THE ROLE OF THE PINNA IN HUMAN LOCALIZATION

by

Dr. Dwight W. Batteau
Professor of Mechanical Engineering
Tutts University

January 1964

This work was performed at United Research Incorporated, Cambridge, Massachusetts, under the sponsorship of the United States Naval Ordnance Test Station, China Lake, California.
THE ROLE OF THE PINNA IN HUMAN LOCALIZATION

by
Dwight W. Batteau

I. INTRODUCTION

The problem of human localization of sound is an old and honored study, and the role of two ears in this performance has been dominant. The early attitudes regarding the pinna are reflected in a statement by Prof. A. Steinhauser (1):

"In its essential construction the pinna of the human ear consists of a funnel to collect the sound and a reflector."

While attention was directed to the pinna and to monaural localization in many early studies (2)(3), there was attention predominantly on intensity differences between the two ears as a mechanism of localization (1)(4)(5). Following the concern with intensities as means of sound localization, a number of investigators turned their attention to phase differences between the two ears (6)(7)(8). The importance of phase continued to appreciate and was shown to be significant beyond intensity differences between the two ears by Langmuir et al (9).

It is important to our consideration that other factors entered the study of hearing as in the case of localization in a reverberant environment (10). It was shown in the reference by Wallach et al that the first signal to arrive in a sequence of reverberant sounds appears to take precedence. In the introduction Wallach remarks "... that localization within a reverberant room is both common and useful. The problem of how this is possible remains however, unsolved." Fortunately, the theory regarding the function of the pinna provides an adequate explanation, since only the time domains differ in the result.
It is also important to our consideration that studies have been made relative to the perception of phase differences alone as sensations of pitch when binaural interaction is considered with noise. In his paper, Cramer refers to an unpublished memorandum of Huggins as the earliest thought in this direction. Since pitch sensations by phase difference alone can also be produced, as was done by the author in 1958 at Harvard University, it is necessary that the mechanisms of hearing be consistent with this observation and fortunate that the model derived for monaural localization also provides an adequate explanation. In a paper by Kock, a pertinent remark is made in the conclusion: "The above experiments suggest that the brain preserves and evaluates time delays (perhaps by the mechanism of delay insertion in one or the other of the nerve paths between the ear and the brain) . . ." It is this mechanism basically, but more complexly arranged which is assumed in the role of the pinna in localization.

II. EXPERIMENTATION

The research reported here began in 1959 with a simple experiment performed for the author by Dr. William B. McLean, which suggested that the external ears or pinnae have a role in human localization. Dr. McLean demonstrated that distortion of the pinnae by bending distorts the perception of the locale of a sound.

It is immediately evident that the perception of locale and attention to a sound source requires a transformation relative to the arriving wave front, since acoustic information arrives at the ear from the impinging sound front. Since a transformation of the incoming sound front is necessary to localization, it seemed reasonable to assume that the external ear served as the acoustical device to perform the requisite transformation. Subsequent work demonstrated that it indeed introduces a significant transformation. Further experimentation showed that the perception of all locales of a sound
position (front, back, up, down, distance) could be provided with a high fidelity electronic system in which the microphones were inserted in casts of human pinnae.

A search of the literature brought to light one paper which was concerned with monaural localization, with experimental evidence that it was quite good. With this encouragement, we conducted experiments with monaural and binaural localization which produced the data presented in Figure 1 (reproduced from reference (15)). Subsequently it was found that monaural localization by persons totally deaf in one ear is commonplace.

It is obvious that there are differences in the exact details of individual pinnae, and the results obtained with a group of four colleagues of varying ages from 25 years to 38 years show a variation in performance with the electronic system of pinnae microphones and headphones. The polar graph shown in Figure 2 presents the results. It is an interesting footnote that the poorest performance was obtained from a man who is color blind. The curves represent averages of 36 trials at each indicated position with a random selection of position. Randomizing was obtained by filling a box with the total test figures on individual slips of paper and drawing them one at a time. The experimental details will be presented in a forthcoming paper by Mr. Roland L. Plante and Mr. William E. Lyle.

