IRCAM summer stage 1980
Course notes

AUDITORY BIOPHYSICS AND PHYSIOLOGY

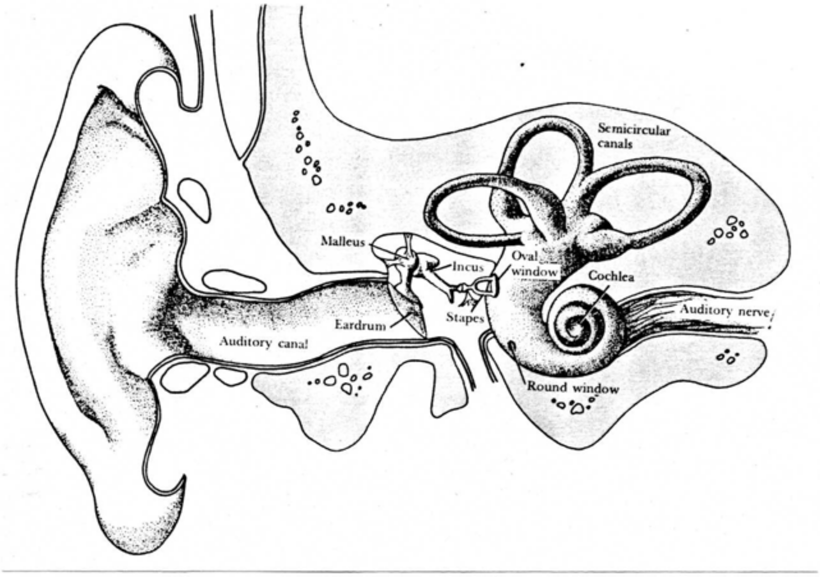
Stephen McADAMS.

Neuropsychology of the auditory system studies the relationship between acoustic vibrations and perception, as well as the method by which the transformation occurs.

It thus studies the processes of :

1) transduction of acoustic information into neural information,
2) transmission and processing of neural information which results in
3) the extraction, interpolation and identification of acoustic events in the environment.

In all cases, the physiological problem of understanding the processing that occurs is one of understanding the coding, transformation and decoding of acoustic information by the nervous system.



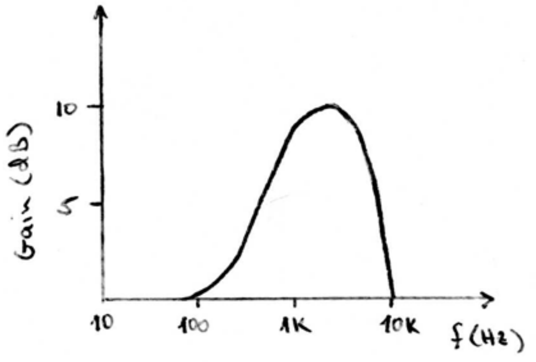
(Figure 1.)

REGION	OUTER EAR	MIDDLE EAR	INNER EAR	BRAINSTEM	CORTEX
ANATOMY	Pinna Ear Canal	Ear Drum (Tympanic membrane) Ossicles (Ear bones)	Cochlea Organ of Corti Hair cells Auditory Nerve	Cochlear Nucleus Superior Olive Lateral Lemniscus Inferior Colliculus Medial Geniculate	Auditory Cortex (temporal lobe)
FUNCTION	Protection Amplification	Amplification Impedance matching	Transduction Preliminary processing Transmission to brain	Neural Processing	Neural Processing Storage (?)
TYPE OF INFORMATION	Air vibration	Mechanical vibration of Bone	Mechanical vibration of fluid and Electro-chemical (analog and digital)	Electro-chemical (analog and digital)	Electro-chemical (analog and digital)

OUTER EAR

Protection from foreign elements (like bugs & insects).

Slight amplification : cavities in pinna and ear canal act as resonance chambers, which provide amplification between approximately 1KHz and 8 KHz.



(Figure 2.)

Localization of sound sources : by structure and positioning of pinnae.

1) Time of arrival and phase difference between the two ears.

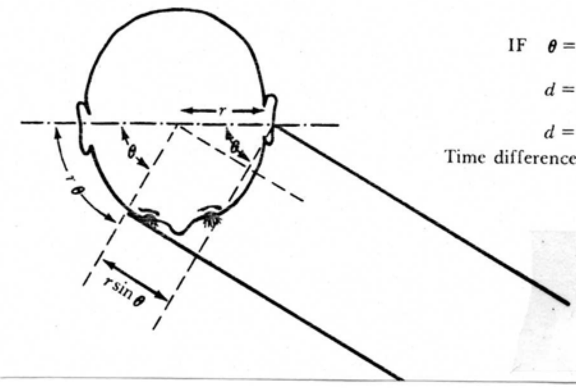
$$\boxed{d = r\theta + r\sin\theta}$$

IF $\theta = 30° = 0.52$ radians

$d = 3.5 \times .5 + 3.5 \sin 30°$

$d = 1.75 + 1.75 = 3.5$ inches

Time difference $= 76 \times d$ msec $= 270$ msec



d = difference in path length

r = radius of the head (around 3.5 inches)

θ = angle of sound source in radians

(Figure 3.)

Since the ears are separated in space and it takes time for sound to travel through space, if the sound comes from the side, it will arrive at one ear before the other. This cue works best for low frequences (< 1500 Hz)

2) Intensity differences between the ears. For higher frequencies (> 2-3 KHz) the head casts a sound "shadow" which means that sound coming from the side will be louder at the closer ear.

Frequency	Ratio of Sound Intensities at the Two Ears (15° off center)
300 Hz	1dB
1100 Hz	4dB
4200 Hz	5dB
10000 Hz	6dB
15000 Hz	10dB

3) Pinna reflections may provide information about front and back (since this would be ambiguous with both time and intensity differences) and about elevation of the sound source above the plane of the head.

MIDDLE EAR

Air-filled, bony cavity in skull, bounded by ear drum , cochlea, and the eustachian tube.

