EECS225d Audio Signal Processing in Humans and Machines **Spring 1995**

Lecture 12: February 14, 1995

Lecturer: Prof. Ben Gold *Scribe: William Chen*

12.1 Physiology of the Ear

The first step to speech perception is to understand how the various parts of the ear works. So in this lecture, we present:

- Anatomy of the Ear: We look at the form and function of the ear. One way to divide the ear up is into the external ear, middle ear, and inner ear.
- Engineering Model: We construct a simple system level model of the ear and show how sound propagates through our model.
- System Characterization: We build upon our simple model by describing functional properties of the neurons that innervate the inner hair cells (the ear is not just a simple linear time-invariant system anymore).

Ultimately, many of the insights gained from human speech perception can lead to better models for machine speech recognition. For example, knowing how the human ear extracts the useful information from a speech signal can help focus our research on machine front ends.

12.2 Historical Development

In 1863, Helmholtz proposed the place concept [H54], which began to shed light on how the ear works in distinguishing among different sounds. The basilar membrane is an important sound conduit in the ear. Helmholtz proposed that different places on the basilar membrane resonate to different tones. In many ways, this place concept led to the filterbank representation that is used to model the ear today.

Von Bekesy won the Nobel Prize for furthering work on the basilar membrane. He made actual measurements of the basilar motion on cadavers and verified the place concept [B60].

Kiang made the next major development, creating electrodes that were fine enough to pierce individual auditory nerves [K65]. This technology reduced the dimensions of the problem, focusing the study of the ear from large surface area measurements to single fibers.

12.3 Anatomy of the Ear

The overall auditory system has been modeled as a chain of nodes, leading from the ear up to the auditory cortex. We are only now studying the lowest member of the chain, that is the ear (refer to slide Main Auditory Pathways).

In the ear, the human auditory nerves contain about 30,000 fibers. But as you reach the auditory cortex, the auditory nerves contain close to 10,000,000 fibers. So as you can see, the complexity of the auditory system increases dramatically once you leave the periphery.

The anatomy and function of the ear is listed below. Please refer to the slides Peripheral Auditory System, Basilar Membrane, Cross-Section through the Cochlea, and Hair Cells at the end.

- Semicircular Canal - provides sense of balance.
- Auditory Canal - acts like an acoustic tube. An acoustic wave that is incident to the ear will travel down the auditory canal. In fact, the length of the canal roughly corresponds to the first mode of a 3kHz acoustic wave, so there is some amplification at this frequency.
- Eardrum- Hammer- Anvil- Stirrup - converts acoustic wave into mechanical vibrations. The incoming acoustic wave vibrates the stapes. Why does such a delicate bone structure exist? There is an impedance mismatch between the eardrum and the fluid of the inner ear. The bones compensate for this to reduce energy loss.
- Round and Oval Window - converts mechanical vibrations into traveling waves in fluid. The vibrating stapes shake the oval membrane, which presses upon the fluid of the inner ear. The resulting motion of the fluid follows that of a traveling wave.
- Basilar Membrane - responds to traveling wave and excites the hair cells. The basilar membrane is a spiralling membrane that is submerged in the fluid of the inner ear. It responds to the traveling wave by vibrating. The traveling wave, however, does not have an equal effect down the length of the basilar membrane. The basilar membrane follows a gradient of stiffness and is more stiff at the base than at the apex. The result is that the base has a greater response to the higher frequency components of the traveling wave, while the apex has a greater response to the lower frequency components. With such a finite amount of real estate, division of the basilar membrane into frequency components is an important concept.
- Hair Cell - produces a receptor potential. The hair cells are sandwiched between the basilar and tectorial membranes. As the basilar membrane vibrates, it creates a shearing effect on the hairs of the cell, which sets off a series of reactions. First, this opens the channels in the hair cells. Once open, the hair cells depolarize and release neurotransmitters. The neurotransmitters, in turn, depolarize the auditory fibers innervating the cell.
- Neurons - produces and propagates an action potential. The neurons surround each of the hair cells and responds to the receptor potential by transmitting an action potential down a network of neurons.