Since there are differences in pinnae, a test was undertaken to indicate the rate of adaptation to the pinnae used. It is interesting to observe in Figure 3 that two well defined groups were formed by the test subjects. In order to prevent room acoustics from influencing the performance, the pickup was reoriented during the test; this did not seem to affect the trends development. Each subject was required to make 288 responses, evenly distributed and randomly selected in azimuth. One may conclude that other people's pinnae can be used.
Figure 1  Localization Histograms for Monaural and Binaural Comparison
In order to obtain information concerning the transformation performed by the pinna, an ear cast and microphone were used to provide oscillograms of the response to a pulse. The form of the pulse as recorded by a bare microphone is shown in Figure 4a and the oscillogram taken through the pinna is shown in Figure 4b. The appearance of the oscillograms is much more suggestive of sets of discrete delays than of a smooth continuum, although consideration is provided in the theory for both.

Figure 2 Azimuth Localization Ability with Electronic Coupling for Four Subjects
Figure 3  The Effect of Experience with Electronic Coupling for Six Subjects
By modeling the ear five times normal size and placing a microphone in it, Mr. Plante and Mr. Lyle were able to obtain a measurement for delays which vary with azimuth and also for those which vary with elevation angle. The results of these measurements are presented in Figure 5, reproduced from reference (15).

### III. DEVELOPMENT OF THEORY

When it is considered that attention is one of the prime functions of localization, a new factor is introduced into the problem. It is possible by time correlation to enhance the signals of a particular interaural time difference, and to select bilaterally from one ear or the other by intensity differences, but the relatively acute localization of attention by the healthy young is not accounted for in these models. The theory developed for the role of the pinna, however, does account for attention.
Figure 5 Time Delay vs. Orientation
The working hypothesis regarding the pinna was very simple. It assumed that the various paths of the sound to the ear canal involved multiple delays due to path differences. The delays introduced were easily seen to be different with the aspect of the sound source to the pinna. In order to examine the significance of delay and orientation, a simple diagram of a system reduced to one delay is shown in Figure 6.

![Diagram of a system reduced to one delay](image)

Figure 6 A Directional Coupler in a Plane
In the simple model shown the pulse separation as a function of angle is given by the following equation:

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

$\tau$ = delay between first and second arrivals at the probe

$s$ = distance between holes in the coupler

$v$ = velocity of sound in the medium

$A$ = angle of normal to wave from $t$ with line between holes: range $0^\circ - 180^\circ$

In this simple model, a means of measuring the time between first and second arrivals would provide the direction of the incident wave. If a spanning set of such measurements is desired, a spanning set of holes can be used, with domain definition for each coordinate. For example, three holes in a plane, arranged in a rectangular coordinate array, with a fixed delay between one of them and the probe provides a domain separation, as given in the following pair of equations:

$$\tau_1 = \frac{s_1}{v} (1 + \cos A_1)$$

$$\tau_2 = \frac{ks_1}{v} + \frac{s_2}{v} (1 + \cos A_2)$$

$s_1$ = spacing between reference and azimuth holes

$s_2$ = spacing between reference and altitude hole

$A_1$ = azimuth angle

$A_2$ = elevation angle

$k \geq 2$
A sketch of the three hole coupler is shown in Figure 7.

When the consideration of range or curvature sensing is added to this system, it is possible to conceive of displaced systems of coordinates providing the sensing of curvature by relative delays. An evolution of a system is shown in the following sketches, Figure 8.

While the resultant geometries are provocatively like the structures of the pinna, these should be taken only as suggestive and not as defined modeling. While the theoretical base appears to be sound, many different geometries are possible providing delay transformations, as may be seen by examination of the base of the ear of dog, cat, mouse, bat, and other mammals.