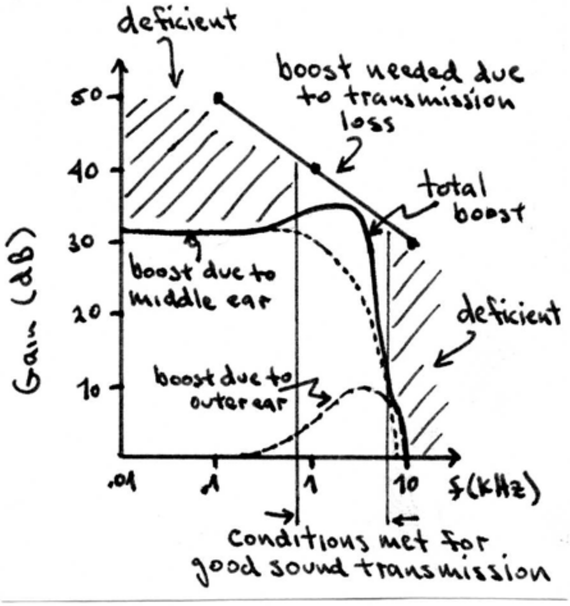
Eustachian tube: pressure equalization between middle ear cavity and the environment.

Ossicles : malleus, incus, stapes

- bones interconnected with ligaments and are suspended in middle ear cavity so they act like a series of levers.
- malleus connected to ear drum.
- incus fitted firmly to malleus.
- leg of incus shorter than leg of malleus : they are parallel and act as a lever which amplifies the force on the eardrum 1.3 times.
- leg of incus pushes on the top of stapes which has a footplate which fits into a hole (the oval window) in the bony wall of the cochlea. The stapes acts like a piston which transmits the vibration of the bones into vibrations in the fluid of the cochlea.
- surface area of ear drum effectively 14 times area of stapes footplate which causes an amplification of pressure.

Impedance matching and transformer function of middle ear :

- acoustic resistance to vibration in fluid of cochlea is 3900 times the resistance in air, which means that if air vibrations hit the oval window directly, there would be a loss in transmission which was dependent on frequency. This loss varies between about 30-50 dB over the frequency range of human hearing.



(Figure 4.)

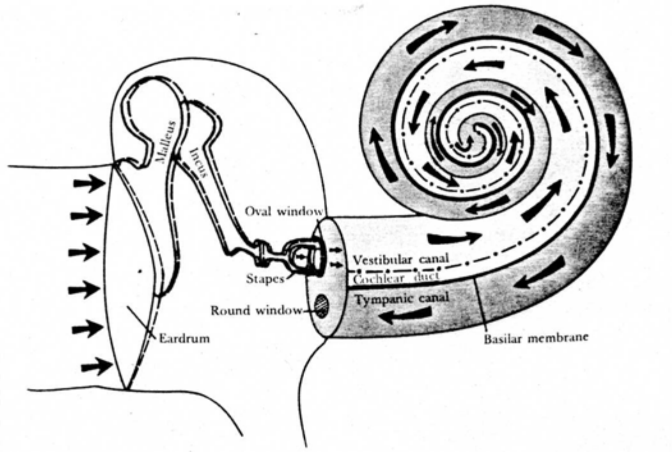
- if we consider the boost in amplitude given by the outer ear and the boost provided by the combined force and pressure amplification provided by the middle ear, we see that the transmission loss is compensated for in the mid-frequency range but there is still a substantial loss at low and high frequencies.

INNER EAR

Transduction of vibrations into languages of the nervous system.

Cochlea : bony spiral with 2,5 turns in humans.

- filled with fluid (perilymph) and contains a membranous spiral (the cochlear duct).
- the cochlear duct is filled with another fluid (endolymph) and contains the basilar membrane and organ of Corti.
- The organ of Corti is filled with a third fluid (cortilymph) which is similar to perilymph and contains the hair cells which actually transduce vibrations into electrochemical signals.



(Figure 5.)

The cochlea is divided into 3 main compartments. The top and bottom contain perilymph while the middle one (the cochlear duct) contains endolymph. All of these fluids are incompressible. The round window releases the pressure applied to the oval window by the stapes. This also allows a mechanical deformation of the cochlear duct. There is a small hole at the apex or tip of the cochlea where the two outer canals communicate.

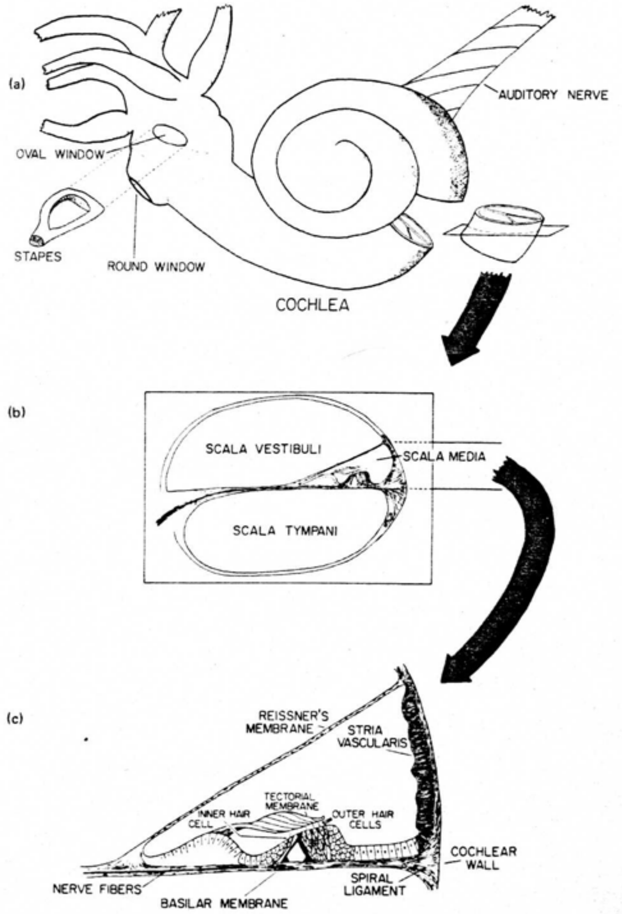


FIG. 3.2 Structural and anatomical features of the cochlea: (a) the cochlea in relation to the middle ear, the vestibular channels (partially illustrated) and the auditory nerve; (b) cross section of the cochlea; (c) the structures within the scala media.

