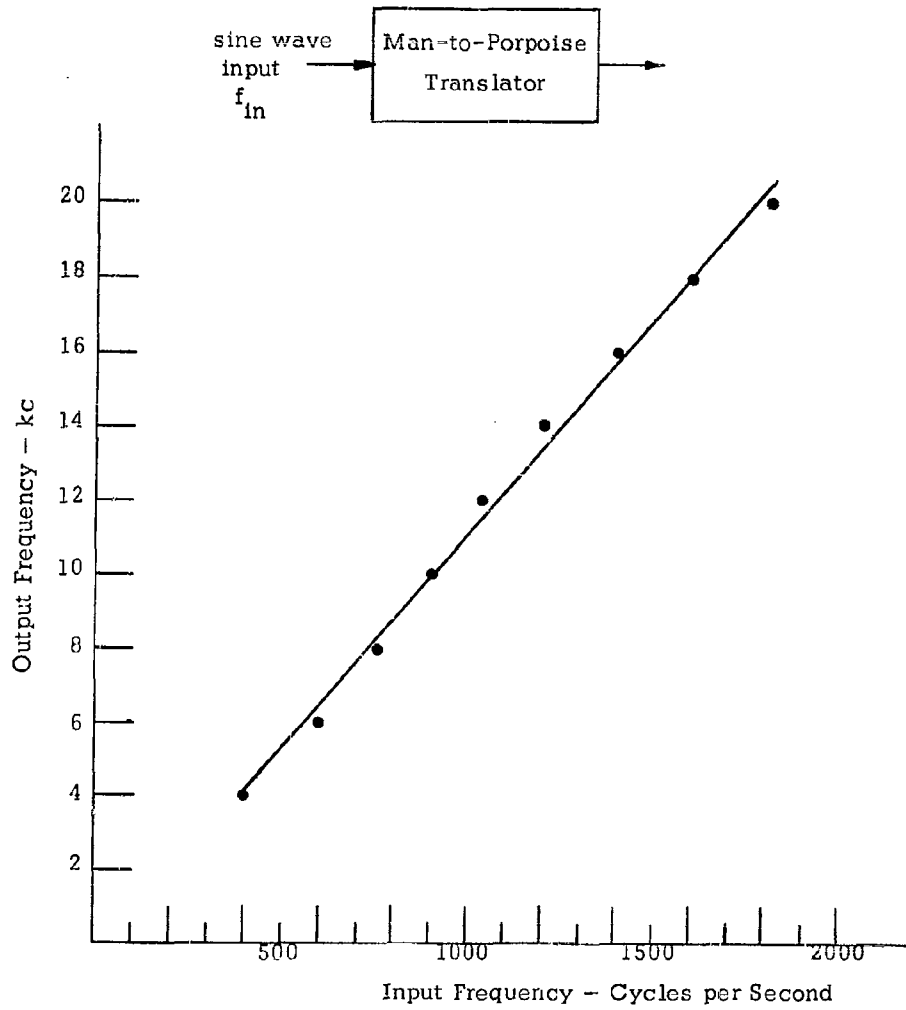


Figure 3.1-11 Characteristic of Man-to-Porpoise Translator

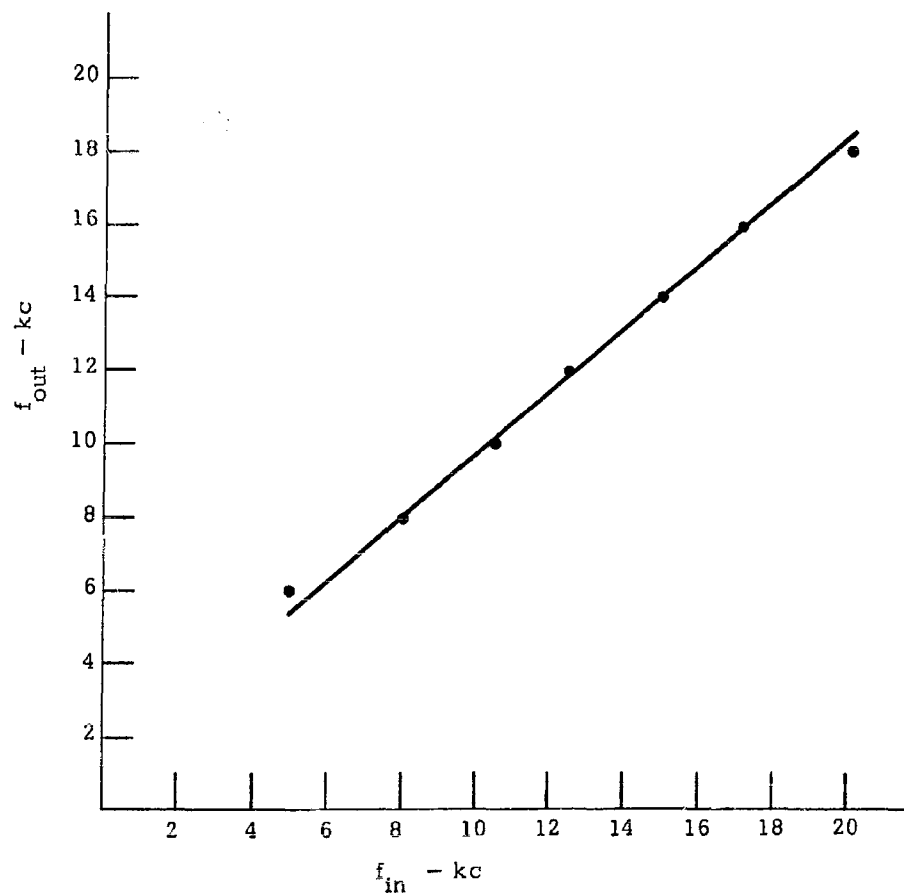
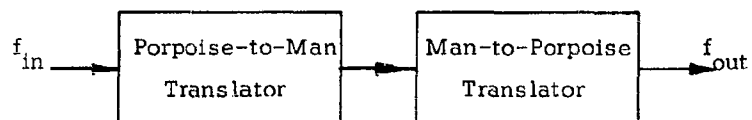


