Chapters 8, 9
The Auditory System and clinical applications

UTD/ SPEECH SCIENCE

COMD 6305

WILLIAM F. KATZ, PH.D.

Chap 8 - issues

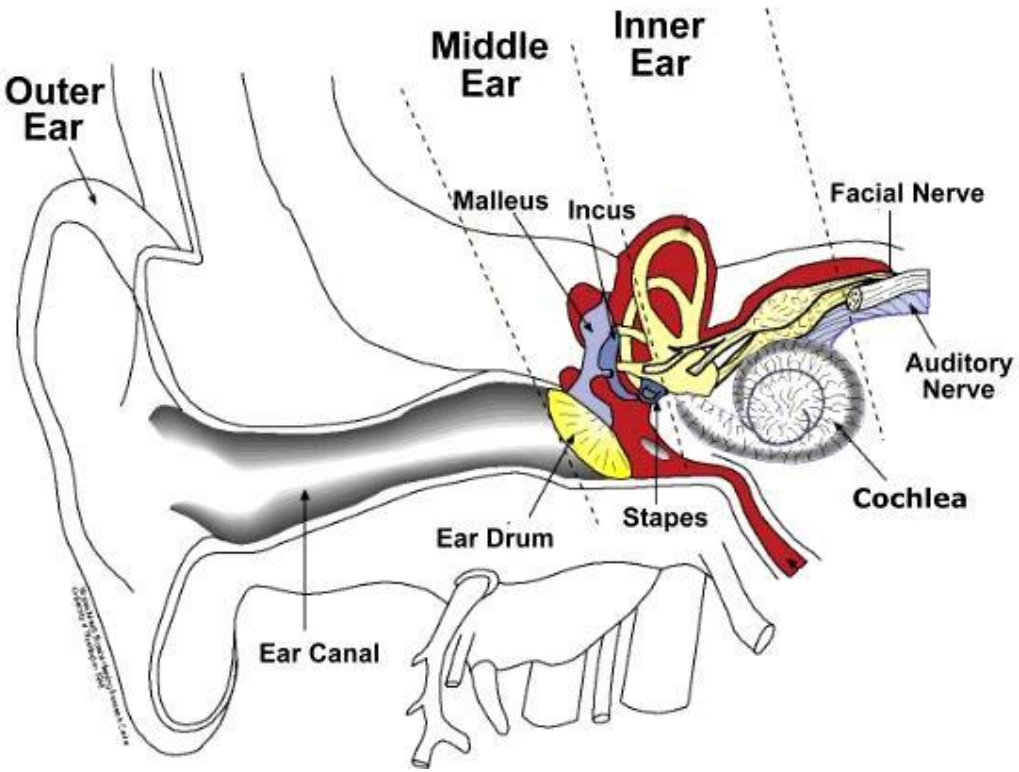
AUDITORY SYSTEM STRUCTURE AND FUNCTION

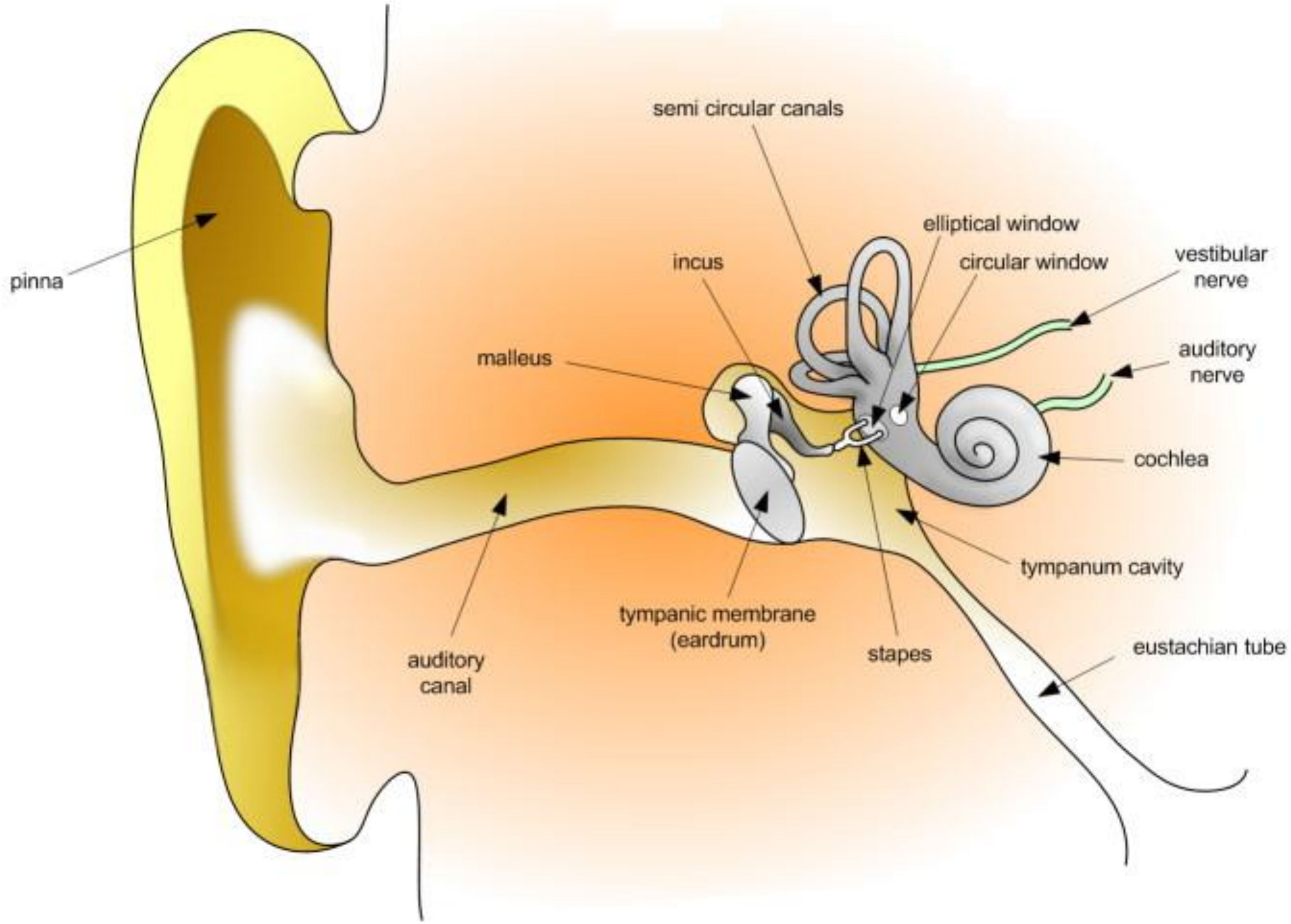
PERCEPTION OF SPEECH

ACOUSTIC PATTERNS OF CONSONANTS AND VOWELS

Outer/ middle/ inner ear

- Pinna, *external auditory meatus*
- (Ear canal) – ¼ wave resonator – boosts high frequency sound
- Together, pinna and ear canal give 10–15 dB amplification for 2500-5000 Hz sounds