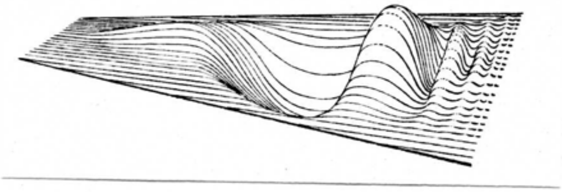
(Figure 6.)

Taking a cross-section of the cochlea we can view the structure of the organ of Corti. The main elements of interest are :

1) the basilar membrane (BM)
2) the hair cells (3 outer, 1 inner per row)
3) the tectorial membrane .

Looking at Figure 5 again the heavy line through the center of the cochlea (bottom border of the cochlear duct) represents the BM. This is a flexible, fibrous structure with some rather amazing biophysical properties.

Stapes motion causes a waveform closely related to the acoustic waveform to travel down the BM toward the apex.



(Figure 7.)

This figure shows what the traveling wave might look like at one instant in time. Note that the peak grows slowly and then dies out rapidly beyond a certain point.

The height of the peak is dependent upon the amplitude of the tone. The smallest response at threshold of hearing has a displacement amplitude of 1 billionth of a cm. Above the threshold of pain the displacement amplitude is 1 hundredth of a cm.

For a sinous tone the traveling wave will stay inside of an envelope of BM displacement amplitude. The place on the BM where the peak of the envelope occurs depends on the frequency of the tone the peak for high frequency tones is close to the base of the cochlea (near the stapes) and the peak for low frequency tones is close to the apex of the cochlea. This means that the BM performs a rough spectral analysis of the sound. For complex sounds, there would be multiple peaks.

The Transduction Process

The hair cells are embedded in a rigid cellular structure (the organ of Corti) which rests on top of the BM. As the BM moves, this structure moves as a whole. The tops of the hair cells are interconnected by a fibrous structure. Protruding from the tops of the hair cells are the hairs, or stereocilia. There is a different arrangement of these hairs for the inner and outer hair cells which may have a functional significance.

Some of the hairs of the outer hair cells are embedded in the overlying tectorial membrane (TM). When the BM is moved in one direction or the other by a pressure wave, both the organ of Corti and the TM move in the same direction. However, they both have separate pivot points on the bone adjacent which causes a shearing motion between the TM and the surface of the hair cells. This causes the cilia to bend .

It is thought that the bending of the cilia causes a deformation of the top of the hair cell which causes electrochemical changes in the hair cell itself. These changes in turn cause the stimulation of the auditory nerve fibers connected to the hair cell.

The way the auditory nerve "fires", or sends impulses to the brain, depends on the frequency, amplitude, and temporal properties of the sound. Remember that the BM traveling wave peaks at different places depending on the frequency. Well, since many auditory nerve fibers are connected to specific hair cells at specific locations in the cochlea, these fibers will generally respond only to a limited band of frequencies. So in the cochlea, frequency is roughly mapped into "place". This means that as far as the higher auditory processing centers are concerned, "which" fiber is firing tells that center something about the frequency content of the sound.

But it turns out that the hairs only cause the cell to stimulate the auditory nerve fibers when they are bent in a certain direction and not in other directions. So as they oscillate back and forth according to a filtered form of the wave passing by, this imposes a certain temporal pattern on the firing of the auditory nerve. For low frequencies (<500-800 Hz) this can be "time-locked" to the waveform itself. Thus auditory fibers can carry spectral information about the signal by their "location" and can carry temporal information about the signal by their pattern of firing. This is a primary example of the parallel encoding of spectral and temporal representations of the acoustic waveform.

IRCAM summerstage 1980
Course notes

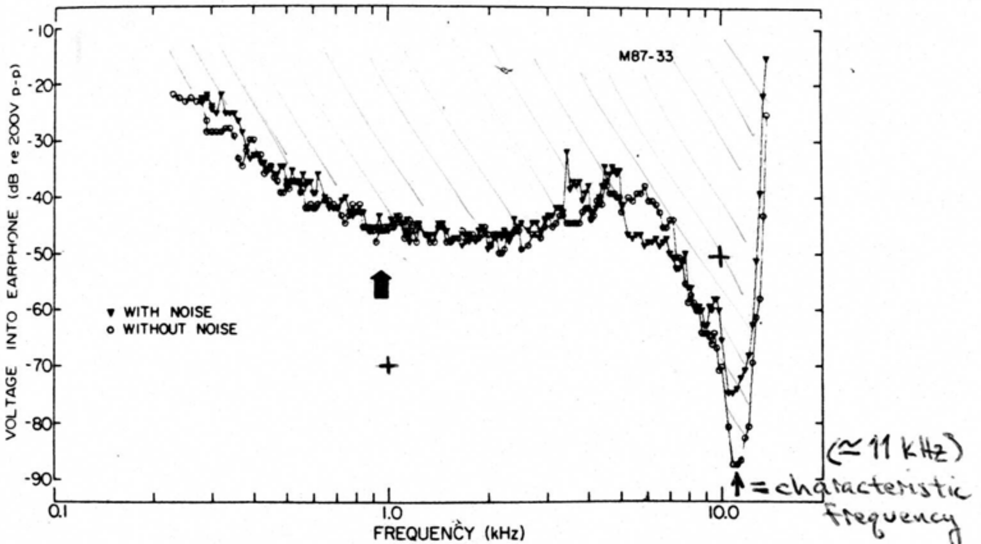
AUDITORY PHYSIOLOGY II.

