The Apple and the Ear, Grasping Sounds in Space—A Theory of Sound Localization

Michelangelo Rossetto

New York, NY, USA

Correspondence to: Michelangelo Rossetto, marossetto@libero.it

ABSTRACT

Sensing the direction of origin of a sound in space has long been attributed to the delay between arrival times between the two ears. This, now discredited two dimensional theory, was put to rest by the observation that a person deaf in one ear can locate sounds in three dimensional space. We present here a new theory of sound localization that has the required three dimensional measurement. It is a theory that interprets the well researched biological structure of the mammalian cochlea in a new and logical way, which leads to a deeper understanding of how sound localization functions.

1. INTRODUCTION

The localization of sounds in space has long been explained by the difference in arrival times in the two ears. This is a reasonable, but overly simplistic, explanation. A review of the current thinking on arrival time localization by Li, X. discusses issues of accuracy and utility [1]. The two ears provide, at most, a two dimensional localization of a sound in space. A two dimensional measurement can only locate a sound source in a single plane, such as the horizontal plane. It is well known, and obvious to the casual observer, that mammals can locate sounds in three dimensional space [2]. In order to achieve an understanding of this ability of mammals to locate a sound in three dimensional space, we must look for mechanism that has three elements, rather than the two elements of the opposing ears.

The nail in the coffin of the two ear theory of sound localization is the observation that profound deafness in one ear does not result in the loss of the ability to locate sounds in three dimensional space [3, 4]. This ability to locate a sound in three dimensional space with one ear can be easily experienced with a simple maneuver. Just disable one ear temporarily with an ear plug, or with a finger, and listen to a sound a moderate distance away. In my experience, it is easy to locate sounds with one ear.

The localization of sounds in space has an obvious value to a free ranging animals' hunting and predator avoidance behavior, just as the ability to grasp an apple and pluck it from a tree is a necessary skill for Eve to fulfill her role.

This note equates these two behaviors, the grasping of an apple and the grasping the origin of a sound

in space, and suggests a parallel in the mechanisms that underlie both of these behaviors. It takes three fingers to grasp an apple. Similarly it takes three independent acoustic measurements to grasp the origin of a sound. This is the task the cochlea is set up to perform (Figure 1).
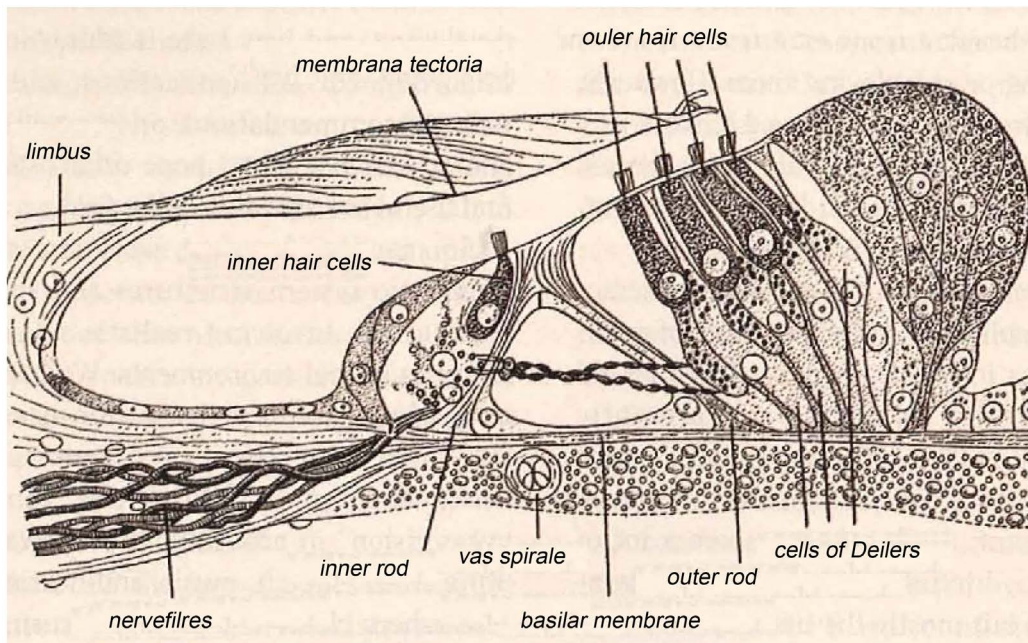