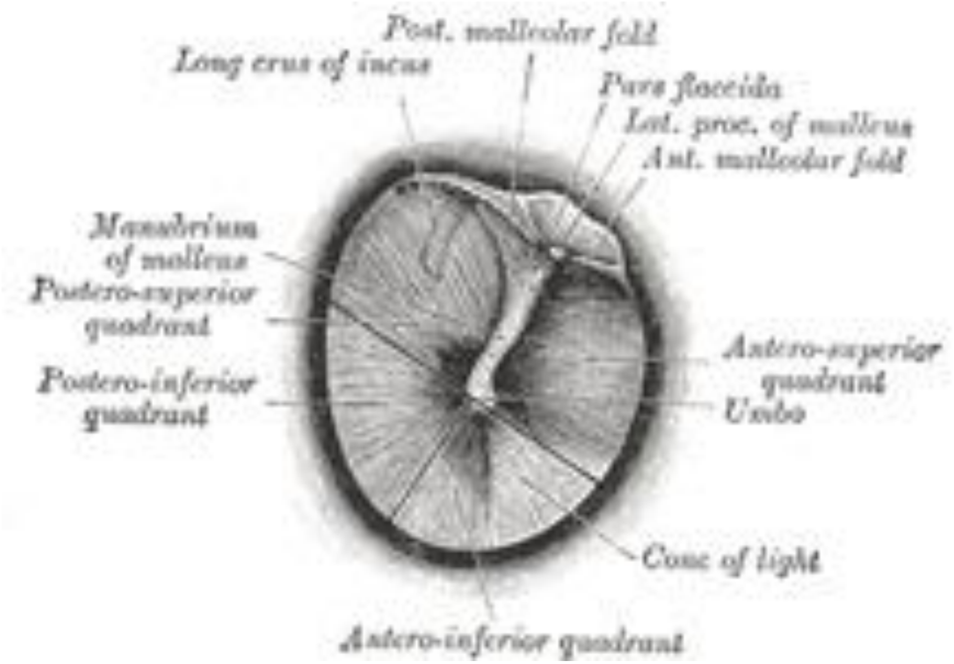




A closer view

Tympanic membrane

GREY'S ANATOMY



PHOTO

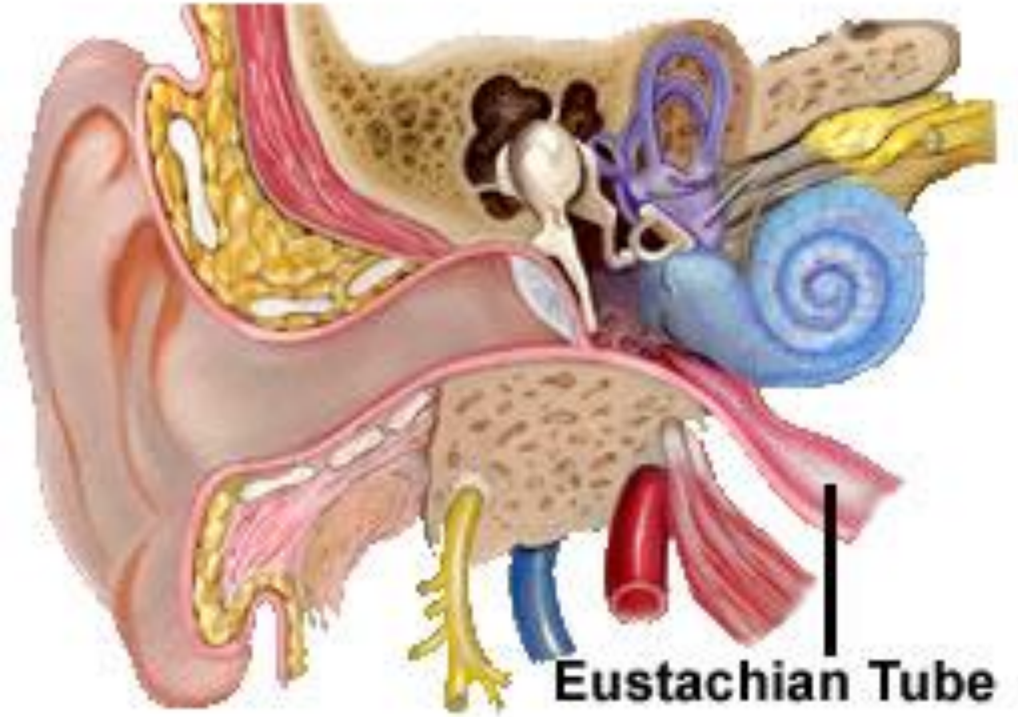
Middle ear - overview

Anatomy

- Ossicles
- Muscles (stapedius, tensor tympani)
- Auditory (Eustacian) tube

Physiology

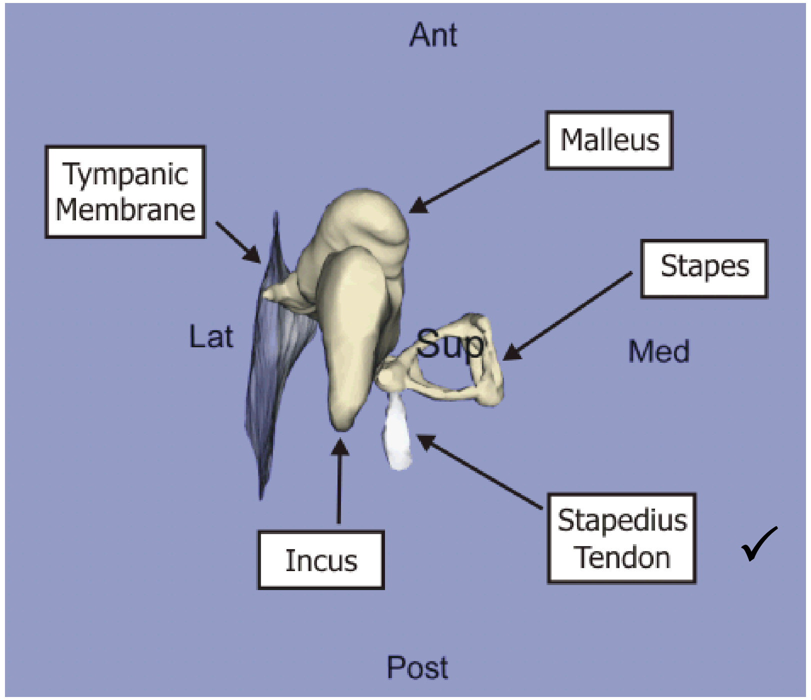
Overcome impedance mismatch of middle/inner ear by increasing pressure changes at oval window



