

Hearing, part 1

Josh McDermott
9.35
April 27, 2004

THE EAR

(Image removed due to copyright considerations.)

Functional schematic of the ear:

(Image removed due to copyright considerations.)

Show movies!

Per ear:

3500 inner hair cells
12,000 outer hair cells
30,000 auditory nerve fibers

Per eye:

5 million cones
100 million rods
1.5 million optic nerve fibers

HEARING LOSS

Normal hearing looks like this:

Image removed due to copyright considerations.
Please see: Sekuler, Robert and Blake, Randolph. Perception
(Fourth edition). New York: McGraw-Hill, 2001.

Figure 10.1, p. 330.

But lots of people, especially old people, have hearing loss.

This is measured as an increase in threshold dB level.

(Image removed due to copyright considerations.)

A 20 to 40 dB loss is significant (100 to 10,000 fold decrease in intensity), similar to what happens if you put in earplugs or put your hands over your ears.

(Image removed due to copyright considerations.)

Two kinds of hearing loss:

Conduction loss - an impairment in the mechanical transmission of sound energy to the inner ear.

-i.e. a problem with the outer or middle ear

Sensory/neural loss - an impairment in the transduction of sound energy to electrical signals, or in their transmission to the brain.

-usually due to problems with the inner ear.

Conduction loss

-puncturing of ear drum (via ear infection)

-otosclerosis

-the gradual immobilization of the stapes (stirrup) bone in the ossicles.

-inherited, incidence is pretty high (10%)

-treated by surgically replacing the stapes

-it can spread to the cochlea, in which case you got a problem.

Sensory/Neural Loss

presbycusis - as people age, they become less sensitive to high frequencies

-most people over 30 can't hear above 15kHz

-most people over 50 can't hear above 12kHz

-worse in men than women

Image removed due to copyright considerations.

Please see: Sekuler, Robert and Blake, Randolph.

Perception (Fourth edition). New York: McGraw-Hill, 2001.

Figure 10.2, p.335.

-unclear what causes presbycusis, perhaps changes in elasticity of cochlea

-appears to be a genetic component

-could also just be prolonged exposure to noise

-one study of an African tribal people living in a natural (non-industrialized) environment found unimpaired hearing in 70-year-olds.

Noise exposure can cause hearing loss, through damage to the organ of Corti.

(Image removed due to copyright considerations.)

Yet another reason for gun control - hunters are missing hair cells!

It is unclear what causes the hair cells to die, but it is clearly related to the noise.

(Image removed due to copyright considerations.)

The right ear is protected by the acoustic "shadow" cast by the head.

Interesting facts about hearing loss:

-Ronald Reagan suffers hearing loss from his days as a Western actor.

-People who regularly attend rock concerts (1/month) have higher thresholds than those who don't, but only by 2dB.

-Workers in noisy environments suffer permanent hearing loss.

-Moderately loud sounds (60 - 100 dB) won't lead to permanent damage, but they do produce temporary increases in thresholds.

-These can be quite large.

-Recovery can take hours to days.

-Unclear whether these are related to permanent hearing loss.

Hearing loss can impair speech perception.

Image removed due to copyright considerations.

Please see: Sekuler, Robert and Blake, Randolph. Perception (Fourth edition). New York: McGraw-Hill, 2001.

Figure 10.2, p.335.

Phones only transmit up to 4 kHz, so frequencies above that are not critical, but those between 1kHz and 4kHz are important, especially for certain consonants.

(Image removed due to copyright considerations.)

So what's the big deal?
We can just use a hearing aid to turn up the volume, right?

Wrong.
Sensory/neural hearing loss is not the same as decreasing the stimulus level.□

Such hearing loss is typically accompanied by damage to the outer hair cells.

Recall that the outer hair cells serve to amplify the response to □ sound, and sharpen tuning.□

Without OHCs, the response to sound is not just lower amplitude, but abnormal in two other ways due to the absence of nonlinear contribution of OHCs:

Loudness recruitment - quiet sounds are inaudible, but loud sounds □ sound just as loud.□

Poorer frequency selectivity due to reduced sharpness of response.

Loudness recruitment is due to absence of nonlinear response in cochlea, due to damaged/missing outer hair cells:

(Image removed due to copyright considerations.)

So, turning up the volume won't work.

-Loud sounds will become uncomfortably loud.

-Current hearing aids adjust amount of amplification depending on level of sound, but it is hard to do this quickly.

-reduced frequency tuning is even harder to fix.

So, even with hearing aids, people with missing outer hair cells have problems.