Figure 1. The organ of Corti, shown here as drawn in ref [5] by the anatomist Gustaf Retzius circa 1884, illustrates how the inner hair cells are poised to respond to the composite signal generated by the rows of outer hair cells. While most of the cochlea has three rows of hair cells, Retzius must have worked with the apical, low frequency, end of the cochlea which often expands to four, or Evan five rows, of outer hair cells (Figure 2).
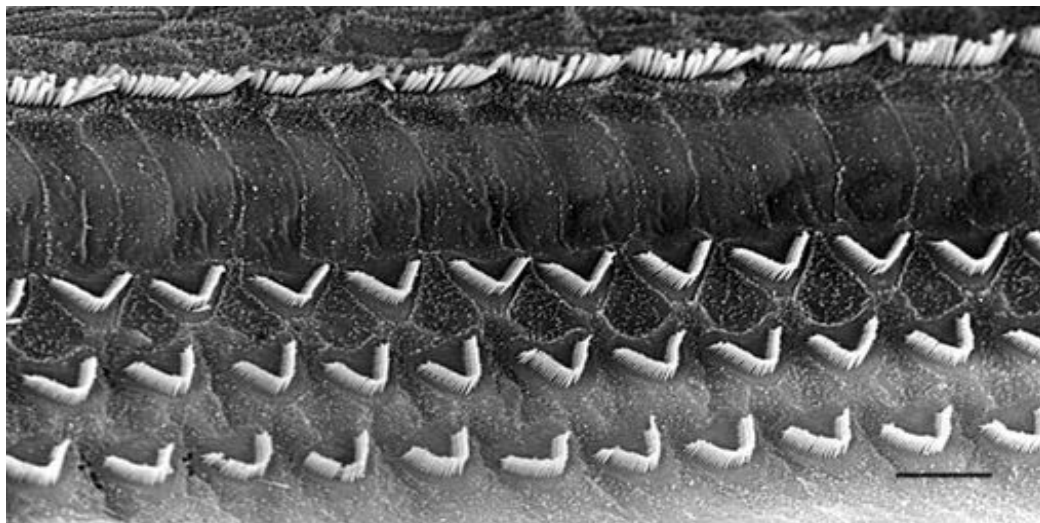

Figure 2. A typical view of the cochlea, Selected from hundreds of hair cell images on "wikipedia/cochlea' showing the three rows of outer hair cells that are overlooked by the single row of inner hair cells." Scale bar = 10 u.

2. ANALYSIS

Looking at the cochlea we see that there are three rows of outer hair cells which send out no afferent signals. Each row of outer hair cells has an independent system of efferent enervation [6]. Overlooking these three rows of outer hair cells is a row of inner hair cells that send afferents to higher centers. The vibrations that the inner hair cell sees is a composite of three simultaneous components of the traveling wave that passes down the three rows of outer hair cells. Each row of outer hair cells can modify its propagation velocity under the control of that rows efferent enervation. The efferent enervation controls the membrane potential of the outer hair cell. The membrane potential controls the lengthening and shortening of the outer hair cell thus stiffening or loosening the loading of that row which will effect the propagation velocity in that row [7, 8]. We now have a mechanism for lining up three components of an acoustic wave with the outer hair cells and a way of sensing, with the inner hair cells, when the appropriate match is achieved.

The question now is "what are these three components" of the acoustic wave that travels down the cochlear membrane?

They are illustrated in Figure 3 which represents the head of a primitive mammal. The three acoustic signals are:

1) The direct reception by the ear that arrives directly from the sound source.

2) The reflection from the wet surface of the nose. A reflection from a solid surface occurs with a 180 degree phase inversion.

3) The reflection from the entrance to the oral cavity (the open mouth) which will be in phase with the arriving acoustic wave.

In many small mammals the areas on the animals face between the ear and the reflective surfaces on the snout are covered in fur. Fur keeps the animal warm, but it also is a poor sound reflector which serves to dampen unwanted sound allowing For a cleaner reflection from the wet nose, oral cavity and possibly the eye.

Let's pause here to look at the nature of these two types of reflection.

3. REFLECTION FROM A SOLID SURFACE

A solid surface cannot support the variations in pressure presented by an arriving acoustic wave. In order to meet the boundary condition of no pressure changes the arriving wave must be canceled by a
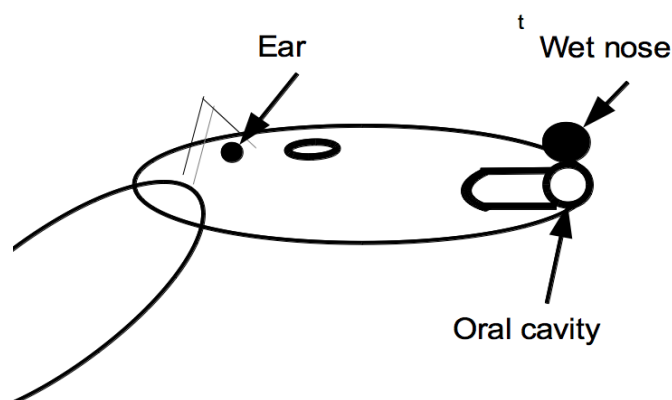

Figure 3. Figure of a primitive mammal showing the three sources of acoustic signals. 1) Direct reception to the ear. 2) Reflection with phase inversion from the wet nose surface. 3) Reflection without phase inversion from the entrance of the oral cavity.

wave leaving the surface 180 deg out of phase with it.