Stephen M c ADAMS

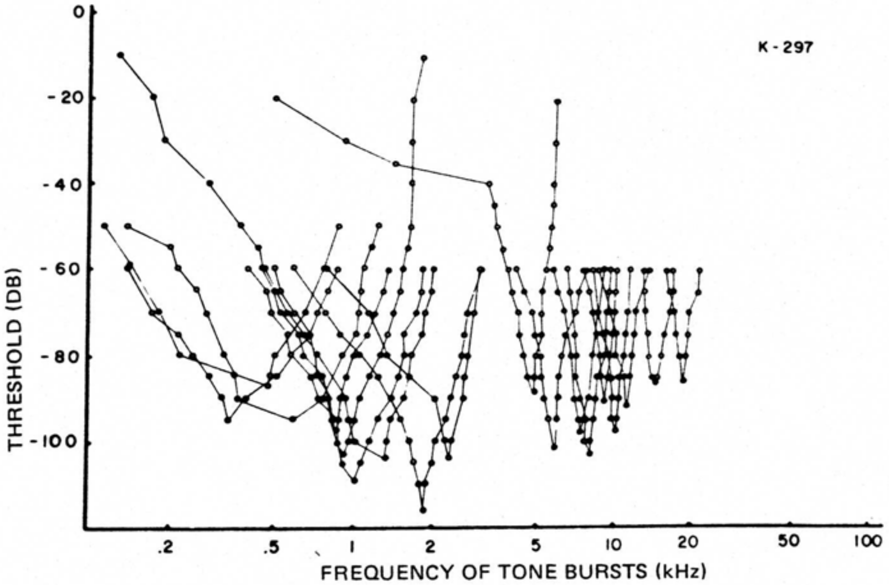
AUDITORY NERVE

The fibers of the auditory nerve (also called "cochlear nerve") are suprisingly homogeneous in their response properties. They all :

1) respond only to a narrow band of frequencies (the exact extent of which is dependent on the intensity)
2) have a relatively small dynamic range (40dB)
3) exhibit spontaneous firing activity (i.e. they are active without any input.
4) exhibit time-locking properties (i.e. respond to temporal properties of the waveform depending on frequency)



This figure shows what is called a "frequency threshold curve" or "tuning curve" of an auditory nerve fiber. Each point on the curve represents the intensity at a given frequency for which the fiber just began responding. At a lower intensity the fiber would not respond. The area above the curve is called the "frequency response area", that is, it represents the frequency-intensity combinations the fiber will respond to . The frequency to which the fiber responds at the lowest intensity is called the "characteristic frequency". Notice that the fiber responds to a narrow band of frequencies until the intensity is increased by more than 40 dB at which point it starts responding to lower frequencies. This broad, low-frequency area means that loud low frequencies can affect the output of a fiber usually "tuned" to higher frequencies. This is due to the asymmetrical shape of the traveling wave envelope on the basilar membrane.

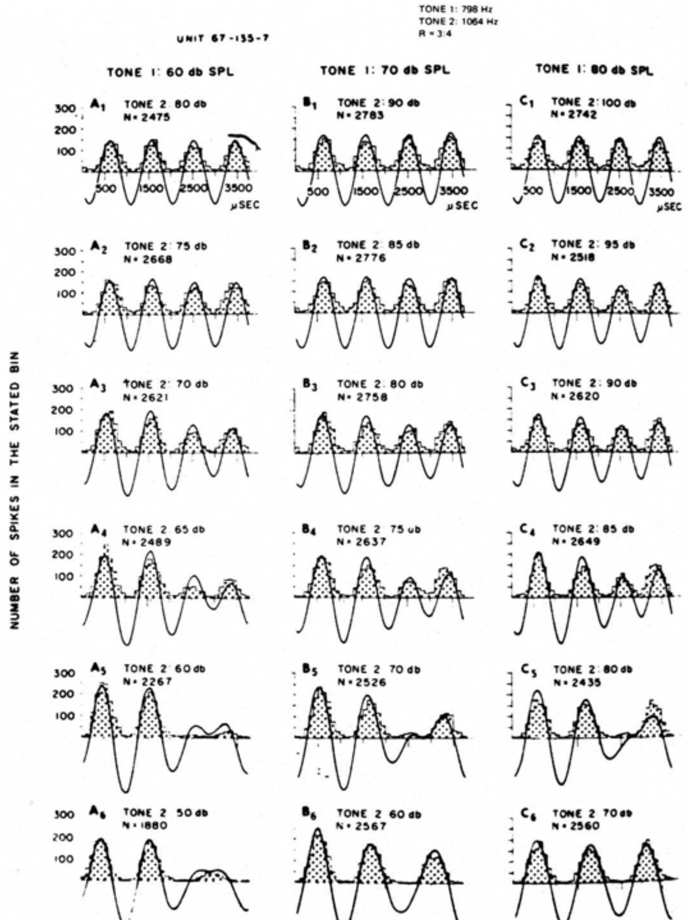


(Figure 2.)

This figure shows a whole lot of tuning curves recorded from different fibers. Note that the bandwidth changes as a function of the center frequency. This may have some relation to what is called the "critical band" in psychoacoustics.

Another interesting thing of musical relevance that occurs in auditory fibers is called "suppression". This means that a tone at a higher or lower frequency can affect the performance of a fiber in response to a tone at the characteristic frequency.
This may be involved in interactions between harmonics or partials and may have an effect on total loudness of complex sounds and spectral aspects of timbre perception.

We noted in the cochlea that frequency gets mapped onto the basilar membrane and that there was a one-to-one connection between inner hair cells and auditory nerve fibers. We call this mapping of frequency to place a "tonotopic" arrangement. This tonotopicity is preserved in the auditory nerve and, to some extent, throughout the entire auditory system. However at higher processing centers there is a convergence of fibers from many frequency regions onto single neurons. This produces broad tuning curves at these higher levels, but it turns out that this broader tuning is good for doing a spectral analysis of time-varying sounds.