COCHLEAR IMPLANTS

Between 1 and 2 million people suffer from sensorineural deafness.

Treatment of choice is the cochlear implant.

Basic idea - if the problem is that the hair cells don't work, but the auditory nerve fibers are mostly normal, then the nerve fibers can be stimulated directly with electrodes, and the person will hear something.

More than 30,000 people have had cochlear implants.

They are controversial within the deaf community.

(Image removed due to copyright considerations.)

(Image removed due to copyright considerations.)

Current cochlear implants use 16-24 electrodes.

Each electrode gets a bandpass filtered version of the speech signal, with the frequency chosen according to its location on the cochlea.

The most successful patients are able to conduct phone conversations with the help of the implant.

Many people are actively studying signal processing techniques to improve cochlear implants.

If the auditory nerve is degenerated, the main recourse is a brainstem implant, in which electrodes are implanted in the cochlear nucleus. These are still highly experimental. Less than 100 people have them.

The output of the cochlea is transmitted to the brain through the auditory nerve.

For high frequencies, auditory nerve fibers fire in proportion to the amplitude of their characteristic frequency.

For low frequencies, their spikes are phase-locked to the stimulus: □

(Image removed due to copyright considerations.)

Phase locking occurs for frequencies under 1kHz.

Most nerve fibers don't fire with every stimulus cycle, especially for higher frequencies.

A stimulus is encoded by many nerve fibers at once. Together, they provide a faithful representation of the stimulus.

This form of population coding may be one reason why there are many more auditory nerve fibers (30,000 per ear) than there are inner hair cells (3500 per ear).

Demo time.

There are thus two cues to the frequencies in a sound:

- 1) The place of excitation in the cochlea.
- 2) The frequency of firing.

Debate on the importance of these two cues has raged for over a century, and is still unresolved.

Both are probably used, but the exact way in which they are combined is still controversial.

A bit more on the cochlea - it's not exactly doing Fourier analysis.

(Image removed due to copyright considerations.)

Bandwidths of auditory nerve fibers are much higher for those carrying high frequencies (note log axis). Called "Q filters".

Bandwidth is roughly proportional to frequency.

This means that low frequencies can be estimated more precisely, but that high frequencies can be better localized in time.

(Image removed due to copyright considerations.)

Onsets and offsets, even of a pure tone (sine wave) generate lots of other frequencies, especially high ones. So to localize sharp transients (e.g. the beginning of a sound), you need to be sensitive to lots of high frequencies at once.

Another way to think about it:

Just as was the case for natural images, it turns out that natural sounds tend to have $1/f$ amplitude spectra:

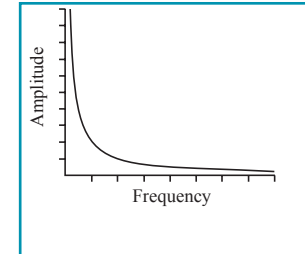


Image by MIT OCW.

So if bandwidth is proportional to frequency, natural sounds will be coded with equal energy in each channel.

Sort of like what we saw in vision.

CRITICAL BANDS

Critical bands are psychophysical analogues of auditory nerve fibers.

They are analogous to spatial frequency channels in vision.

Psychophysical evidence for them comes from masking experiments, first done around 1940.

Observers are asked to detect a pure tone on top of noise filtered in various ways.

Image removed due to copyright considerations.
Please see: Sekuler, Robert and Blake, Randolph.
Perception (Fourth edition). New York: McGraw-Hill, 2001.

Figure 10.9, p.344.

Thresholds for detecting the tone are measured for different noise bandwidths. The noise is centered on the tone frequency.

Image removed due to copyright considerations.

Please see: Sekuler, Robert and Blake, Randolph. Perception (Fourth edition). New York: McGraw-Hill, 2001.

Figure 10.10 upper panel, p. 345.

Noise outside a critical band of frequencies does not affect the detection of the tone.

Just like spatial frequencies sufficiently separated from an adapting grating's freq. were unaffected by the adaptation (remember the bite out of the CSF).

The critical band resembles (and is presumably due to) the tuning curves of auditory nerve fibers.

Image removed due to copyright considerations.

Please see: Sekuler, Robert and Blake, Randolph. Perception (Fourth edition). New York: McGraw-Hill, 2001.

Figure 10.10 lower panel, p. 345.

So, both psychophysical and physiological evidence suggests there are frequency channels in the auditory system, just as in vision.

SOUND LOCALIZATION

How do people tell where a sound is coming from?

Unlike in vision, spatial information is not inherent to the sensory input.

Image removed due to copyright considerations.