12.4 Engineering Model

Modeling the ear can teach us how the human ear conditions the acoustic signal for speech perception. A very general black box model of the ear is shown in Figure 12.1 and a more detailed model is shown the slide The Auditory System.

Putting it all together, sound travels down the acoustic tube and vibrates a mechanical system. The mechanical system starts the fluid motion in the inner ear. Since the basilar membrane follows a gradient of stiffness at a ratio of 100:1 (base:apex), the base of the basilar membrane reacts more to higher frequencies than the apex. In a way, the high frequency components of the fluid motion are lost down the length of the basilar membrane. So the fluid motion can be modeled as passing through a serial filterbank; the center frequency of each subsequent filter is shifted down and the frequency range is reduced. Out of each filterbank, a resonator mimics the basilar membrane motion and excites the auditory nerve.

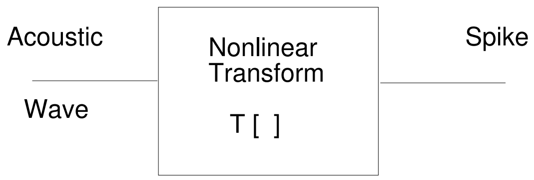


Figure 12.1: Black Box Model

12.5 System Characterization

The following lists serveral properties of the firing patterns of the neurons that surround the hair cells (refer to the slides Functional Properties of afferent Fibers that Innervate Inner Hair Cells, Min. Sound Pressure for Spontaneous Firing, Adaptation, Tuning Curves, and Synchrony).

- Spontaneous Spike Generation - the neurons can fire even without excitation. Measurements of these spontaneous firings tell us that the hair cells naturally divide up according to a low, middle, and high threshold of firing. Since the range of hearing is approximately 100dB and the dynamic range of a hair cell is 20dB, different nerves from our three groups must be used to cover the entire range of human hearing in addition to compression and adaptation.
- Adaptation - the ear is a time varying system. The auditory nerve discharge rate decreases over time in response to a constant stimulus.
- Tuning - the tuning curves verify the place concept.
- Synchrony - if a continuous tone is applied to the ear, the spikes bundle up near the peaks. This gives the auditory system a neural phase locking mechanism; ie both time and frequency information.

Next time: We discuss how the auditory system takes the information presented to it by the ear and integrates it so that we can distinguish foreground sound from background noise.