We also noted in the cochlea that due to the functional polarization of the hair cells temporal information could be encoded in auditory fibers. This could be time-locking to low frequencies or to amplitude envelopes in higher frequency signals. Figure 3 shows a fiber that actually follows the complex waveform produced by two sinusoids as their absolute and relative intensities are varied.

So if we imagine a complex signal, such as speech or musical sounds in which the spectrum varies over time, we might see that this would be represented in the auditory nerve as a variation over time of which fibers are firing and in what temporal patterns they are firing. Everything that gets to the brain from the ear must go through the auditory nerve. So the information must be coded here in some manner.

SENSORY CODING

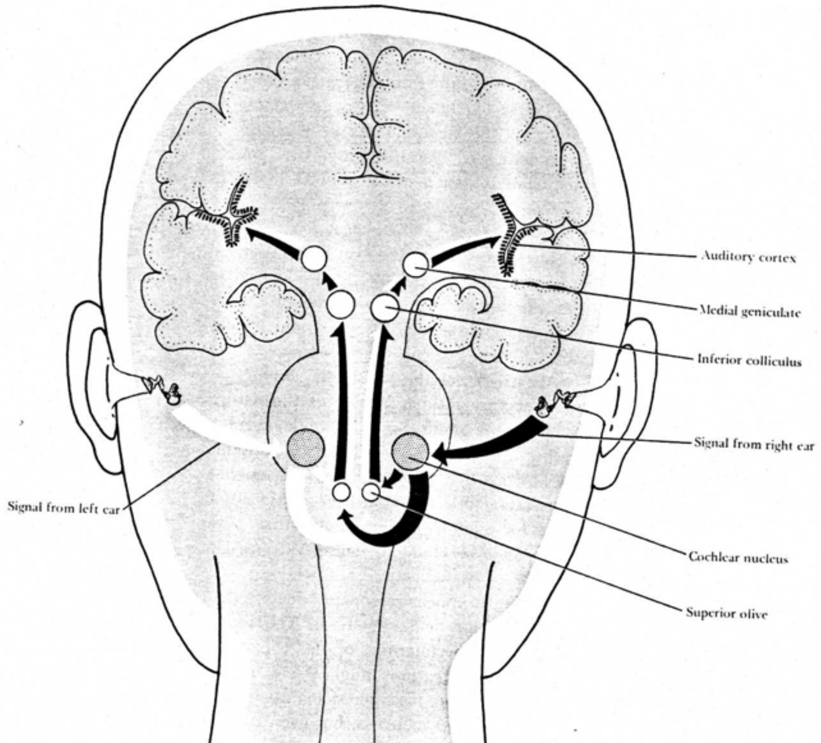
The problem of understanding the physiological processing of acoustic information is one of understanding the coding, transformation (recoding) and decoding performed by the nervous system. Obviously, by the time the information about the acoustic environment gets to auditory cortex its representation is entirely different from what it was when it first encountered the outer ear.
It would be wise at this point to mention a few limits of the methods used to extract information about sensory coding. First, animals are always anesthetised, which can have dramatic effects on the operation of differents parts of the nervous system, especially the higher stations where there is a lot of interplay with the rest of the brain. Also, the types of recordings made have serious limitations with regard to extrapolating to the operation of the whole system. One kind of recording uses micro-electrodes which are stuck into or nearby single cells or fibers and which record the electrical activity of that cell . The problem here is that the number of cellular units which make up the auditory system is enormous and it is quite likely that perception results from the operation of hundreds, thousands, or millions of cells simultaneously. Another kind of recording is with gross electrodes placed on the scalp. What these record

is likely to be the summed activity of thousands of cells at a time. So to extract the activity of a particular group of cells from the complex, composite record is at this point nearly impossible in all but a few special cases (though advances *are* being made in this direction).

To accurately characterize the activity of the auditory nervous system we would need to be able to describe the time-varying activity over hundreds of thousands of cells and neural fibers. And we would need to be able to characterize the parallel processing of temporal and spectral information throughout the system which is necessary to extract and/or synthesize the salient, or relevant, features of the acoustic environment. There is a lot of noise out there that needs to be ignored, and that is no easy task for any system.

CENTRAL AUDITORY NERVOUS SYSTEM

After leaving the cochlea there are approximately four main processing stations before the neural signals reach cortex. These are the cochlear nucleus, the superior olive, the inferior colliculus and the medial geniculate body of the thalamus.



As far as we know, all auditory nerve fibers synapse in the cochlear nucleus. There are 2 or 3 main branches (depending on your classification scheme) leaving cochlear nucleus. After passing either directly or indirectly through superior olive and inferior colliculus all ascending pathways synapse in the medial geniculate which sends projections to cortex. These are general features of most of the other sensory systems : projection to a preliminary processing nucleus, further processing, synapse in some appropriate nucleus in thalamus and then projection to sensory cortex. It should also be noted that most information from the ears eventually gets sent to both sides of the brain.

There are also descending feedback systems between most of the main stations which eventually project all the way back to the cochlea itself, where incoming information can be directly modified before it ever gets to the brain. This may be important for such higher level processes as are involved in attention, which is basically the neural filtering of irrelevant information. If attentional processes are projecting as far back as the cochlea this might imply that some information may never even <u>get</u> to the brain.