With the proposition that the external ear acts as a reflector and diffractor producing a variety of delays, the function of attention could be provided if there were a particular inverse to a given transformation due to a particular aspect, so that only that one sound source were properly to be reconstituted. In a simple system of one delay, the inverse can be computed as follows:
Figure 8  Evolution of a Diffraction System Providing Localizing Delays in Azimuth, Altitude and Range
\[ f(t) = \text{original sound at source} \]

\[ f(t) + af(t-\tau) = \text{sound at eardrum due to } f(t) \text{ coming through the simple delay system} \]

\[ \tau = \text{time delay in reflection path} \]

\[ a = \text{attenuation factor at path reflection} \]

Construct the series

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

\[
= \sum_{n=0}^{M} (-1)^{n} a^{n} f(t-n\tau) + \sum_{n=0}^{M} (-1)^{n+1} a^{n+1} f(t-\tau-n\tau)
\]

\[
= f(t) + \sum_{n=1}^{M} (-1)^{n} a^{n} f(t-n\tau)
\]

\[
+ \sum_{n=1}^{M+1} (-1)^{n-1} a^{n} f(t-n\tau)
\]

\[
= f(t) + \sum_{n=1}^{M} (-1)^{n} a^{n} f(t-n\tau) - \sum_{n=1}^{M} (-1)^{n} a^{n} f(t-n\tau)
\]

\[
+ (-1)^{M+1} a^{M+1} f(t-\tau-M\tau)
\]

\[
= f(t) + (-1)^{M+1} a^{M+1} f(t-\tau-M\tau)
\]

\[
= h(t)
\]

\[
\lim_{M \to \infty} h(t) = f(t) \quad a < 1
\]
A second method of computing the inverse includes the provocative series of octave delays, i.e., $\tau, 2\tau, 4\tau, \ldots, 2^n\tau$, which suggest a basis for octave identity in music.

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

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

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

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

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

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

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

And

$$\left[a^{2^{n-2}}\right]^2 = a^{2^{n-1}}$$

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

\[
\lim_{n \to \infty} P_n = f(t), \quad a < 1
\]

In extending the simple one delay system to the role of the pinna, we must consider that a spanning set of delays for localization in three space has a minimum of four linearly independent components, and this dimensionality must be preserved by the pinna, requiring a minimum of four independent paths for the sound to reach the eardrum. It is convenient at this point to introduce the Laplace transform notation. The signal in this representation is

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

\[ H(s) = \text{sound heard at eardrum} \]

\[ F(s) = \text{source sound} \]

for a spanning set. However, we might generalize

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

The inverse for this in transform representation becomes

\[
T^{-1}(s) = \sum_{n=0}^{\infty} (-1)^n \left[ \sum_{k=0}^{N} a_k e^{-s \tau_k} \right]^n
\]

The time domain realization of \( T^{-1}(s) \) is then a series of delays with particular attenuation factors. Obviously an approximation

\[
T^{-1}(s) \approx \sum_{n=0}^{L} (-1)^n \left[ \sum_{k=0}^{N} a_k e^{-s \tau_k} \right]^n
\]

converges to some tolerance when the reflection factors are less than unity.
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:

\[ \begin{align*}
F(s) &= \text{sound at source} \\
R(s) &= \text{sound at perceiver as modified by the room} \\
R(s) &= F(s) \sum_{n=0}^{\infty} a_n e^{-s \tau_n}
\end{align*} \]

If we consider first the effect of two parallel walls of infinite extent as shown in Figure 9, the equation can be simplified

![Figure 9: Reverberation from Two Infinite Parallel Walls](image)
\[ R(s) = F(s) \left[ a_0 e^{-s\tau_0} + a_1 e^{-s\tau_1} + a_2 e^{-s\tau_2} \right] \]

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

\[ \tau_M = \text{maximum delay} \]

\[ R(s) = F(s) \sum_{n=0}^{M} a_n e^{-s\tau_n} \]

and the neural transformation \( T(s) \) constructed

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

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

\[ P(s) = R(s) T(s) \]

\[ = F(s) \sum_{j=0}^{M} a_j e^{-s\tau_j} \sum_{k=0}^{M} a_k e^{-s(\tau_M - \tau_k)} \]

\[ = F(s) e^{-s\tau_M} \sum_{j=0}^{M} a_j e^{-s\tau_j} \sum_{k=0}^{M} a_k e^{+s\tau_k} \]

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

\[ \sum_{j=0}^{M} a_j e^{-s\tau_j} \sum_{k=0}^{M} a_k e^{+s\tau_k} = \sum_{n=0}^{M} a_n^2 + \text{cross terms} \]
cross terms = \sum_{k=0}^{M} \sum_{j=0}^{M} a_k a_j e^{-\tau_k} e^{+\tau_j} \quad j \neq k