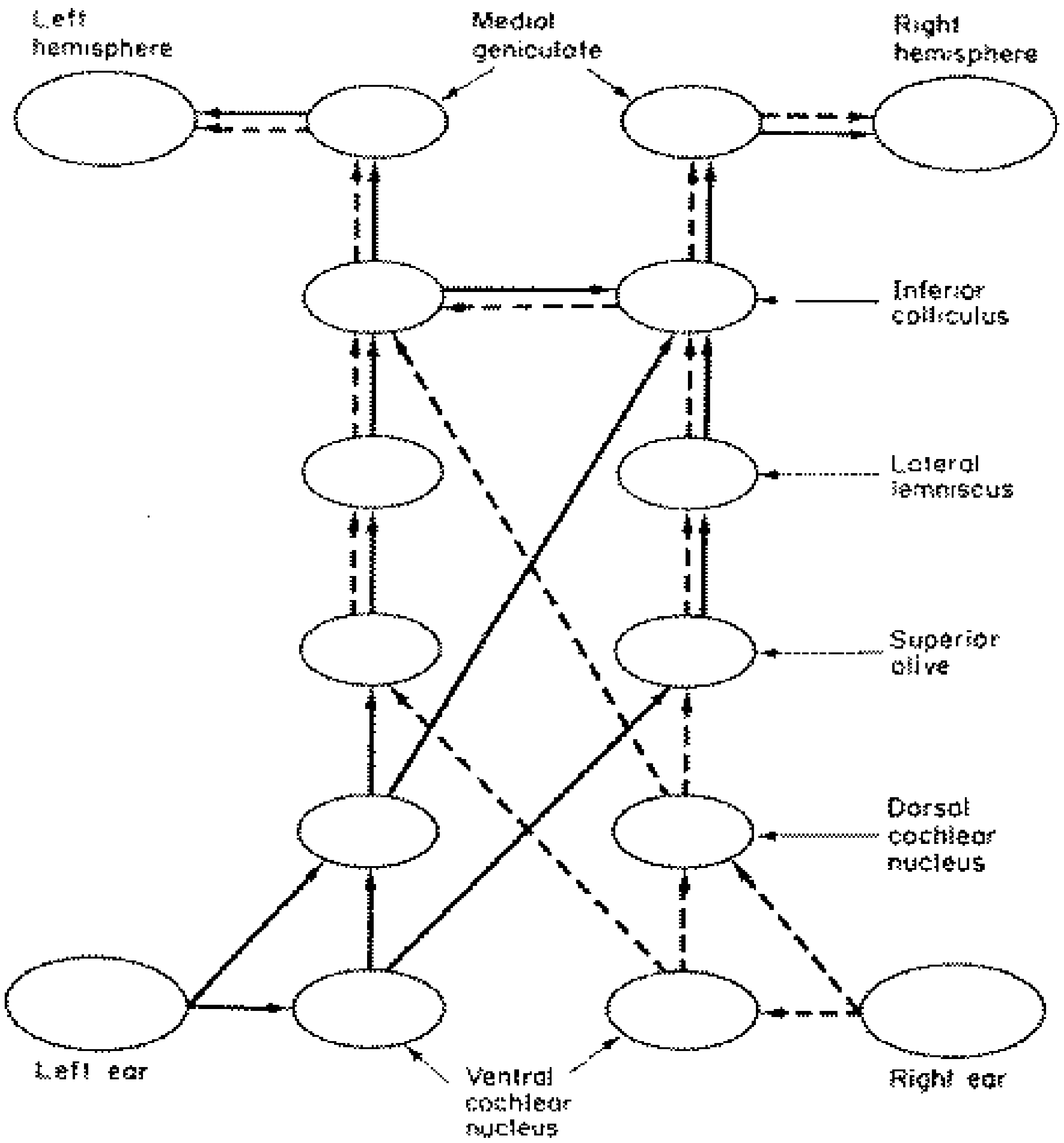
References

[B60] von Bekesy, **Experiments in Hearing**, McGraw Hill. 1960.

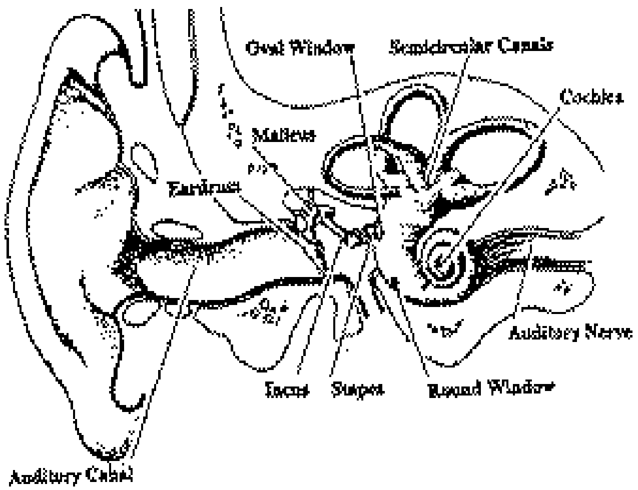
[H54] Helmholtz, **On the Sensations of Tone**, Dover. 1954.