Figure 3.1-12 Tracking Characteristic of Translators

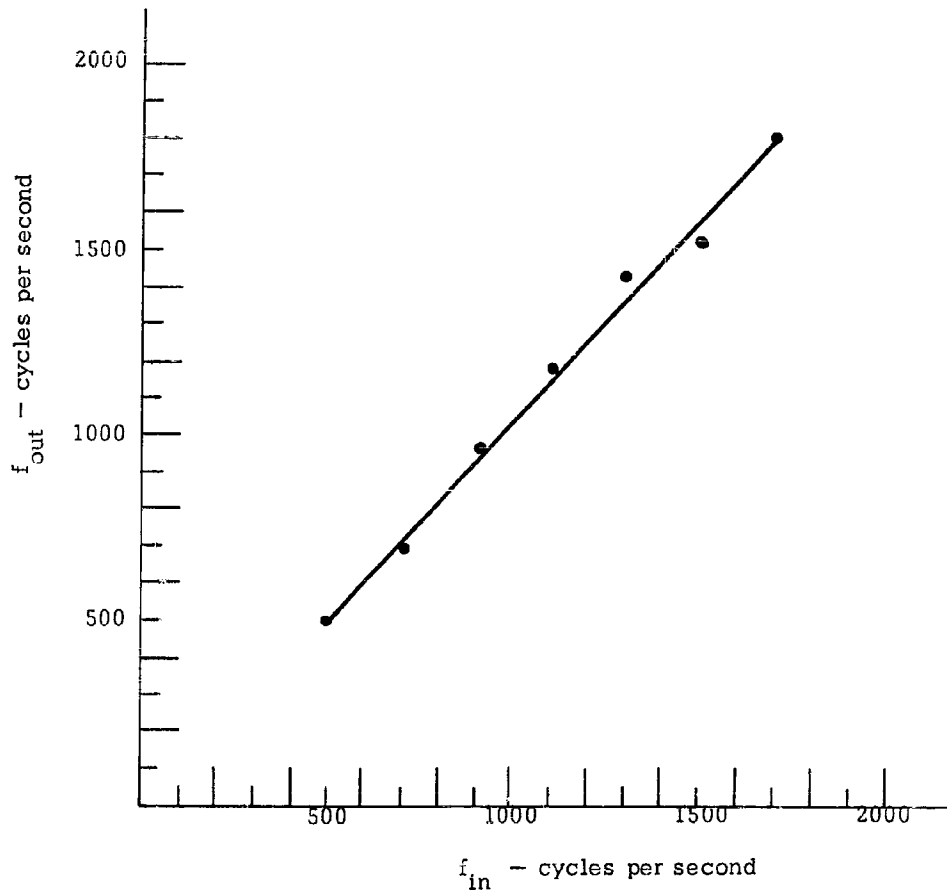
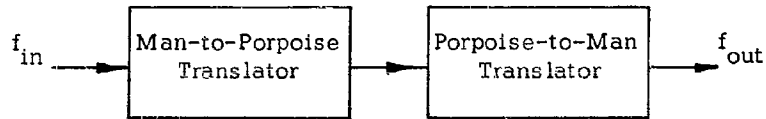


Figure 3.1-13 Tracking Characteristic of Translators

3.2 Theoretical Basis for Translators

The mechanism of time delays which has been found to be essential in sound localization was extended to speech recognition. As reported earlier [Ref. 2], speech is presumed to be formed by a delay mechanism similar to that provided by the pinna wherein the sound originating in the larynx reaches the front of the mouth by diverse paths defined by the tongue, cheeks, teeth, lips, nasal passages, etc. An oscillographic study of phoneme wave forms revealed that each phoneme has a characteristic time delay irrespective of the speaker. The delay times for phonemes "oh" to "ee" were estimated to be from 2000 μ sec to 400 μ sec. Thus, in the translation scheme used, it was decided to process the frequency modulated porpoise whistles so that a 4 kc frequency produces a 2000 μ sec characteristic time and an 18 kc frequency produces a 400 μ sec characteristic time in the synthetic speech output from the porpoise-to-man translator. Intermediate frequencies produce a period inversely related to frequency. Vocal cord repetition rate was arbitrarily set at approximately 100 pulses per second.

The reverse mode, the translation of voiced sounds to porpoise-like whistles, is programmed in a similar way. A characteristic time of 2000 μ sec is translated to a whistle frequency of 4 kc and a 400 μ sec characteristic time is translated to an 18 kc frequency. Translation of intermediate frequencies is on a frequency proportional to inverse period basis.

3.3 Test of Translators

The experimental work with porpoises using the translators was most gratifying. The tests were conducted at the U. S. Naval Ordnance Test Station facility at Point Mugu.

A set of words was selected for use in testing mimicry with the porpoise. The words used were "ball," "ezra," "peterson," and "focina." Of these only "ball" was reproduced accurately by the porpoise-to-man translator. Dependence on simple recognition allowed the remaining words to be used. A half-four tape recording of porpoise sounds was made for reference. Then the word "ball" was spoken into the man-porpoise translator. Silence ensued, but after three repetitions of the word, the porpoise vocalized a similar sound as translated by the porpoise-to-man unit. Several minutes of this exchange followed, with reinforcement by feeding mackerel, until it seemed that mimicry was present. The selection of vocabulary by the porpoise during this time was limited almost entirely to "ball"-like words. The same process was repeated with "ezra" with apparent mimicry by the porpoise, but with more frequent use of his standard vocabulary. The word "peterson" followed.

Upon the introduction of "peterson," still more of the standard vocabulary was used by the animal, and mimicry did not seem to occur. However, an interesting thing did occur. When silence reigned, "peterson" was spoken into the translator and the porpoise responded immediately with "why-oh." This was repeated at least ten times. It was concluded that "peterson" would not be repeated. The same reaction occurred with "focina."

The experiment was ended at this point with reservation. Responsiveness of the porpoise was certain whereas mimicry was somewhat less certain. Communication was definitely present but the game suffered from lack of meaning for the porpoise.

Perhaps the most important result was the knowledge that the design and conduct of future experiments must include consideration of the animals' reactions to intrusions on his environment. Meaningful results can be expected only when the test is designed so as to insure that the porpoise retains complete control of his environment even while responding to outside influences.

CHAPTER 4

THEORY

4.1 Introduction

At this reporting point it seems desirable to review the study of the role of the pinna in human localization in order to make a coherent whole of what has been done. Although the study of this particular topic began with the demonstration that distorting the pinnae distorts the perception of locale, these are several background ideas which are essential to the formation of hypotheses and theories regarding the function.

The first background theorem concerns the number of independent indices necessary to define a point and the consequent requirements regarding transformations of a space of points. If the point is three-dimensional, requiring three indices to define, then not less than three indices must be preserved in transformation if the character of the space is to be preserved. If a sound source is located in a room, then a coordinate system of referent and three referreds is necessary to specify its location.

If the information space is acoustical, then there must be a referent and three referred acoustical features to the point in order to make a correspondence in location between the physical space and the acoustical space. The sound from the sound source must therefore be complex, consisting of at least four sine wave components of differing frequencies. This requirement is met by transient sounds, usually consisting of a broad spectrum of components, more than sufficient to form a basis for measurement at a point.

There may be a joint arrangement of physical and acoustical features in the measurement. If measurement is made at two points in the physical space, the requirement on the acoustical space is reduced by one dimension,

and so on. If measurement is made at a coordinate set of points in physical space, the sound source may consist of a single sine wave.

The second background theorem concerns the transformation inverse to the one of perception. The initial or perceptual transformation forms a correspondence between the source and the observer, it is then necessary to form a correspondence between the observer and the source which properly identifies the source coordinates. If the inverse transformation cannot be constructed, then the location of the source remains unknown. In general, only approximate or probabilistic inverse transformations are constructable. This may be considered as correspondence in measured subsets of the point sets concerned.