Muscles of middle ear

STAPEDIUS

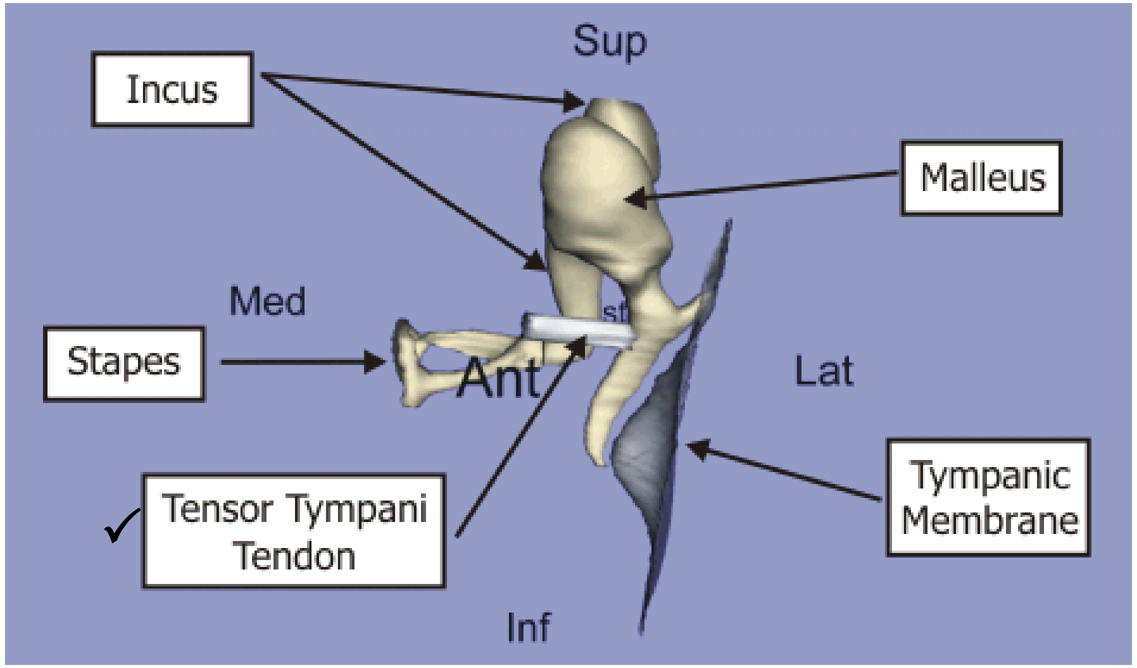
Left Ear



Superior View

TENSOR TYMPANI

Left Ear



Anterior View

Acoustic (stapedial) reflex

Stapedius

In response to 80 dB or more

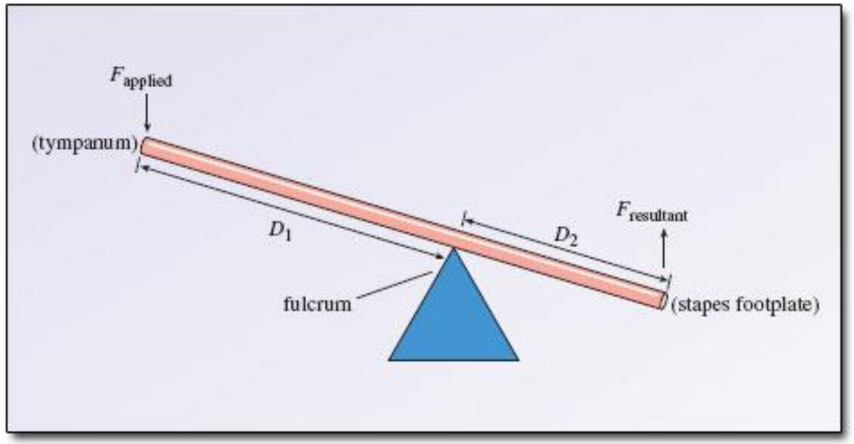
Contraction pulls stapes to the side

Reduces sound pressure by ~ 10 dB

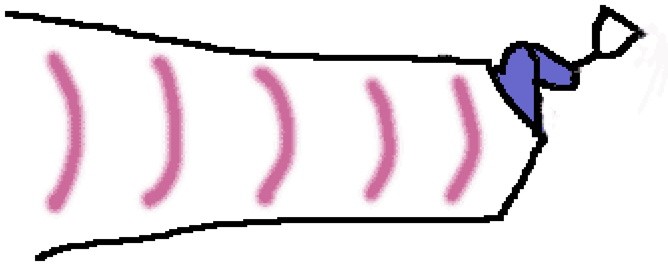
Initiates in fractions of a second, fatigues after time