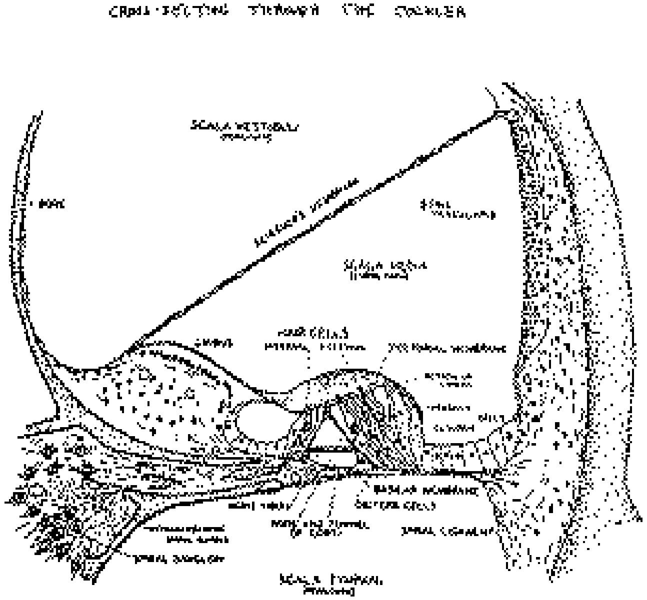
[K65] Kiang, **Discharge Patterns of Single Fibers in a Cat's Auditory Nerve**, MIT Press. 1965.

MAIN AUDITORY PATHWAYS (Moore 1982)

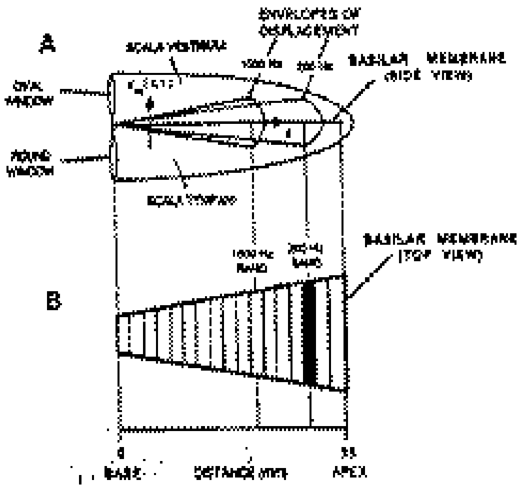


PERIPHERAL AUDITORY SYSTEM (MOORE 1982)

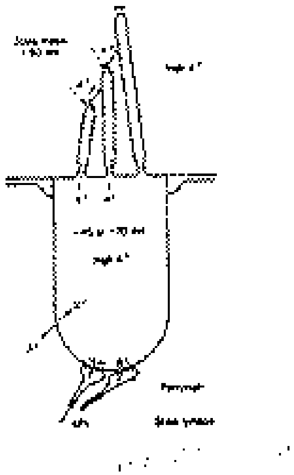




BASILAR MEMBRANE (FRISHKOPF 1989)



HAIR CELL (PICKLES)



IONIC CHANNELS IN STEREOCILIA OPEN WHEN BASILAR MEMBRANE VIBRATES.

POTASSIUM IONS ENTER CELL, CAUSING DEPOLARIZATION (VOLTAGE INCREASE).

DEPOLARIZATION OPENS CALCIUM CHANNELS NEAR BASE OF CELL.

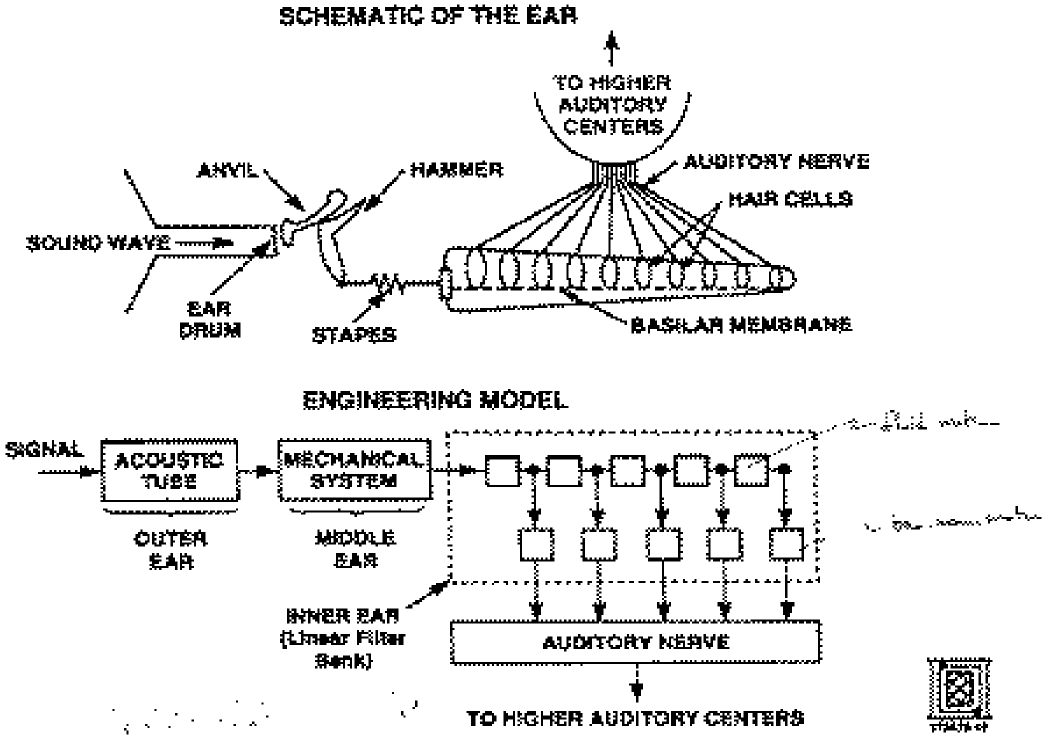
INFLUX OF CALCIUM RELEASES (UNKNOWN) NEUROTRANSMITTER.