Please see: Sekuler, Robert and Blake, Randolph. Perception (Fourth edition). New York: McGraw-Hill, 2001.

Figure 10.14, p. 359.

Three main sources of information about sound location:

interaural intensity differences

interaural time differences

spectral cues from the filtering properties of the pinna

Interaural Intensity Differences

The head creates an acoustic shadow, reducing the sound intensity on one side.

(Image removed due to copyright considerations.)

The intensity differences are negligible for low frequencies, whose wavelengths are on the order of the width of the head.

So this cue is useful for high frequencies only. Hence the single subwoofer.

Image removed due to copyright considerations.
Please see: Sekuler, Robert and Blake, Randolph.
Perception (Fourth edition).New York: McGraw-Hill, 2001.

Figure 10.15, p. 360.

Interaural Time Differences

A sound wave coming from the side will hit the nearer ear first, producing a time lag between the two ears.

(Image removed due to copyright considerations.)

Image removed due to copyright considerations.
Please see: Sekuler, Robert and Blake, Randolph.
Perception (Fourth edition).New York: McGraw-Hill, 2001

Figure 10.16, p. 361.

For pure tones, this cue is ambiguous for high frequencies.

But for complex sounds, and low frequencies, it is less ambiguous.

Jeffress proposed a neural model of how ITD's might be detected, suggesting it could happen via delay lines in the brainstem:

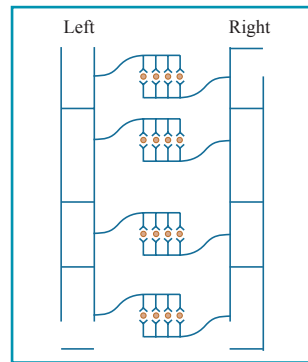


Image by MIT OCW.

Show movie

A Jeffress-like circuit has been found in the medial superior olive, part of the superior olivary complex in the brainstem.

(Image removed due to copyright considerations.)

The central auditory pathways

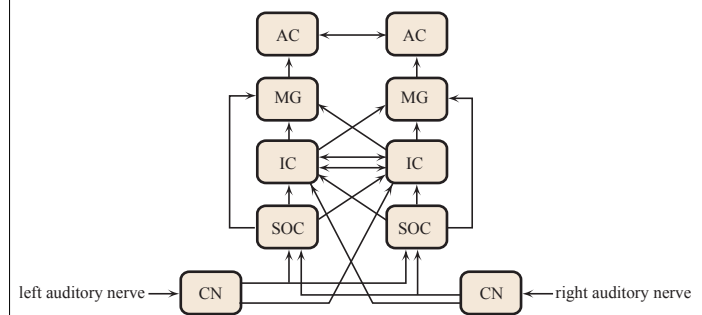


Image by MIT OCW.

A map of auditory space exists in the midbrain nucleus of the barn owl.

The midbrain nucleus is the homologue of the inferior colliculus in mammals.

No one has found such a map in mammals yet.

(Image removed due to copyright considerations.)

So, for low frequencies we can use ITD; for high frequencies we can use IID.

This duplex theory of localization is supported by performance on localization tasks, which is better at high and low frequencies than middle frequencies.

Both only specify the azimuth of a sound, however. The elevation is left ambiguous:

Image removed due to copyright considerations.

Please see: Sekuler, Robert and Blake, Randolph.

Perception (Fourth edition). New York: McGraw-Hill, 2001.

Figure 10.17, p. 362.



To determine elevation, we make use of the filtering properties of the pinna.

Broadband sounds will be “colored” differently depending on their elevation.

The filtering is different for different ears.

(Image removed due to copyright considerations.)

Evidence that we use spectral cues from the pinna comes from studies in which localization is impaired by plastic ear inserts.

You can verify this for yourself by putting on earmuffs and testing your ability to localize sounds.

Also, if you go into a sound-proofed room and listen to the world through microphones, it sounds funny.

(Image removed due to copyright considerations.)

Make a life-sized model of your head and ears, and put the microphones in the ears.

When you listen to the sound through earphones, it sounds natural again, and you can accurately localize sounds.

This won't work as well if you use someone else's head or ears because the sounds are filtered differently from what you are used to.

(Image removed due to copyright considerations.)

Stereo Reproduction of Audio

Stereo audio, introduced in 1958, makes use of the brain's ability to infer sound direction from intensity differences between the ears.

For it to work perfectly, the two audio channels must mimic what the two ears would hear if you were listening to the music being played live. So they must be recorded with two microphones in the ears of a model of the head, and listened to with headphones.

Because music is almost never recorded this way, instruments heard through headphones are mislocalized (they often sound like they are inside your head).