Two means of overcoming impedance mismatch

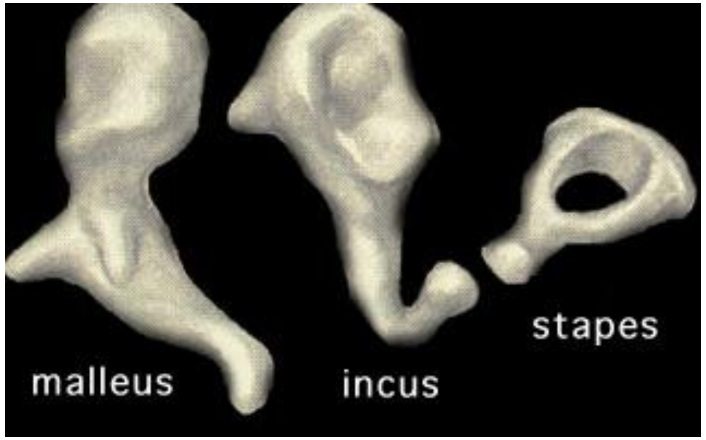
LEVER ACTION OF OSSICLES



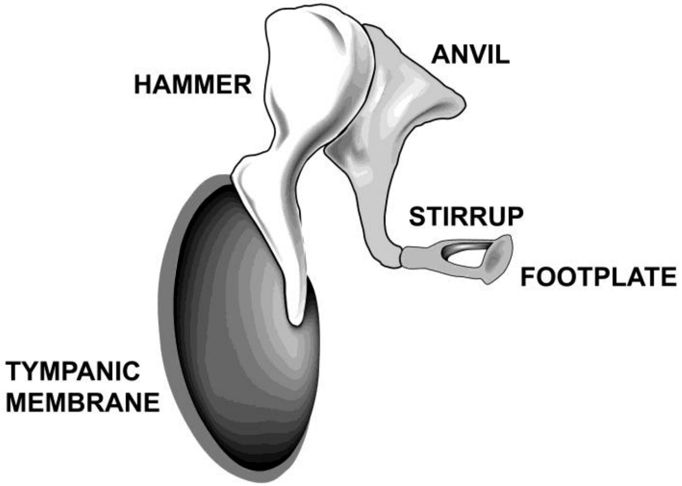
SURFACE AREA DIFFERENTIAL
PRESSURE AMPLIFICATION



LIBERTY
IN GOD WE TRUST
1992
P



malleus
incus
stapes

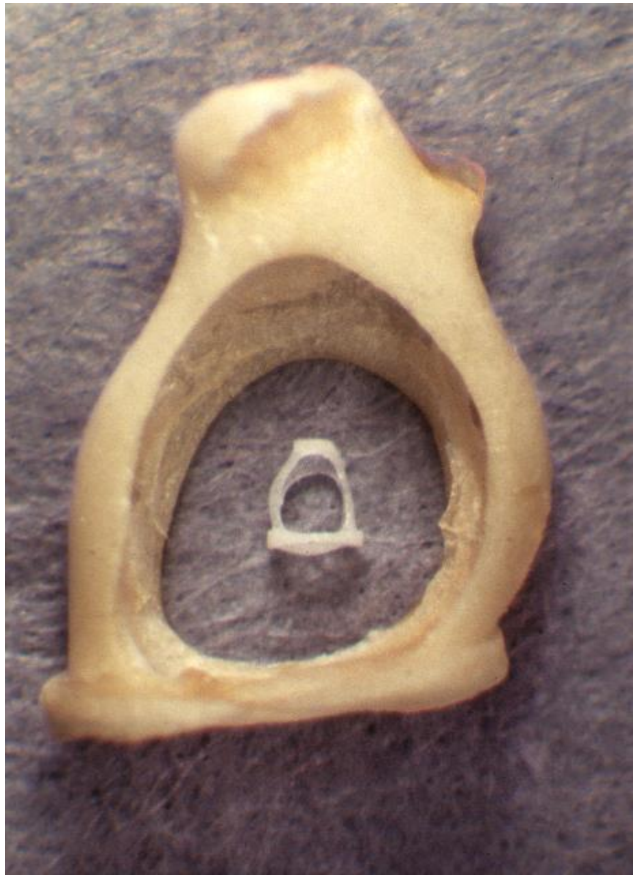


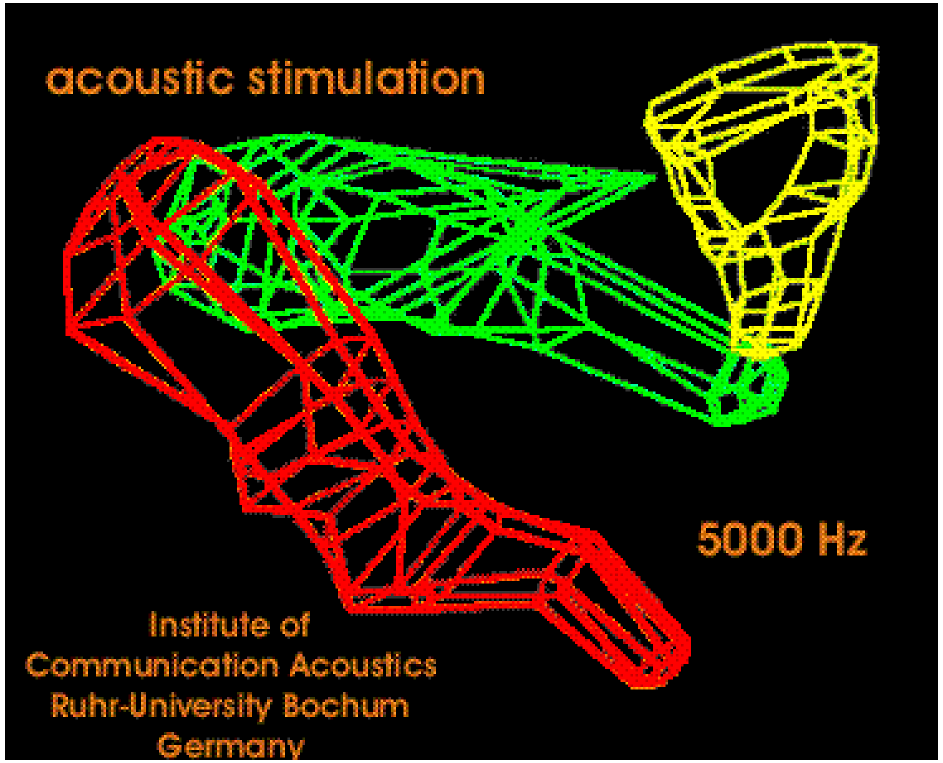
HAMMER
ANVIL
STIRRUP
FOOTPLATE
TYMPANIC MEMBRANE