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

\[ \sum_{x} a_x (e^{-\tau_x} e^{+\tau_x}) \]

\[ \tau_x = \tau_k - \tau_j \quad j \neq k \]

\[ a_x = a_k a_j = a_j a_k \]

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

\[ P(s) = F(s) e^{-s\tau} M \left[ \sum_{n=0}^{M} a_n^2 + \text{coloration} \right] \]

The perceived signal would then be increased in amplitude by

\[ \sum_{n=0}^{M} a_n^2 \]

and colored by terms of the form

\[ a_j a_k (e^{-sx} e^{+sx}) \]

none of which could alter the result by more than

\[ \pm a_j a_k \]
For example, let

\[
\begin{align*}
    a_0 &= 1 \\
    a_1 &= .9 \\
    a_2 &= .8 \\
    a_3 &= .7 \\
    a_4 &= .6
\end{align*}
\]

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 \( \tau_M \)
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.
IV. APPLICABILITY TO NERVOUS SYSTEM

In order to examine the construction of inverses, or attention functions, it is again convenient to use a single fixed delay system. If it is assumed that a characteristic delay \( \tau \) has been inserted by the pinna, a possible nerve system constructable at the basilar membrane, based on the first mathematical model is as shown in Figure 10. This model assumes that the cochlea provides a delay line, with the basilar membrane carrying the delay connections.

![Diagram of theoretical nerve end and synaptic distribution at the basilar membrane as a computational system](image)

Figure 10 Theoretical Nerve End and Synaptic Distribution at the Basilar Membrane as a Computational System

It should be noticed that the attenuation constants \( a^n \) behave as if there were constant attenuation per unit delay, providing a time-intensity equivalence in the network.

If we use the second mathematical model, it is necessary to introduce nerve delays, and computation may be performed en route from the nerve.
endings at the basilar membrane to the auditory center. Figure 11 shows a model of this network.

![Diagram](image)

Figure 11 Computation of Inverse to Pinna Transform by Octaves

One of the provocative outcomes of this research has been the construction of mathematical models which appear to be realizable in the human nervous system by rather simple means.

When the less simplified view of the pinna as a continuous system of delays is taken, it becomes convenient to assume a continuous distribution of delays between 0 and some maximum $\tau_{\text{max}}$, and associated with each a reflection factor, $a(\tau)$ d$\tau$. 

\[ f(t) = \sum a(\tau) f(t - \tau) \]
The expression for the complete transformation provided by the pinna may be given by summing over all the delays, giving

\[ T_{\text{max}} \]
\[ \int_0^\infty a(\tau) e^{-s\tau} d\tau = \int_0^\infty a(\tau) e^{-s\tau} d\tau = a(s) \]

\[ a(\tau) = 0 \quad : \quad \tau > \tau_{\text{max}} \]

If the incoming signal is presented in its Laplace transform notation \( F(s) \), the sound reaching the eardrum becomes

\[ F(s) A(s) = H(s) \]

\( H(s) \) = sound transform at ear drum

The inverse function to be performed by mental computation, provides

\[ F(s) = \frac{H(s)}{A(s)} \]

The continuous inverse may be approximated to any degree of tolerance by either of the computational models already presented.

The three mathematical models already given, two discrete and one continuous, provide means of constructing functions on hearing. In the consideration of realizing these computations, a fourth model was conceived having considerable provocation. Again using the Laplace transform notation, a delay feedback system is constructed as shown in Figure 12.

\[ \text{Figure 12 A Feedback Delay Computation} \]
With this system, let

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

be the signal at the eardrum, being reflected and delayed by \( \tau \). The error signal is given

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

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

\[ E(s) = F(s) \]

With this model, a system was constructed to recognize and remove a "reverberation" of delay \( \tau \) and operated as predicted, Figure 13.

When a system of delays is introduced

\[ H(s) = \sum_{n=0}^{3} a_n e^{-s\tau_n} F(s) \]

\( \tau_0 = 0 \)

\( a_0 = 1 \)

Figure 13  Removal of Reverberation

a. Signal with one echo
b. Signal with echo removed
The requisite network is shown in Figure 14.