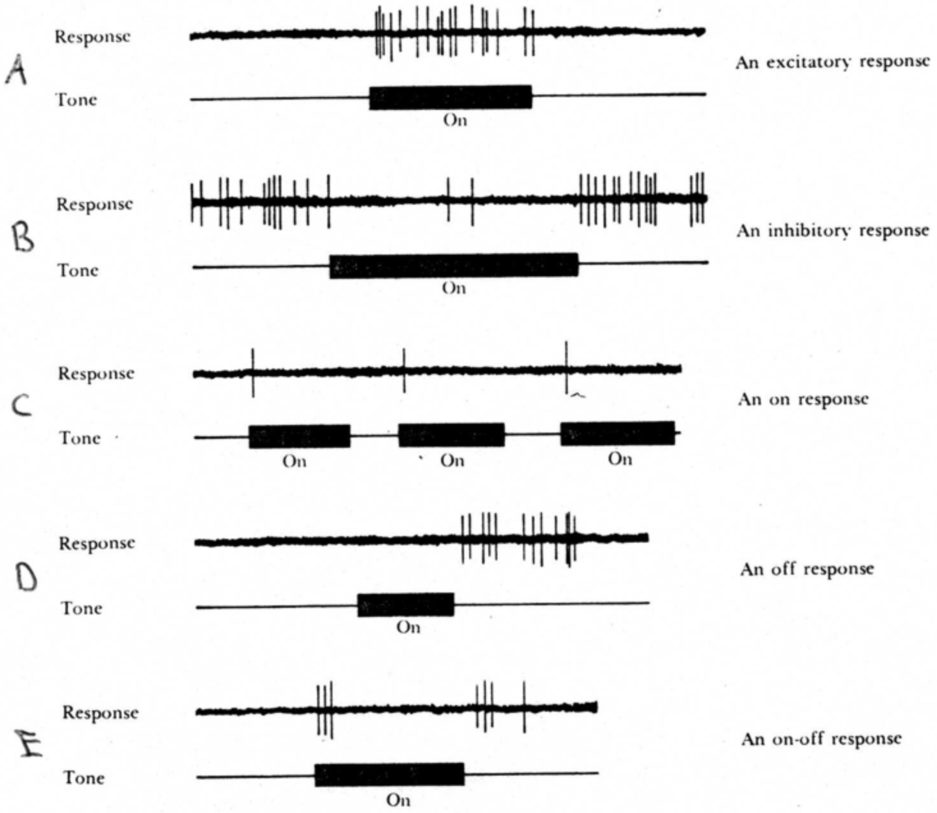
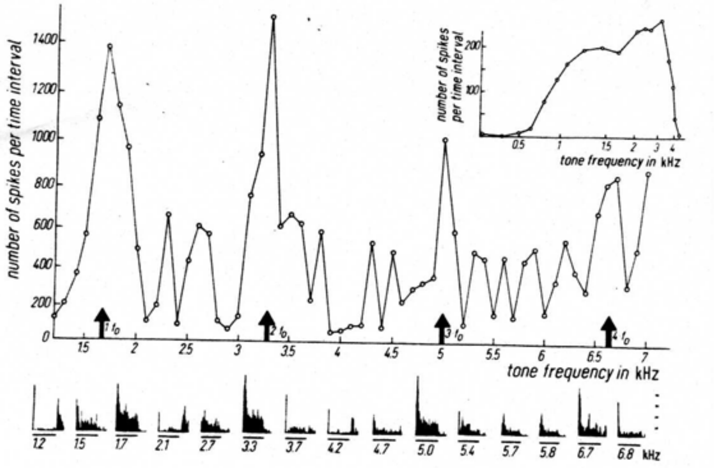


Figure 5 shows a variety of types of response to sound different types of cells in the auditory system have. These were actually recorded from cortex but may be found throughout the system. Cell A only responds when a sound is present , sustains its response during the sound, and is silent otherwise. Cell B is exactly the opposite: it turns off its otherwise high level of spontaneous activity when a sound is present. Cell C only turns on briefly at the beginning of a sound while Cell D turns on at the end of the sound. Other ways of characterising cells are by the time course of their response, their frequency response area and by their response to variations in amplitude and frequency.

It was mentioned previously that the tonotopic organization of the cochlea is preserved (with varying degrees of accuracy) all the way to auditory cortex. In each of the auditory nuclei (as these higher stations are called -+-) it has been found that the spectrum map is preserved in the organization of the cells, with high frequency cells being in one region, low frequency cells in another with all the mid-frequencies in between. While this is generally true up through cortex, what does seem to change is both the frequency resolution of individual cells (i.e. the frequency range to which they respond; higher resolution means narrower range of response) and the complexity of sound necessary to elicit a response in the first place. As a general rule frequency resolution decreases the higher the cell is in the system and this is accompanied by an increase in complexity which is responded to, eg. some cells in cortex will not respond to pure tones but will respond to the sound of the refrigerator door opening. Also, the frequency response curves get more complicated. In inferior colliculus and medial geniculate one finds response curves with multiple peaks (see figure 6) and also with areas where tones inhibit rather than excite the cell. So frequency components in inhibitory regions compete with frequency components in excitatory regions. The final response of the cell is the sum total of the effects of all these inputs.

(+) A nucleus in the central nervous system denotes a collection of cell bodies into some functional unit; remember that a nucleus is also an inner part of any single cell which controls the metabolism, regrowth, reproduction of the cell - it contains the cells DNA.