4.2 Transformation by the Pinna

Given the background requirements, the pinna must introduce a characterizing transformation between the sound source and the eardrum which preserves the dimensionality pertinent and for which an inverse can be constructed. Our early observations showed that the pinna introduced a number of paths of different lengths between the sound source and the eardrum. The transformation thus defined showed variation with azimuth angle, altitude angle, and range such that the delay lengths for azimuth angle and altitude angle could be measured, but the range information was less obvious and remains to be well specified.

In order to understand the apparent character of the transformation, we introduced simplified models. Of these, the first was the "two-hole coupler" shown in Figure 4.2-1. This model was patterned after the "directional coupler" familiar to users of wave guides, but has a range of angles between 0-180° instead of only the two limit points.

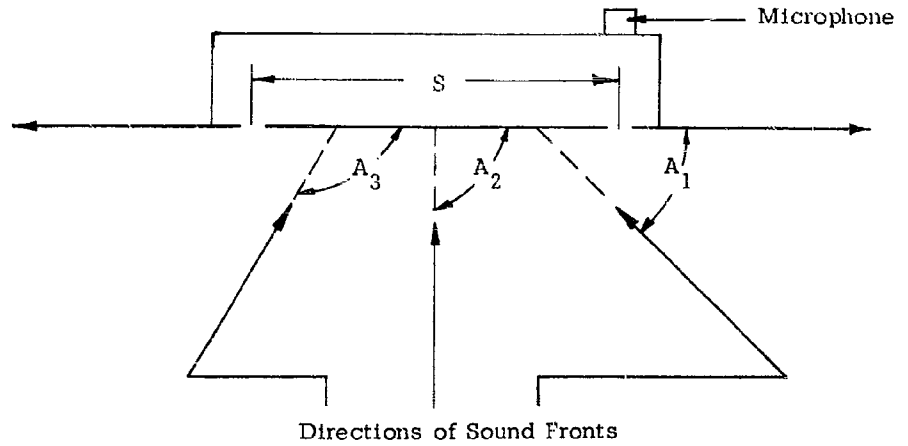


Figure 4.2-1 Two Hole Coupler

If a sound pulse is propagated at the angles indicated, the resultant output from the microphone (neglecting reverberation) will be as shown in Figure 4.2-2.

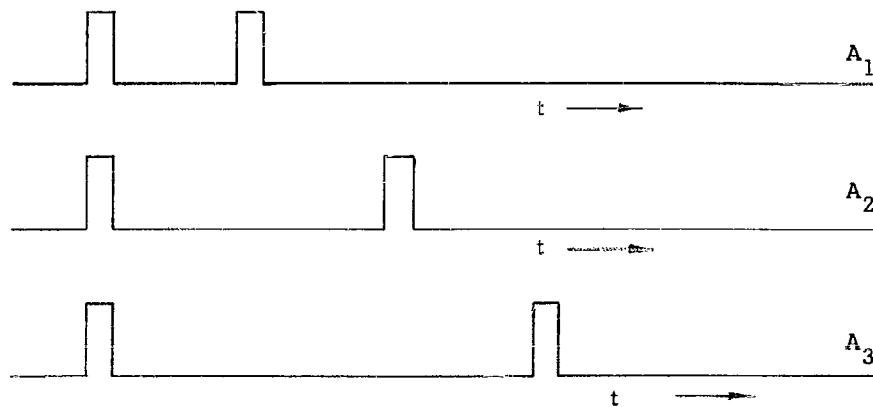


Figure 4.2-2 Signals from the Two Hole Coupler

A measurement of the time between pulses would provide the information from which the angle of arrival could be computed. The time between pulses at the microphone follows equation (4.2-1)

$$\tau = \frac{s}{v} (1 + \cos A) \quad (4.2-1)$$

$$\tau = \text{time between pulses} \quad (4.2-2)$$

$$s = \text{distance between holes} \quad (4.2-3)$$

$$v = \text{velocity of sound} \quad (4.2-4)$$

$$A = \text{angle of normal to wave front} \quad (4.2-5)$$

with hole line

Following the two hole coupler, a "three hole coupler" was examined as a model. The three hole coupler may be considered as a coordinate pair of two hole couplers with a common origin or referent. A three hole coupler is shown in Figure 4.2-3.

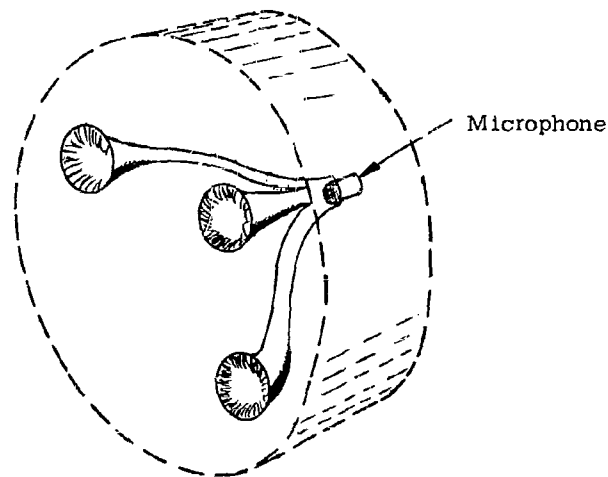


Figure 4.2-3 Three Hole Coupler

In order for the inverse transformation to provide a unique correspondence between pulse time separation and azimuth angle and altitude angle, it is necessary to separate the domains by a fixed delay path as shown in the figure. The equations pertinent are as follows:

$$\tau_1 = \frac{S_\theta}{v} (1 + \cos \theta) \quad (4.2-6)$$

$$\tau_2 = \frac{kS_\theta}{v} + \frac{S_\Psi}{v} (1 + \cos \Psi) \quad (4.2-7)$$

$$\tau_1 = \text{delay corresponding to } \theta \quad (4.2-8)$$

$$\tau_2 = \text{delay corresponding to } \Psi \quad (4.2-9)$$

$$S_\theta = \text{distance between holes on } \theta \text{ line} \quad (4.2-10)$$

$$S_\Psi = \text{distance between holes on } \Psi \text{ line} \quad (4.2-11)$$

$$K \geq 2 \quad \text{domain separating delay path}$$

These models provided insight into possible manners of transformations defining angles, but did not provide for range information. There are a number of possibilities as yet unsettled regarding the inference of range. In examining the requirements for acoustic transformation to provide localization, a geometric form was evolved which is similar to an ear. For localization in azimuth, the transforming mechanism should be altitude independent and function over 180 degrees. Figure 4.2-4 was sketched. The reflector-diffractor was made half-round to provide 180 degree effects, whereas a full round would be ambiguous because of symmetry.

A similar structure would provide altitude angle information, but must be displaced. The focal point could be ducted to the microphone to produce the geometry sketched in Figure 4.2-5.