Swapping the inputs to the two ears results in sound mislocalization, but only if the eyes are closed.

Image removed due to copyright considerations.

Please see: Sekuler, Robert and Blake, Randolph. Perception (Fourth edition). New York: McGraw-Hill, 2001.

Figure 10.19, p. 366.

If there is a discrepancy between vision and audition, vision dominates. Localization in vision is more reliable, e.g. because of echoes.

This is the ventriloquism effect.

Judging a Sound's Distance

When outdoors, the main cue to a sound source's distance is the sound's intensity. Not so useful if the sound is unfamiliar.

Indoors, echoes are used to infer distance. The intensity of the energy reaching the ears directly from the source is compared to that of the echo.

People are pretty good at this if there are multiple reflecting surfaces.

If echoes follow the source sound by a small enough lag, they will be heard as a single sound.

Listening in rooms: Echo suppression and the precedence effect

- In many environments, the direct path of a sound to our ears is only one of many. Yet we mostly hear only one source at a location usually corresponding to the that of the source.

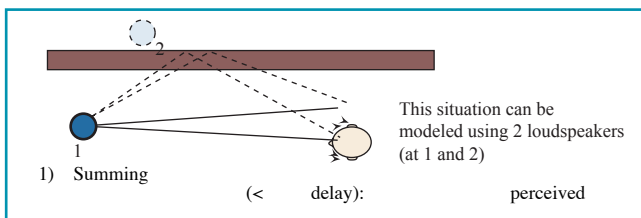


Image by MIT OCW

- location is a weighted sum of the two.
- 2) Precedence effect (ca. 1-5 ms delay): Only one sound perceived: direction of first sound dominant.
- 3) Echo threshold (> 5 ms delay): Two sounds heard.

Localization Errors

Errors tend to occur at locations where IID and ITD cues are ambiguous - e.g. straight ahead and behind.

Errors are more common for narrowband sounds, as broadband sounds can often be localized from the coloring by the pinna.

Head movements can resolve the ambiguities, if sound duration is long enough.

Sound Localization Summary

Binaural cues: interaural intensity and time differences
 IID occurs for high frequencies
 ITD unambiguous for low frequencies
 ITD is detected with delay lines in the brainstem

Binaural cues leave elevation ambiguous.

Spectral cues from pinna filtering help to determine elevation.

Can judge elevation with only one ear, but azimuth requires both.

The distance of a sound source is estimated via sound loudness and echoes (when indoors).

Vision dominates audition when the two are in conflict.

What about the rest of the auditory pathway?

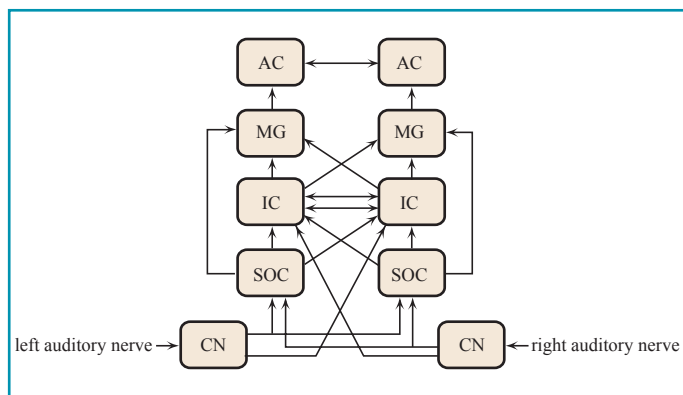


Image by MIT OCW

In vision, there is an important, obvious new feature of processing that is found only in cortex - orientation selectivity.

The theoretical ideas about simple and complex cells, and beyond, are fundamental to how we think about visual processing.

We know of no such fundamental change in processing that occurs between subcortex and cortex in the auditory pathway, and have no corresponding framework with which to think about how the auditory cortex works.

Auditory neuroscience is still awaiting a breakthrough akin to the discovery of orientation selectivity in the visual system.

There are lots of interesting findings, but no big picture.

A few things you should know about auditory cortex:

Neurons are frequency selective, just as they are in the auditory nerve.

Cortex is arranged in tonotopic maps - the frequency tuning of neurons changes gradually across cortex. Adjacent maps are mirror reversed, as in the visual cortex.

Primary auditory cortex (A1), which receives the auditory projections from the thalamus, is located on Heschl's gyrus in humans.

Some of the fields beyond A1 respond poorly to pure tones. They respond best to complex tones. Tonotopy still exists there, but it has to be mapped with bandpass filtered noise.