NEUROTRANSMITTER OPENS CHANNELS IN AFFERENT NERVE FIBER.

SODIUM FLOWS INTO FIBER, DEPOLARIZING FIBER.

DEPOLARIZATION OF FIBER CAUSES ACTION POTENTIAL (SPIKE).

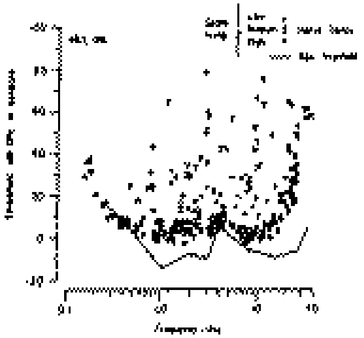
THE AUDITORY SYSTEM



FUNCTIONAL PROPERTIES OF AFFERENT FIBERS THAT INNERVATE INNER HAIR CELLS

- SPONTANEOUS SPIKE GENERATION OF A STOCHASTIC NATURE
- ADAPTATION
- TUNING
- SYNCHRONY
- NON-LINEARITIES
  a) SATURATION
  b) TWO-TONE SUPPRESSION
  c) COMBINATION TONES

[illegible] SOUND PRESSURE LEVEL NEEDED TO EXCITE FIBERS ABOVE SPONTANEOUS



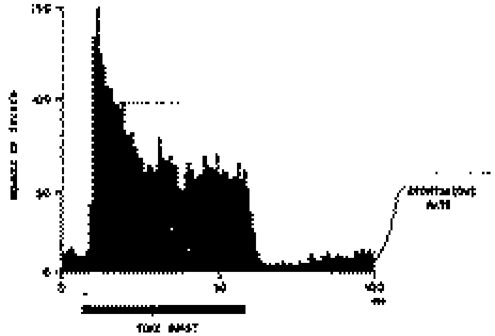
TUNING CURVES OF SIX AUDITORY NERVE FIBERS (KIANG & MOXON 1974)

METHOD

1. APPLY 50 MSEC. TONE BURSTS EVERY 100 MSEC.
2. INCREASE SOUND LEVEL UNTIL DISCHARGE RATE INCREASED BY 1 SPIKE/SEC.
3. REPEAT FOR ALL FREQUENCIES

ADAPTATION

POST STIMULUS TIME HISTOGRAM FROM AN AFFERENT AUDITORY NERVE INNERVATING AN INNER HAIR CELL (KIANG 1965)



POST STIMULUS TIME HISTOGRAM

1. APPLY TONE TO EAR MANY TIMES
2. PLOT HISTOGRAM OF TIME OF SPIKE REFERRED TO INITIATION OF TONEBURST

SYNCHRONY (JOHNSON 1974)
INTERVAL HISTOGRAMS FOR RESPONSES TO TONES