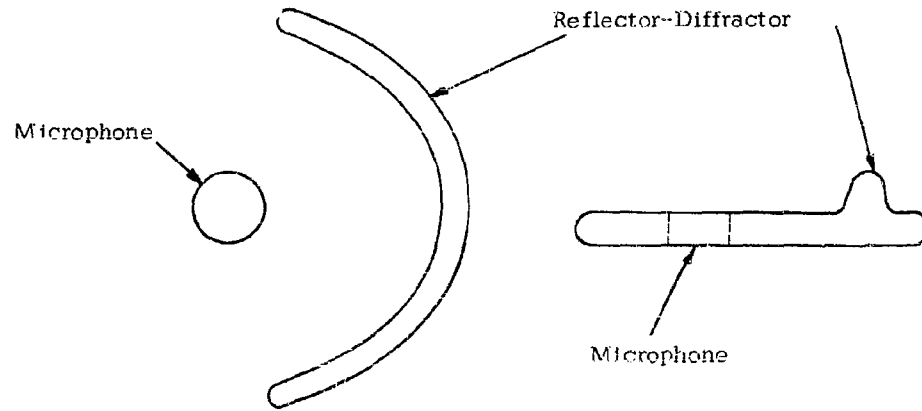


Figure 4.2-4 A Diffraction System Providing Azimuth Delays

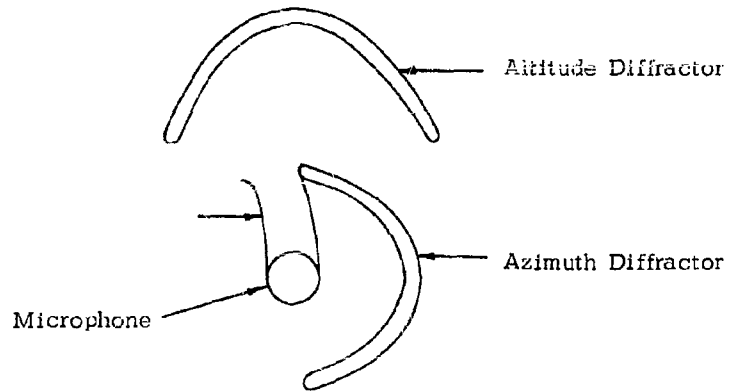


Figure 4.2-5 A Diffractor System Providing Delays in Azimuth and Altitude

It is evident that only the curvature of the incident wave front can provide range information to the transformation being considered, so that such measurement will be poorly resolved at best. A second diffractor, Figure 4.2-6, would provide curvature information by small delay differences, shown by double pulses in Figure 4.2-7. Note that the coordinates may also be skewed rather than orthogonal, suggesting the geometry of the human ear.

The pertinent equations are formed by two sets of three hole coupler equations.

4.3 Inverses to the Pinna Transformation

Assuming that the transformations described are pertinent, as supported by observation, the means of forming the inverse correspondence becomes the next question. This was approached by attempting mathematical formulations of the inverse transform. The inverse first constructed is essentially that for the two hole coupler, and can be done in several ways. The first way is as follows:

Given a signal after passing through a delay

$$f(t) + af(t - \tau)$$

$$f(t) = \text{source sound}$$

$$\tau = \text{time delay in reflection path}$$

$$a = \text{attenuation factor at path reflection}$$

Construct the series

$$\begin{aligned} & \sum_{n=0}^M (-1)^n a^n [f(t - n\tau) + af(t - n\tau - \tau)] \\ &= \sum_{n=0}^M (-1)^n a^n f(t - n\tau) + \sum_{n=0}^M (-1)^n a^{n+1} f(t - \tau - n\tau) \end{aligned}$$

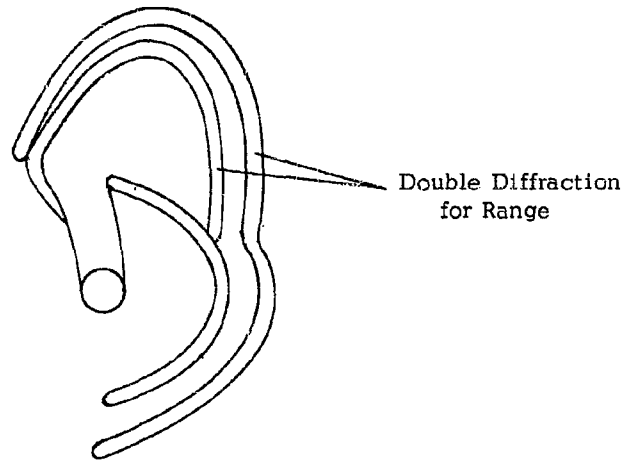


Figure 4.2-6 A Diffraction System Providing Delays in Azimuth, Altitude, and Range

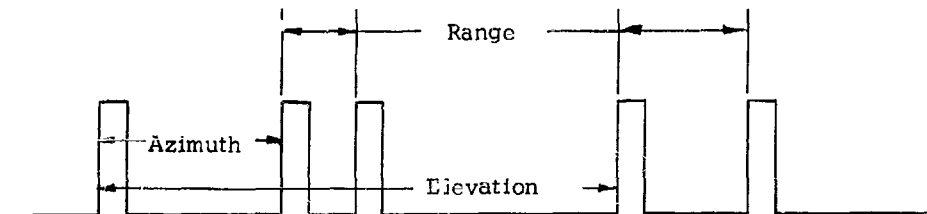


Figure 4.2-7 Hypothetical Presence of Range

$$\begin{aligned}
 &= f(t) + \sum_{n=1}^M (-1)^n a^n f(t - n\tau) \\
 &\quad + \sum_{n=1}^{M+1} (-1)^{n-1} a^n f(t - n\tau) \\
 &= f(t) + \sum_{n=1}^M (-1)^n a^n f(t - n\tau) - \sum_{n=1}^M (-1)^n a^n f(t - n\tau) \\
 &\quad + (-1)^{M+1} a^{M+1} f(t - \tau - M\tau) \\
 &= f(t) + (-1)^{M+1} a^{M+1} f(t - \tau - M\tau) \\
 &= h(t)
 \end{aligned}$$

$$\lim_{M \rightarrow \infty} h(t) = f(t) \quad a < 1$$

$$M \rightarrow \infty$$

Although the first method produces a satisfactory inverse, there is a second way which converges more rapidly.