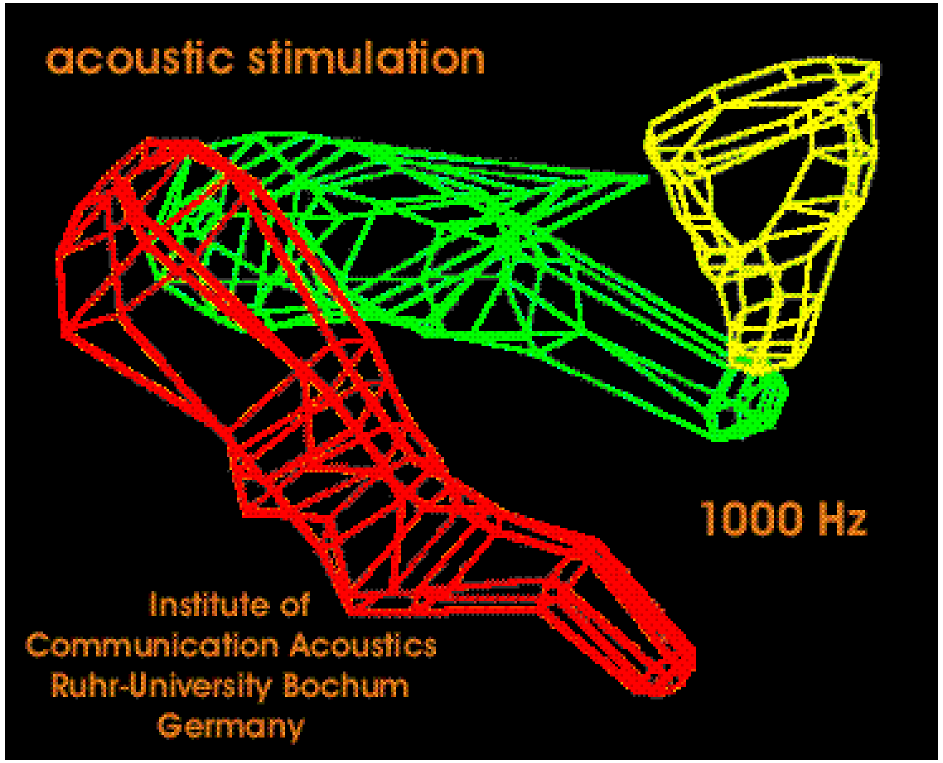
Ossicles

Man vs. mouse

… I couldn't resist…





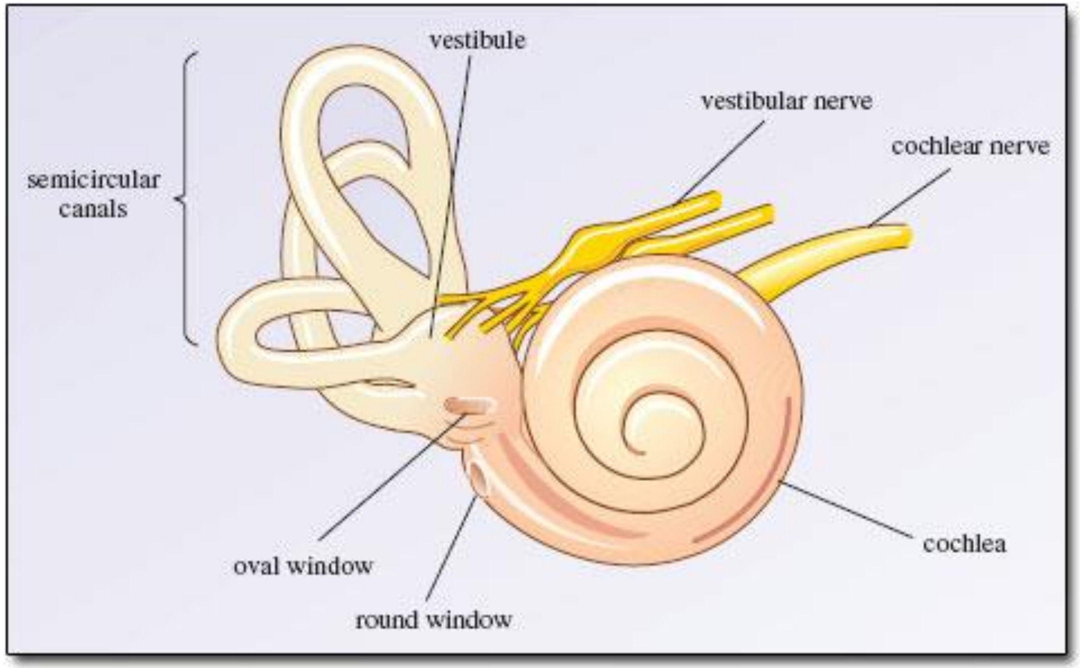


Simulations of ossicle function

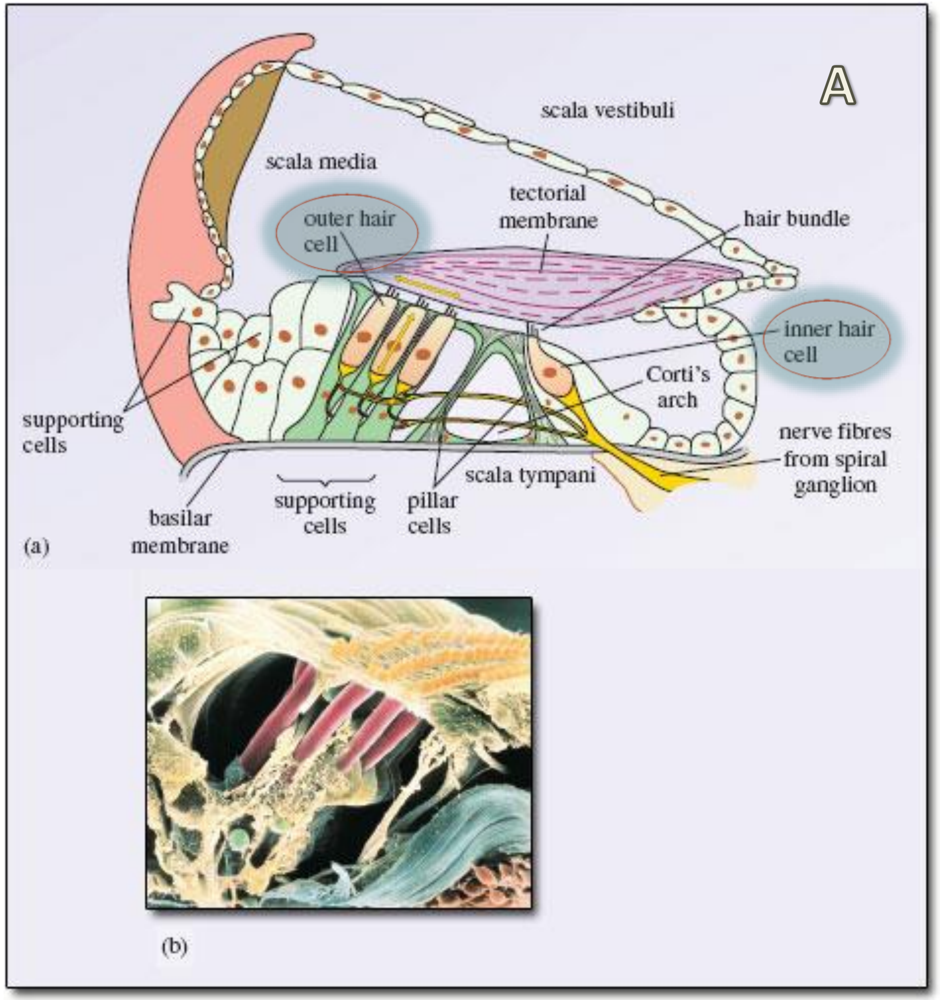