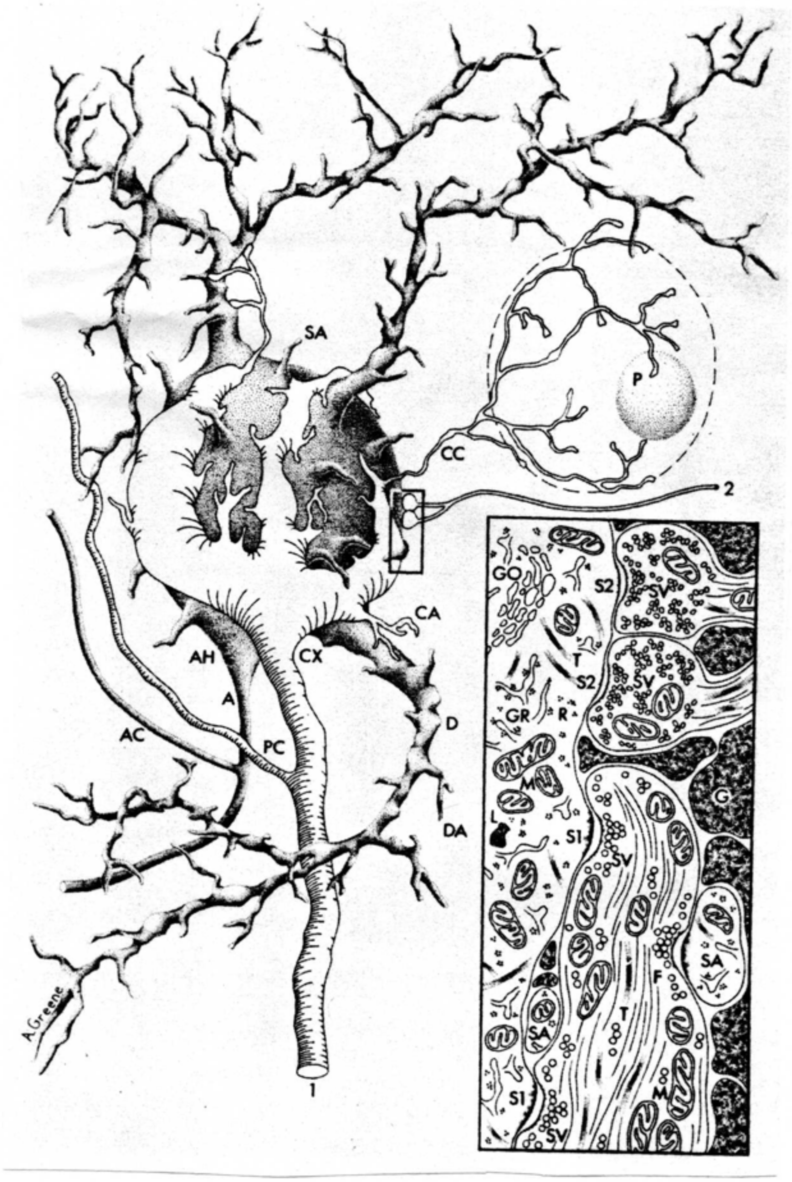


(Figure 6.)

In medial geniculate there is evidence (Figure 6.) for cells that may act like "harmonic templates". That is, they would respond when any of the harmonics of a certain fundamental were present. If this is a more general case in auditory processing it would imply that,in a sense, the auditory system is biased toward hearing (or at least extracting information from) harmonic sounds.

TEMPORAL PROCESSING

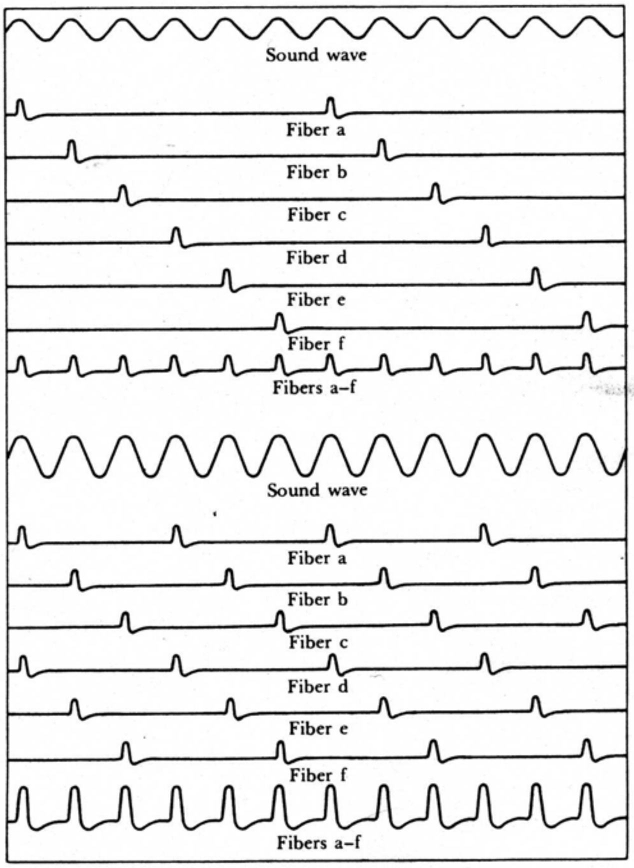
Temporal processing is extremely important in audition, in fact, the auditory system can probably process information faster than any other sensory system . If precise temporal information is to be preserved in the central auditory nervous system there must be special kinds of synapses to ensure that preservation. Indeed in cochlear nucleus and superior olive one finds very large synapses which virtually cover the cell body. These are called "chalice" or calyx endings and they have the property that whenever they fire, the cell they synapse on fires as well with a very predictable delay (which is accounted for by the time it takes the transmitter substance released from the calyx to reach the cell body and cause the electrical change to take place which causes the cell to fire). Such endings are found on cells in cochlear nucleus and at the ends of the axons of these cells on cells in the superior olive. In fact there is evidence that these cells in sup.olive are arranged such that temporal information from both ears gets cross-correlated, which could provide an ideal mechanism for dealing with the temporal cue in localization.



(Figure 7.)

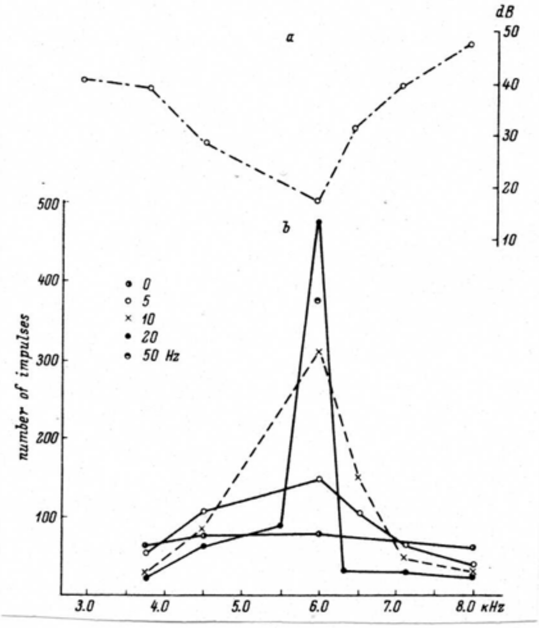
There is also evidence that groups of cells may cooperate in "volleying" temporal information at rates higher than any one fiber could carry it. (see Figure 8. next page). This nucleus is about the highest level at which phase locking to sine tones is observed.