Let $P_1(t) = f(t) + af(t - \tau)$ be the signal after delay and addition. Then construct a series

$$P_n(t) = P_{n-1}(t) + a^{2^{n-2}} (-1)^{2^{n-2}} P_{n-1}(t - 2^{n-2}\tau)$$

$$2 \leq n \leq \infty$$

And by examination, we find that a consequence is that

$$P_n = f(t) - a^{2^{n-2}} f(t - 2^{n-2}\tau)$$

Using the rule of construction and P_n , we find

$$\begin{aligned}
 P_{n+1} &= f(t) - a^{2^{n-2}} f(t - 2^{n-2} \tau) \\
 &\quad + a^{2^{n-2}} (-1)^{2^{n-2}} f(t - 2^{n-2} \tau) \\
 &\quad - a^{2^{n-2}} a^{2^{n-2}} (-1)^{2^{n-2}} f(t - 2^{n-2} \tau - 2^{n-2} \tau) \\
 &= f(t) - a^{2^{n-2}} f(t - 2^{n-2} \tau) \\
 &\quad + a^{2^{n-2}} f(t - 2^{n-2} \tau) \\
 &\quad - a^{2^{n-2}} a^{2^{n-2}} f(t - 2^{n-2} \tau - 2^{n-2} \tau)
 \end{aligned}$$

And

$$(a^{2^{n-2}})^2 = a^{2^{n-1}}$$

and

$$2^{n-2} \tau + 2^{n-2} \tau = 2^{n-1} \tau$$

so

$$P_{n+1} = f(t) - a^{2^{n-1}} f(t - 2^{n-1} \tau)$$

By finite induction, starting with P_2 , the expression of P_n , P_{n+1} in terms of $f(t)$ etc. is correct.

Of course

$$\lim_{n \rightarrow \infty} P_n = f(t) \quad a < 1$$

The first mathematical model can be realized by an arrangement of delays, attenuations, and signed additions according to the sketch of Figure 4.3-1.

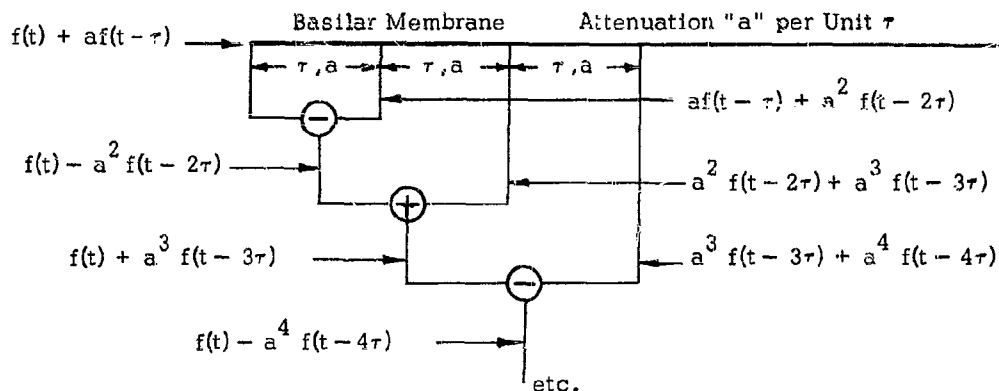


Figure 4.3-1 Construction of Series Inverse at Basilar Membrane

The second mathematical model can also be realized by an arrangement of delays, attenuations, and signed additions according to the sketch of Figure 4.3-2.

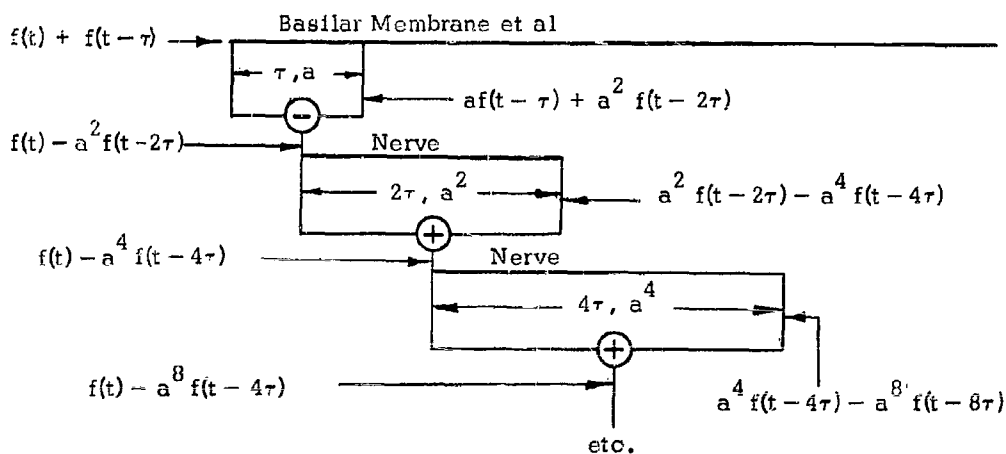


Figure 4.3-2 Construction of Octave Inverse in Nerve Trunks

If we introduce the Laplace transform, the mathematical expressions can be greatly simplified, although the details of construction are not evident as in the time models.

A delay is expressed as

$$e^{-s\tau}$$

The transformed signal of the first model is thus

$$F(s)(1 + a e^{-s\tau}) = H(s)$$

From which the inverse is obviously

$$F(s) = \frac{H(s)}{1 + a e^{-s\tau}}$$

Using the Laplace transform permits the expression of the more complex system of a spanning transformation.

The initial signal in this representation is

$$H(s) = F(s)(1 + a_1 e^{-s\tau_1} + a_2 e^{-s\tau_2} + a_3 e^{-s\tau_3})$$

for a spanning set. However, we might generalize

$$H(s) = F(s) \sum_{k=0}^N a_k e^{-s\tau_k}$$

The inverse for this in transform representation becomes

$$F(s) = H(s) \sum_{h=0}^{\infty} (-1)^h \left[\sum_{k=1}^N a_k e^{-s\tau_k} \right]^h$$

The time domain realization of $F(s)$ is then a series of delays with particular attenuation factors. Obviously an approximation

$$F(s) = H(s) \sum_{h=0}^{\infty} (-1)^h \left[\sum_{k=1}^N a_k e^{-s\tau_k} \right]^h \quad \begin{array}{l} a_0 = 1 \\ \tau_0 = 0 \end{array}$$

converges when the reflection factors are less than unity.

The inverse can be constructed as shown by the equation as continued networks of networks similar to the one shown in Figure 4.3-3.

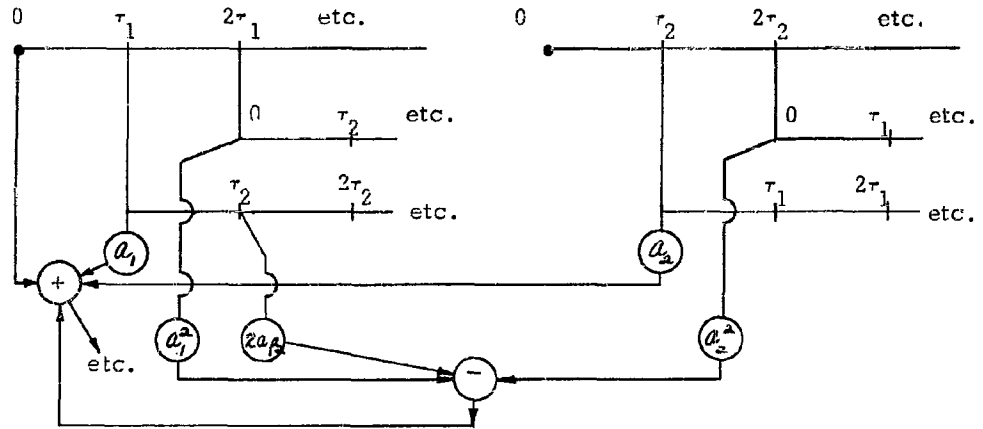


Figure 4.3-3 Series Inverse for Several Delays

However, a third mathematical model is possible using a feedback system as shown in Figure 4.3-4.