Inner Ear

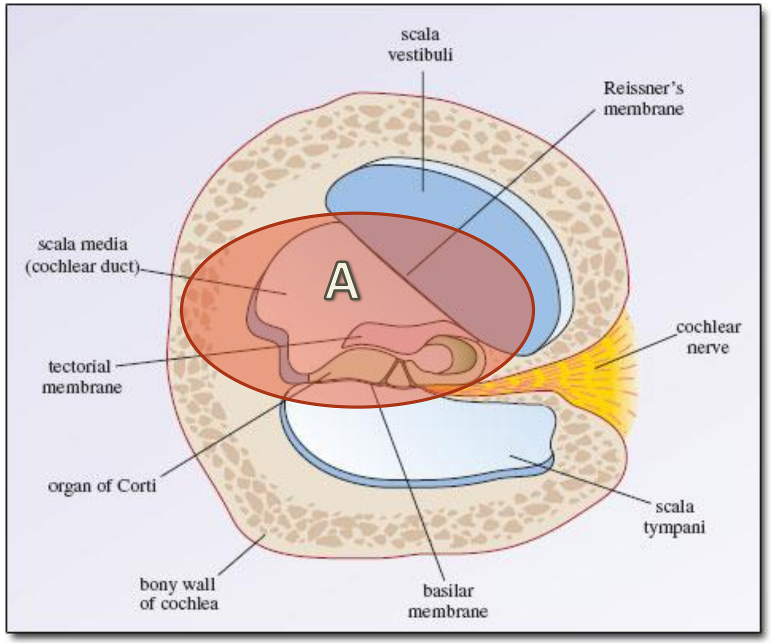
- Cochlea ✓
- Basilar membrane
- Cochlear function
- Tonotopic organization

(Semicircular canals and vestibule – for balance)