![Figure 14](image)

Figure 14  A Feedback Computation Inverting Three Delays

This delay network model cannot apply to signals at the cochlea, since the feedback must operate on the signal. It can, however, be constructed in any subsequent trunk.

V. EXTENSIONS OF THEORY

In considering the significance of the pinna, attention, and computation, it seemed useful to extend the model of inverse computation to include other domains. Six domains are identified tentatively:

- **Pinna**: 2μ sec to 300μ sec
- **Interaural**: 0μ sec to 800μ sec
- **Speech**: 0.6 millisec to 2.6 millisec
- **Shape**: 2.6 millisec to 10 millisec
- **Reverberation**: 10 millisec to 40 millisec
- **Memory**: 0 to ∞

These domains certainly may overlap, but the pinna domain can be synthesized by pulse techniques and sensations of motion obtained, so that it may be called a "where" domain. The same is true and well known for the interaural
domain which may also be called "where." The speech and shape domains include the lengths characteristic of the vocal tract and musical instruments and may be called "what." The reverberation domain, of course, includes the system of reflections from walls and objects in the environment and may be called "place." The memory domain includes all delays as well as factors of long delay used in personal and environmental recognition. A paraphrase could be made regarding the domains as the recognition of "something" (what), "somewhere" (localization) in "someplace" (environmental) as having some significance (memory).

When "where" and "what" are considered, the significance of the median plane becomes apparent. Sound sources lying in the median plane have identical pinnal transforms on the left and right, so that no separation is provided between "where" and "what." Away from the median plane, each ear transforms differently, so that "where" is different for each ear, but "what" is the same, providing distinguishability.

The characterization of reverberation, in particular, has the identical form of the expressions given for the pinna, so that mental inversion of the resultant transform can provide attention dependent upon location within the environment, the delay characteristics being unique from point to point. Simple tests have shown two significant facts: (1) in monaural subjective tests, reverberation is reduced by a pinna compared with no pinna, (2) subjective loudness is raised in a reverberant environment for the same source acoustic power (17). The same mathematical models apply to these processes as apply to the pinna transformation.

A theoretical process including the role of the pinna in normal hearing can be stated as follows:

(1) The incoming sound is characteristic of its source, transformed by the environment.
(2) The incoming sound is transformed by the pinna and each
direction of arrival has a characteristic.

(3) The transformations pertinent to the source to which attention
is to be paid are inverted for all directions of arrival, giving a set of
environmental transformations.

(4) Each environmental transformation for the sound to which
attention is paid is inverted and the resultant set of elements combined.

(5) The same process is accomplished by each channel (right and
left ears) and the results combined.

(6) The resultant signal is inverted to the most acute stimulus and
interpreted or recognized by the requisite transformation used to accomplish
the inverse.

When combination is mentioned, it connotes identity in signal except for
time shift, so that simple delay and addition are sufficient. Since the inverse
of the environmental transforms provides a combining set, a precedence effect
is a simple consequence.

The neural networks necessary to accomplish this process require
delays, attenuations, and signed additions. It is apparent that tone recognition,
both pure tones and colored noise, can be perceived in this manner, and it is
theorized that no mechanical resonances of the basilar membrane or other
acoustic structures are significant in the process.

SUMMARY

It may be said in summary that, in theory, the role of the pinna in
localization is to introduce, by means of delay paths, a transformation of the
incoming signal which is mentally inverted to provide attention, and that the
inverse transform required defines the location of the sound source. It may be
further shown that relatively simple systems of delays, attenuations, and
signed additions may be used to construct the inverse transformations, and that these could easily be realized in the nervous system. It may be further theorized that the same method of constructing inverse transformations can apply to monaural and binaural localization, sound recognition, and the utilization of reverberation.
REFERENCES


*Quoted by Thompson
REFERENCES (Cont.)

13. Ears in Action, a motion picture produced by the U. S. Naval Ordnance Test Station, China Lake, California.


17. "Intelligibility Experiments in Reverberation," by Terence M. McNay, L. E. Stiles, Jr., Earle Yaffa, Vernon E. Ware, Arial George and Robert E. Baker, former students of Dr. D. W. Batteau in mechanical engineering at Tufts University, Medford, Massachusetts.