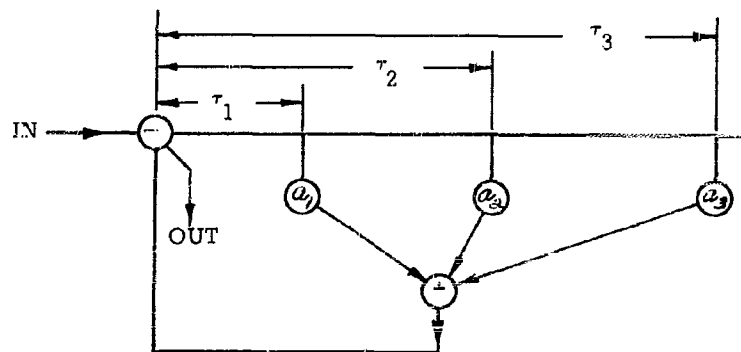


Figure 4.3-4 Feedback System for Inverting a Set of Three Delays

The mathematics are as follows for the two hole coupler

$$H(s) = F(s) (1 + ae^{-s\tau})$$

The "error signal" in the feedback system at the subtractor, \ominus is

$$E(s)$$

The error signal after delay is

$$E(s) e^{-s\tau}$$

The feedback equations are

$$E(s) = H(s) - E(s) ae^{-s\tau}$$

$$E(s) (1 + ae^{-s\tau}) = H(s)$$

$$E(s) = \frac{H(s)}{1 + ae^{-s\tau}} = F(s)$$

And for the more complex network shown:

$$E(s) (1 + a_1 e^{-s\tau_1} + a_2 e^{-s\tau_2} + a_3 e^{-s\tau_3}) = H(s)$$

$$E(s) = F(s)$$

4.4 Extensions of Theory

With the transformations introduced by the pinna defined in theory and measurement, and the inverses constructable by any of the models given, it is possible to extend the examination of the function of human hearing to the treatment of reverberation. One may observe that the external ear or pinna provides a characteristic reverberation by the introduction of multiple paths for the sound to reach the eardrum. Similarly, a room provides multiple paths between the sound source and the perceiver. In theory then, the significant difference between the role of the pinna and room reverberation is one of time domain and multiplicity of paths.

The reverberation of a room can be characterized as follows:

$F(s)$ = sound at source

$R(s)$ = sound at perceiver as modified by the room

$$R(s) = F(s) \sum_{n=0}^{\infty} a_n \epsilon^{-s\tau_n}$$

If we consider first the effect of two parallel walls of infinite extent as shown in Figure 4.4-1, the equation can be simplified

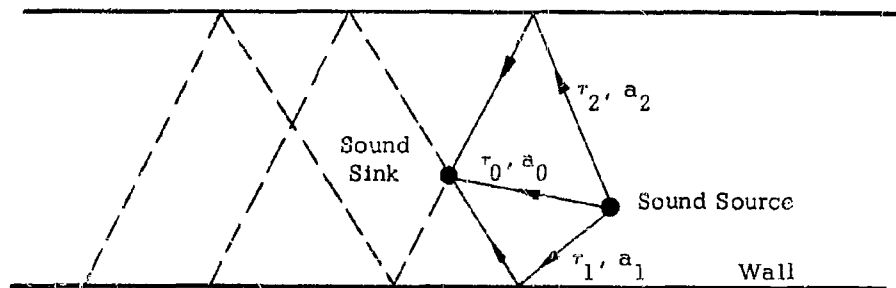


Figure 4.4-1 A Simple Reverberant Situation

$$R(s) = F(s) [a_0 \epsilon^{-s\tau_0} + a_1 \epsilon^{-s\tau_1} + a_2 \epsilon^{-s\tau_2}]$$

The inverse to this equation is simply constructed by the methods given, thus removing the effect of reverberation. It would, however, be preferable to use the incident sound from the several paths to improve the signal-to-noise ratio for a particular condition.

If a finite set of delayed signals is considered, having

$$\tau_M = \text{maximum delay}$$

$$R(s) = F(s) \sum_{n=0}^M a_n \epsilon^{-s\tau_n}$$

and the neural transformation $T(s)$ constructed

$$T(s) = \sum_{n=0}^M a_n \epsilon^{-s(\tau_M - \tau_n)}$$

the resultant mental signal $P(s)$ is given

$$\begin{aligned} P(s) &= R(s) T(s) \\ &= F(s) \sum_{j=0}^M a_j \epsilon^{-s\tau_j} \sum_{k=0}^M a_k \epsilon^{-s(\tau_M - \tau_k)} \\ &= F(s) \epsilon^{-s\tau_M} \sum_{j=0}^M a_j \epsilon^{-s\tau_j} \sum_{k=0}^M a_k \epsilon^{+s\tau_k} \end{aligned}$$

The result is the original signal delayed by τ_M and modified by the product of sums. The product of sums can be represented as follows:

$$\begin{aligned} \sum_{j=0}^M a_j \epsilon^{-s\tau_j} \sum_{k=0}^M a_k \epsilon^{+s\tau_k} &= \sum_{n=0}^M a_n^2 + \text{cross terms} \\ \text{cross terms} &= \sum_{k=0}^M \sum_{j=0}^M a_k a_j \epsilon^{-s\tau_k + s\tau_j} \quad j \neq k \end{aligned}$$

In this product, the cross term result for $j \neq k$ can be given as follows

$$\begin{aligned} \sum_x a_x (\epsilon^{-s\tau_x} + \epsilon^{+s\tau_x}) \\ \tau_x &= \tau_k - \tau_j \quad j \neq k \\ a_x &= a_k a_j = a_j a_k \end{aligned}$$

This form indicates a symmetrical shift about the maximum delay in the general expression, and in the frequency domain indicates no phase shift of components but an amplitude change between $+a_j a_k$ and $-a_j a_k$. Thus such terms can be called "coloration." The general expression then becomes

$$P(s) = F(s) \epsilon^{-s\tau} \sum_{n=0}^M a_n^2 + \text{coloration}$$

The perceived signal would then be increased in amplitude by

$$\sum_{n=0}^M a_n^2$$

and colored by terms of the form

$$a_j a_k (\epsilon^{-sx} + \epsilon^{+sx})$$

none of which could alter the result by more than

$$\pm a_j a_k$$

For example, let,

$$\begin{aligned} a_0 &= 1 \\ a_1 &= .9 \\ a_2 &= .8 \\ a_3 &= .7 \\ a_4 &= .6 \end{aligned}$$

The increase in amplitude, I , is then

$$I = 3.30$$

and the maximum coloration, C_M

$$C_M = a_0 a_1 = .9$$

The subjective result of this method of attention would be

- a) Increased loudness
- b) Perceived with delay τ_M
- c) Colored somewhat in spectrum

Where sounds are weak, this method can be preferred to the direct inverses and appears to be used, from our subjective test regarding position and coloration.