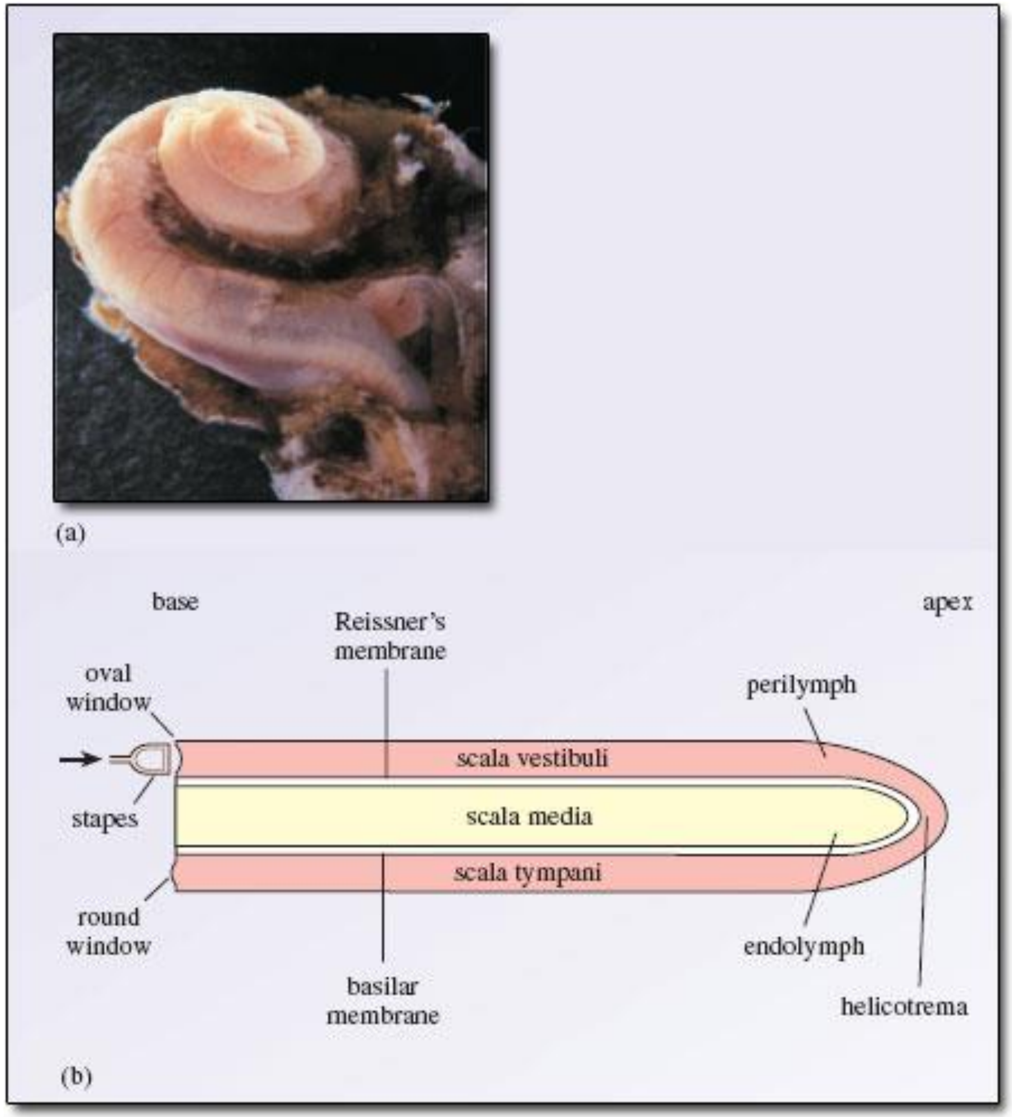


Cochlea – cross section



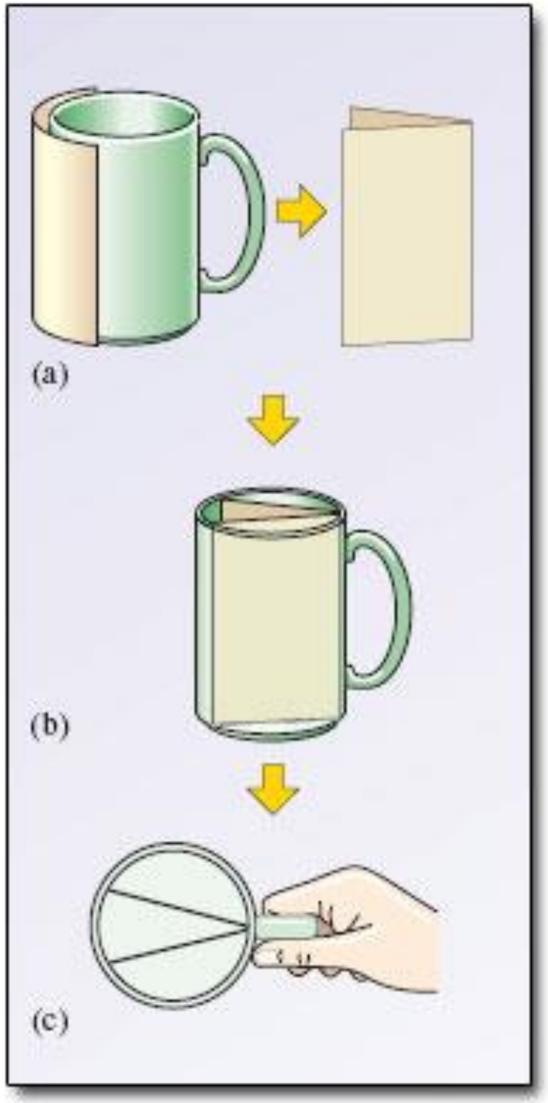


← <u>Organ of Corti</u>
(electron-micrograph) showing
4 rows of hair cells supported by
pillar cells



Cochlea - terms

- Perilymph
- Endolymph
- Cochlear duct
- Basilar membrane (BM)
- Vestibular membrane
- Round window
- Helicotrema
- Organ of Corti
- Tectoral membrane



Make your own cochlea

- You have just divided the mug into 3 chambers: the *scala vestibuli*, *media*, and *tympani* (top, middle, bottom)
- Top arm of the 'V' is *Reissner's (Vestibular) membrane,* the bottom is the *Basilar membrane*

Cochlea – continued

Broad → narrow from lateral →medial

However, BM is opposite (shown in grey), from narrow -> broad as it approaches the apex

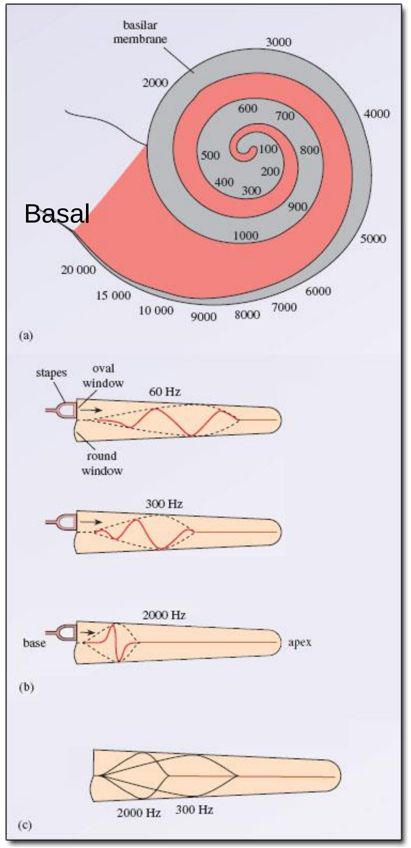
Stiffer at narrow (basal) end; 100x more compliant at wider (apical)

Tonotopic organization

Lying on BM is *Organ of Corti* (the organ of hearing!)

- 3500 inner hair cells ✓ (to 8th Nerve)

- 20,000 outer hair cells …?