In cochlear nucleus cells have been found which respond to the AM rate of gated tones and noise up to modulation frequencies of 500-800 Hz. In fact some cells seem to have a "best modulation frequency" as well as a characteristic frequency, which means the cell is "tuned" to a certain frequency range with certain AM characteristics. It seems that adjacent units with otherwise very similar properties can have different phase responses which may provide important information about features of the sound to higher processing centers.



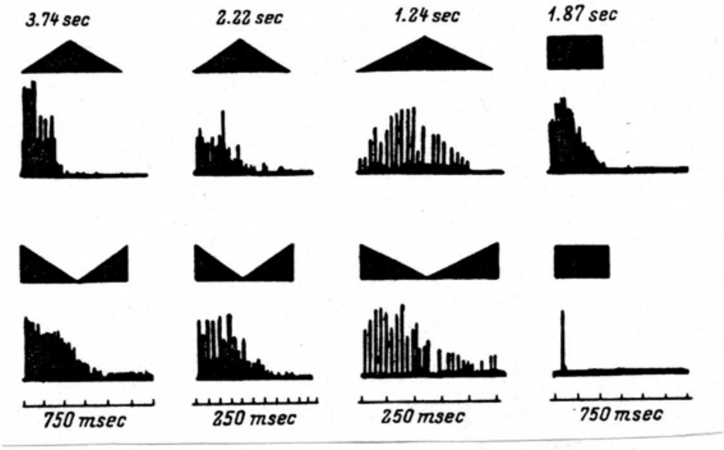
(Figure 8.)

In inferior colliculus cells of different response types seem also to respond differently to AM signals. Sustaining cells (like cell A, Figure 5) synchronize their output with AM rates up to 30-50 Hz. Onset cells (cell C, Figure 5) synchronize with AM rates up to 30-100 Hz. The actual cell discharge pattern depends on the gross structure of the sound indicating that features are more important than fine structure in the inferior colliculus.



(Figure 9.)

The figure shows the broad tuning of a cell in inferior colliculus (top graph) and also shows the complex relationship between response to carrier frequency-modulation frequency combinations (bottom graph) Each curve represents a different modulation rate. Note that the total response of the cell depends on both the carrier frequency and the modulation frequency and that there is a slight shift in the best carrier frequency with a change in modulation frequency. This means that there is a specific constellation of sounds to which this cell will respond maximally. This constellation could be thought of as an important feature of some sound relevant to this animal (cat, in this case). Note also that these kinds of characteristics of a cell are only seen with time-varying sounds. If one presented only steady-state sine waves these characteristics would never become apparent.



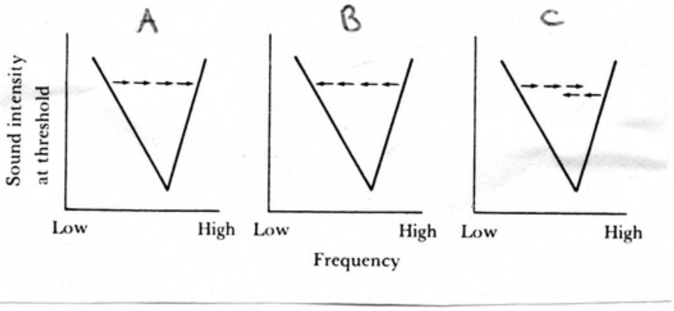
(Figure 10.)

Here is another example of a complex response of one single cell in inferior colliculus to a constellation of different types of amplitude modulation durations. The solid black forms represent the amplitude envelopes of the sound. Note in this case that as one changes the duration which changes rate of change of amplitude, the type of response to change in amplitude varies. Other types of cells in inferior colliculus will exhibit a complex response to rising and falling amplitudes which changes with respect to overall intensity rather than duration as is the case in this last figure. So each cell responds to a complex family of structural features of the sound.

Cells in these auditory nuclei also have complex responses to FM signals. Just as cells in cochlear nucleus seemed to have "best AM" rates, so do some of them have "best FM" rates. So each cell may be thought to respond to a best carrier frequency, a best AM frequency and a best FM frequency.

Cells in inferior colliculus, medial geniculate and auditory cortex all respond best to sounds which are continually changing and readily habituate to steady sounds (they get bored and go to sleep).

Remarkable FM responses are found in cortex ;cells have been found which have one of 4 modes of response to tones sweeping in frequency (glissandi) :



(Figure 11.)

1) sweeping across the boundaries of the frequency response area determined with steady tones,
2) sweeping within the frequency response area : either up (A in Figure 11) or down (B) or split (C),
3) outside the frequency response area; that is, the response area is bigger for moving tones than for steady ones, and
4) responds to FM tones ,but will not respond at all to steady sine tones.

Thus the response to complex sounds is not necessarily predictable from the response to pure tones.

SUMMARY

We have seen that the response to pure tone stimuli becomes increasingly complex higher in the system. The responses to complex sounds in-may play an essential role in pattern recognition. There are (at least) two main paths leaving the cochlear nucleus (remember, all auditory nerve fibers stop off here on the way up):

1) the quick path to superior olive preserves spectral and temporal information intact and with great acuity and may thus be involved with a "where" mechanism for locating sound sources, and
2) a more complex path heading through higher stations with much processing may be involved with a "what" discriminative and pattern-recognition/ feature-extraction mechanism.

These processes of feature extraction, data reduction and identification of sounds in the light of past experience are crucial for functioning in our everyday lives.

The complexity of interconnection between and within the auditory nuclei almost defies description and functional investigation. But this richness of processing is surely (at least in part) responsible for the vast complexity of auditory experience available to us.