It should also be evident that transformations by the pinna can be treated in the same way.

In an examination of the extensions of the theory it now seems reasonable to state tentatively several domains in time regarding correlations and inverses:

$0 \mu \text{ sec} < D_1 < 300 \mu \text{ sec}$	Pinna
$0 < D_2 < 800 \mu \text{ sec}$	Interaural
$600 \mu \text{ sec} < D_3 < 2200 \mu \text{ sec}$	Speech
$2.2 \text{ ms} < D_3 < 10 \text{ ms}$	Music
$10 \text{ ms} < D_4 < 200 \text{ ms}$	Reverberation

The remaining concerns for human hearing have been briefly presented in the previous final report. It is hoped that these may all be combined in the near future into a coherent statement regarding human hearing, localization, attention and recognition as well as models of possible nerve and synaptic systems to accomplish the theoretical and experimental results.

CHAPTER 5

CONCLUSIONS

5.1 Hydrophones

It is now apparent that a hydrophone suitable for underwater sound localization must be designed as an integral part of the mechanism which introduces the localization transformation. Implied is not only the performance specifications outlined earlier, but also the material and the configuration of that mechanism. Furthermore, it is now felt that enlarged ear replicas do not provide the optimum mechanism. While their use allows subjective evaluation of remote environment localization, they do not model accurately in water as do normal size ears in air. It may be stated simply that for preservation of time delays an enlargement by the sonic velocity ratio is sufficient. However, accurate acoustic modeling requires enlargement by the square of the ratio. Hence, a greater emphasis will be placed on the synthesis of the ear's function by a combination of hydrophones and fixed delays.

5.2 Headphones

The condenser headphones described earlier adequately transfer the localization function. Nevertheless, further improvements can be made in dynamic range, low end response, and sensitivity; however, no further research is contemplated in this area. It is expected that manufacturers of available products will introduce in the near future, headphones which will hopefully surpass the performance of the condenser headphones.

5.3 Effects of Structure

The reflection and diffraction effects of structure between the ears has been shown to be ineffective in changing the localization ability with electronic systems. In addition it has been found that the interaural distance may be changed without affecting localization.

5.4 Theory

The theoretical aspects of localization research have provided a deeper insight into the general area of acoustics. Some of these ideas with mathematical support are collected below. Others are included in Reference 3.

Given a signal modified by one delay and attenuation

$$H(s) = F(s)(1 + a\epsilon^{-s\tau})$$

An inverse can be constructed directly

$$F(s) = \frac{H(s)}{1 + a\epsilon^{-s\tau}} = H(s) \sum_{n=0}^{\infty} a^n (-1)^n \epsilon^{-sn\tau}$$

Utilizing feedback techniques to find the inverse gives

$$E(s) = H(s) - E(s)a\epsilon^{-s\tau}$$

$$E(s) = \frac{H(s)}{1 + a\epsilon^{-s\tau}} = F(s)$$

When feedback circuits are used, the inverse to a multiple delay is simply

$$F(s) = H(s) \sum_{n=0}^M a_n \epsilon^{-s\tau_n} \quad \begin{array}{l} a_0 = 1 \\ \tau_0 = 0 \end{array}$$

$$E(s) = H(s) - E(s) \sum_{n=1}^M a_n \epsilon^{-s\tau_n}$$

When correlation and recognition are both desirable a correlation function can be constructed. Given a signal $F(s)$ multiply delayed

$$H(s) = F(s) \sum_{n=0}^M a_n \epsilon^{-s\tau_n}$$

construct the correlation function

$$T(s) = \sum_{n=0}^M a_n \epsilon^{-s(\tau_M - \tau_n)}$$

τ_M = maximum delay in original signal

then

$$H(s)T(s) = F(s) \epsilon^{-s\tau_M} \sum_{n=0}^M a_n^2 + \text{coloration}$$

If the signal is reverberant on a simple length

$$H(s) = F(s) \sum_{n=0}^{\infty} a^n \epsilon^{-sn\tau}$$

then the inverse, or removal of reverberation can be done by

$$H(s)I(s) = F(s)$$

where

$$I(s) = 1 - a\epsilon^{s\tau}$$

If the reverberation includes a sign change

$$H(s) = F(s) \sum_{n=0}^{\infty} (-1)^n a^n \epsilon^{-sn\tau}$$

then

$$H(s)I(s) = F(s)$$

when

$$I(s) = 1 + a\epsilon^{-s\tau}$$

If an object has an acoustic shape, such as a cube, tetrahedron, or fish, the character is

$$H(s) = F(s) \sum_{n=0}^M a_n \epsilon^{-s\tau_n}$$

which may be called "external reverberation." The object is recognized by

$$H(s) T(s) = \text{maximum peak}$$

If the object is the same shape but variant in size, the scale factor k is introduced

$$H_2(s) = F(s) \sum_{n=0}^M a_n \epsilon^{-sk\tau_n}$$

$$C_2(s) = C_1(ks)$$

Thus the forms of the recognition transforms are identical for the same shape.

If the object is viewed at different aspects, the aspect transformation becomes

$$H_a(s) = F(s) \sum_{n=0}^M a_n \epsilon^{-s\tau_n} \epsilon^{-s(\cos\theta + \Psi_n)\tau_n}$$

θ = angle of view

If the aspect transformation is multiplied by a rotation matrix which can be inverted by a counter rotation matrix, recognition is possible from any angle. Thus size, shape, and aspect are all recognizable. Distance does not alter the characterizations.

If the object is internally reverberant, as a violin, then

$$H(s) = F(s) \sum_{n=0}^M \sum_{k=0}^{\infty} a_n^k \epsilon^{-sk\tau_n}$$

and rotation corresponds to time relationship shifts in n , but each string k remains invariant with aspect. A violin can be recognized easily at any aspect, but aspect colors the sound.

If reflecting walls are acoustically colored, then the attenuation coefficients are functions of s

$$H(s) = F(s) \sum_{n=0}^M a_n(s) \epsilon^{-s\tau_n}$$

The forms of correlations and inverses are unchanged.

If the delays are not discrete, but smeared, then the form becomes

$$\begin{aligned} H(s) &= F(s) \int_0^{\infty} a(\tau) \epsilon^{-s\tau} d\tau \\ &= F(s) a(s) \end{aligned}$$

or

$$F(s) = \frac{H(s)}{a(s)}$$

and the time domain is recovered by

$$f(t) = L^{-1} [F(s)] = \int_s \frac{H(s)}{a(s)} \epsilon^{s\tau} ds$$

which suggests the correlation transform.