4. REFLECTION FROM THE OPEN END OF AN ACOUSTIC CAVITY

Reflection from the open end of an acoustic 1/4 wave cavity (think of blowing across the open end of a beer bottle) is a little more complicated. The arriving signal hits the open end and travels to the bottom of the cavity. If the cavity is a 1/4 wavelength of the incoming vibration there will be a 90 deg phase shift when it reaches the cavity bottom. The wave then reflects off the bottom, which is a solid surface, with a 180 deg phase shift. This reflected wave is shifted by 90 deg by the trip back to the open end of the cavity. All these phase shifts add up to 360 degrees. The emerging wave (the reflected wave) is in phase with the incident wave, but delayed by 360 degrees of the incoming wave. This delay results in the illusion that the reflection originates one wavelength in front of the opening of the oral cavity.

The ear is in effect looking with three ears spaced far enough apart to allow triangulation. Since only one ear is required to perform this triangulation we have an answer as to how people deaf in one ear can locate sounds in space.

5. DEDUCTIONS

When eve goes for that apple, signals are sent by motor neurons to the muscles that move the fingers until tactile signals say that the apple is contacted. There are three independent feedback networks, one per finger, that act simultaneously to insure that each finger provide the correct force to accomplish that fateful task of plucking the apple.

In the mammalian ear the efferent signals to the outer hair cells are sent by neurons derived from motor neurons. They are in effect "motor neurons" driving what are three acoustic fingers that are reaching out to grasp a sound in space.

Since the advent of the Cochlear Implant which elicits neural activity along the cochlea it is increasingly important to understand the fine structure of how sound vibrations travel and are processed through the organ. This paper aims to expand our understanding of the complex behavior of the cochlea.

In mammals the vibrations, which produce a traveling wave, contain temporal components from reflections originating at different parts of the head with different delays based the distance of the reflecting surface to the ear drum. The particular sound that is received at the ear drum is the sum of the original wave and the reflections from prominent reflecting surfaces on the animals head. This sound vibration now has a temporal dimension that sweeps across the extended surface of hair cell sensors in the inner ear of the animal. The sensors are now able to simultaneously sense both the original wave and its delayed reflections as the wave travels across the extended sensory surface and are able to, it is speculated, infer the direction of the sounds origin. In lower animals, such as turtles and reptiles, reflections originate on parts of the body requiring he animal to keep its' body very still while listening. In mammals, however the acoustic reflections required for sound localization originate from points on the head allowing the animal to keep his head focused on the sound while allowing the body freedom of motion.

It is the job of the auditory part of the brain to institute the necessary feedback functions to control propagation velocity of each row of outer hair cells independently. It must be able to independently adjust the propagation velocity of each row to allow it to achieve simultaneity in the part of the spectrum of interest. By keeping track of the propagation velocities that are required to achieve simultaneity, the location of an incoming sound can be determined.

Occam's Razor, {"the simplest explanation is the best"} [9], does not seem to hold for the theory presented here that piles complexity upon complexity to solve a very difficult problem. Yet looking at Figure 2 we see a clean direct design that can carry the spirit of Occam's Razor.

CONFLICTS OF INTEREST

The author declares no conflicts of interest regarding the publication of this paper.

REFERENCES

1. Li, X., Deng, Z.D., Rauchenstein, L.T. and Carison, T.J. (2016) Contributed Review: Source-Localization Algorithms and Applications Using Time of Arrival and Time Difference of Arrival Measurements. *Review of Scientific Instruments*, **87**, 041502.

2. Middlebrooks, J. and Green, D. (1991) Sound Localization by Human Listeners. *Annual Review of Psychology*, **42**, 135-129. https://doi.org/10.1146/annurev.ps.42.020191.001031

3. Wightman, F.L. and Kistler, D.J. (1997) Monaural Sound Localization Revisited. *The Journal of the Acoustical Society of America*, **101**, 1050-1063. https://doi.org/10.1121/1.418029

4. Slatterly, W. and Middlebrook, J.C. (1994) Monaural Sound Localization: Acute versus Chronic Unilateral Impairment. *Hearing Research*, **75**, 38-46.

5. Retzius, G. (1881-1884) Das Gehororgan Der Wirbethiere: Morphologishe histolotologische, Vol 2. Samson & Wallin, Stockholm.

6. Spoendlin, H. (1963) Electronmicrospic Study of the Efferent and Afferent Innervation of the Organ of Corti in the Cat. *Annals of Otology*, **72**, 660-686.

7. Ashmore, J.F. (1987) A Fast Motile Response in Guinea-Pig Outer Hair Cells: The Cellular Basis of the Cochlear Amplifier. *The Journal of Physiology*, **388**, 323-347. https://doi.org/10.1113/jphysiol.1987.sp016617

8. Ashmore, J. (2018) Outer Hair Cells and Electromotility. Cold Spring Harb Perspect Med. 4 September 2018.

9. Hoffmann, R., Minkin, V. and Carpenter, B. (1997) Ockhams Razor and Chemistry. *HYLE International Journal for Philosophy of Chemistry*, **3**, 3-28.