If speech is characterized by delays, the recognition transform $T(s)$ does not change with differentiation of the signal

$$s^p H(s) = s^p F(s) \sum_{n=0}^M a_n \epsilon^{-s\tau_n}$$

which is recognized by

$$T(s) = s^{-q} \sum_{n=0}^M a_n e^{-s(\tau_M - \tau_n)}$$

The recognition delays $\tau_M - \tau_n$ are invariant to p and q . However, the condition $q = p$ is a coloration adjustment which might be made if p were known.

In resume, it is possible to localize, pay attention, recognize in noisy, reverberant surroundings, with cognizance of size and aspect, for sounds, objects, and environments, with some freedom from particular spectrum requirements, by correlations for maximum peakedness, and inverses for maximum bandwidth, by methods using delays, attenuations and signed additions, which appear to be the counterpart of the human system of hearing. This room sounds like this room whether from air noise, speech, traffic, or clatter; a thing is such and such by pulse, sweep, or any spanning stimulus. A violin has the sound shape of a violin. Porpoises recognize fish, bats find their food, people listen to speech. This by means simply conceived, easily realizable in the nervous system, but likely to require large numbers of elements for general use.

REFERENCES

1. "Characteristics of Human Localization of Sound," Final Report, May, 1961, prepared under Contract No. N123-(60530) 23545A, for U. S. Naval Ordnance Test Station, China Lake, California, by United Research Incorporated, Cambridge, Massachusetts.
2. "The Mechanism of Human Localization of Sounds with Applications in Remote Environments," Final Report, March, 1962, prepared under Contract No. N123-(60530) 27872A for U. S. Naval Ordnance Test Station, China Lake, California, by United Research Incorporated, Cambridge, Massachusetts.
3. "Localization of Sound: A New Theory of Human Audition," September, 1963, prepared under Contract No. N123-(60530) 30283A, for U. S. Naval Ordnance Test Station, China Lake, California, by United Research Incorporated, Cambridge, Massachusetts.

INITIAL DISTRIBUTION

- 6 Chief, Bureau of Naval Weapons
 - DLI-31 (2)
 - RRRE (1)
 - RU (1)
 - RUTO (1)
 - RUTO-2 (1)
- 3 Special Projects
 - SP 20 (1)
 - SP 22 (1)
 - SP 27 (1)
- 1 Chief, Bureau of Ships (Code 560)
- 1 Chief of Naval Operations
- 2 Chief of Naval Research
 - Code 104 (1)
 - Code 466 (1)
- 1 David W. Taylor Model Basin
- 1 Fleet Anti-Air Warfare Training Center, San Diego (Guided Missile Section)
- 1 Naval Advanced Undersea Weapons School, Key West
- 1 Naval Air Development Center, Johnsville
- 1 Naval Engineering Experiment Station, Annapolis
- 1 Naval Ordnance Laboratory, Corona
- 1 Naval Ordnance Laboratory, White Oak
- 1 Naval Postgraduate School, Monterey (Library, Technical Reports Section)
- 1 Naval Research Laboratory
- 1 Naval Torpedo Station, Keyport (Quality Evaluation Laboratory, Technical Library)
- 1 Naval Underwater Ordnance Station, Newport
- 1 Naval Underwater Weapons Systems Engineering Center, Newport
- 2 Navy Electronics Laboratory, San Diego
- 1 Navy Mine Defense Laboratory, Panama City
- 1 Navy Underwater Sound Laboratory, Fort Trumbull
- 1 Navy Underwater Sound Reference Laboratory, Orlando
- 1 Norfolk Naval Shipyard (Underwater Explosion Research Division)
- 1 Pacific Missile Range, Point Mugu
- 1 Submarine Development Group 2
- 1 Atlantic Missile Range
- 20 Defense Documentation Center (TISIA-1)
 - 1 National Aeronautics & Space Administration
 - 1 Acoustics Laboratory, Harvard University, Cambridge
 - 1 Aerojet-General Corporation, Azusa, Calif., via BWR
 - 1 Applied Physics Laboratory, JHU, Silver Spring

<p>U.S. Naval Ordnance Test Station <u>Localization of Sound, Part 4. Further Developments in Sound Localization (Final Report)</u>, by United Research, Inc. China Lake, Calif., NOTS, April 1964. 88 pp. (NOTS TP 3109, Part 4), UNCLASSIFIED.</p> <p>ABSTRACT. The theoretical considerations of the localization function utilizing time delays have been extended to other areas of acoustics to provide a comprehensive theory useful in explaining speech recognition, attention, and the importance of reverb-</p> <p style="text-align: right;">○ (Over) 1 card, 4 copies</p>	<p>U.S. Naval Ordnance Test Station <u>Localization of Sound, Part 4. Further Developments in Sound Localization (Final Report)</u>, by United Research, Inc. China Lake, Calif., NOTS, April 1964. 88 pp. (NOTS TP 3109, Part 4), UNCLASSIFIED.</p> <p>ABSTRACT. The theoretical considerations of the localization function utilizing time delays have been extended to other areas of acoustics to provide a comprehensive theory useful in explaining speech recognition, attention, and the importance of reverb-</p> <p style="text-align: right;">○ (Over) 1 card, 4 copies</p>
<p>U.S. Naval Ordnance Test Station <u>Localization of Sound, Part 4. Further Developments in Sound Localization (Final Report)</u>, by United Research, Inc. China Lake, Calif., NOTS, April 1964. 88 pp. (NOTS TP 3109, Part 4), UNCLASSIFIED.</p> <p>ABSTRACT. The theoretical considerations of the localization function utilizing time delays have been extended to other areas of acoustics to provide a comprehensive theory useful in explaining speech recognition, attention, and the importance of reverb-</p> <p style="text-align: right;">○ (Over) 1 card, 4 copies</p>	<p>U.S. Naval Ordnance Test Station <u>Localization of Sound, Part 4. Further Developments in Sound Localization (Final Report)</u>, by United Research, Inc. China Lake, Calif., NOTS, April 1964. 88 pp. (NOTS TP 3109, Part 4), UNCLASSIFIED.</p> <p>ABSTRACT. The theoretical considerations of the localization function utilizing time delays have been extended to other areas of acoustics to provide a comprehensive theory useful in explaining speech recognition, attention, and the importance of reverb-</p> <p style="text-align: right;">○ (Over) 1 card, 4 copies</p>

NOTS TP 3:09, Part 4

eration on speech intelligibility, acoustic coloration, noise masking, and other phenomena.

The knowledge of speech recognition was used in initial work to establish inter-species communication with the porpoise.

NOTS TP 3:09, Part 4

eration on speech intelligibility, acoustic coloration, noise masking, and other phenomena.

The knowledge of speech recognition was used in initial work to establish inter-species communication with the porpoise.

NOTS TP 3:09, Part 4

eration on speech intelligibility, acoustic coloration, noise masking, and other phenomena.

The knowledge of speech recognition was used in initial work to establish inter-species communication with the porpoise.

NOTS TP 3:09, Part 4

eration on speech intelligibility, acoustic coloration, noise masking, and other phenomena.

The knowledge of speech recognition was used in initial work to establish inter-species communication with the porpoise.