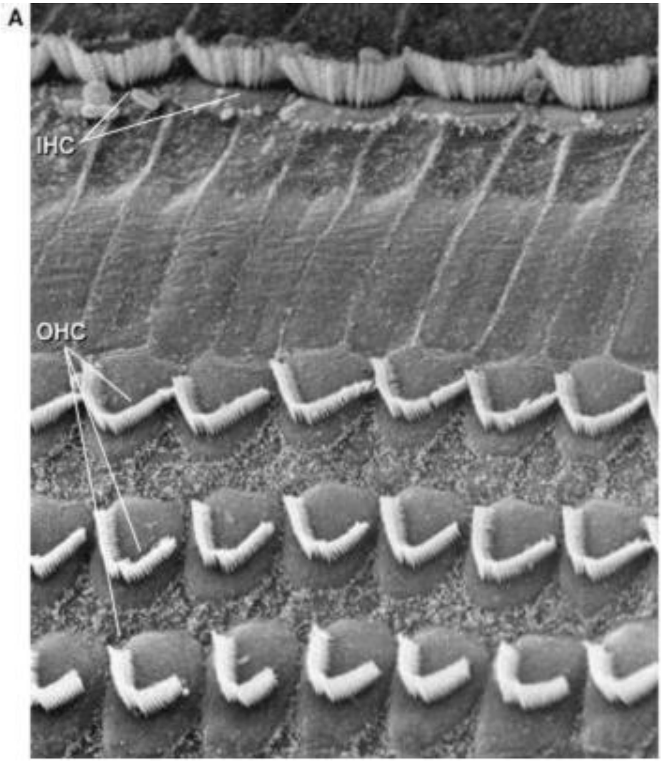
Hair cell function

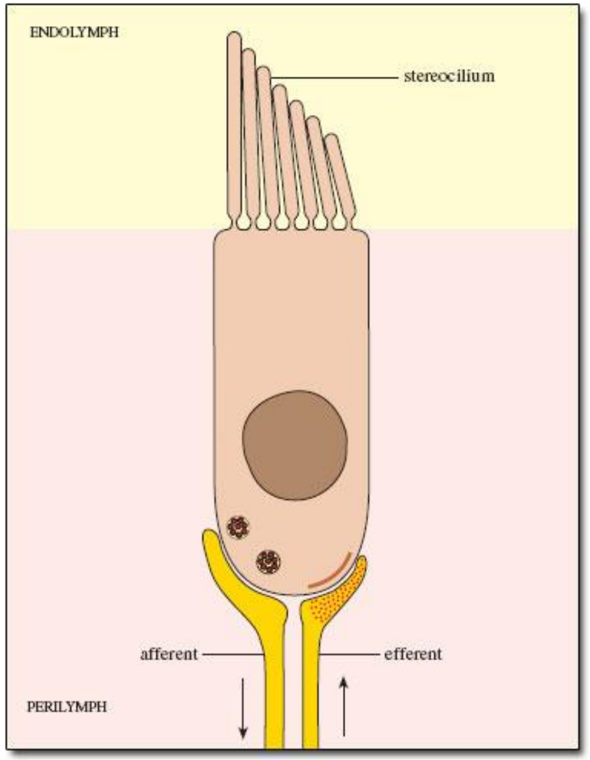
Inner

- Directly reach brain via 8th nerve
- Fewer

Outer

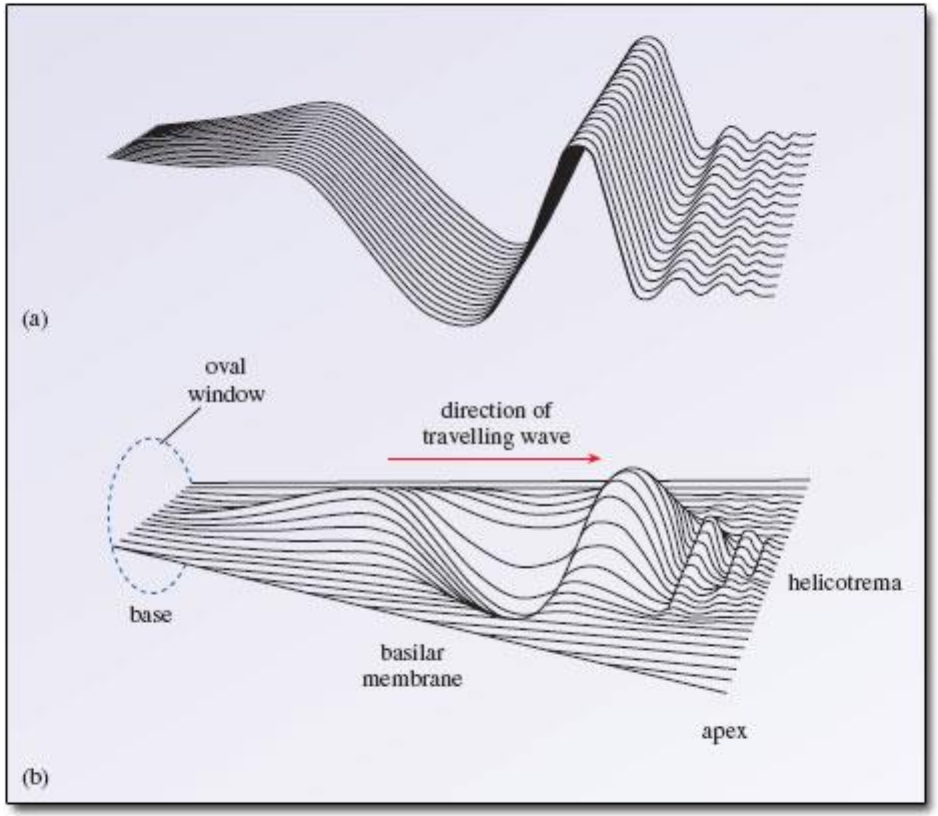
- Long a mystery
- Ear actually creates sound (external otoacoustic emissions)
- OHCs amplify them
- “Cochlear amplifier”
- Refines frequency and sensitively of mechanical vibrations


A
IHC
OHC


ENDOLYMPH
stereocilium
afferent
efferent
PERILYMPH

Hair cells

Cochlear function – continued



- Cochlea essentially performs a Fourier analysis
- With hearing loss, spectral sensitivity typically compromised, “spectral smear”

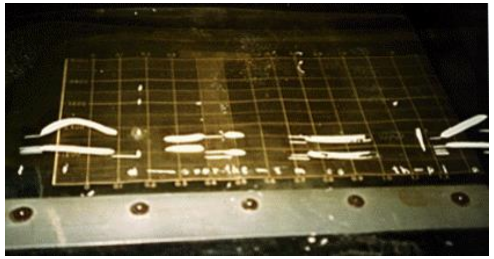
Video: https://www.youtube.com/watch?v=dyenMluFaUw

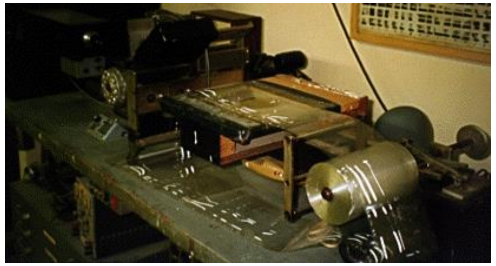
Speech Perception – some key issues

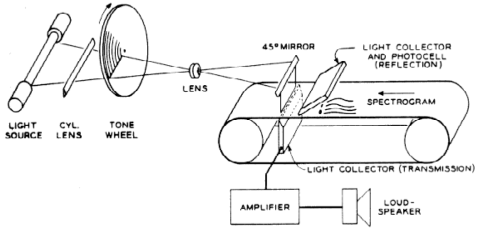
* How do we derive meaning from the stream of sounds we hear? ("the segmentation problem")

* Role of redundancy

* What are the "units?"
  - Individual sounds (phonemes)
  - Syllables
  - Words
  - Sentences



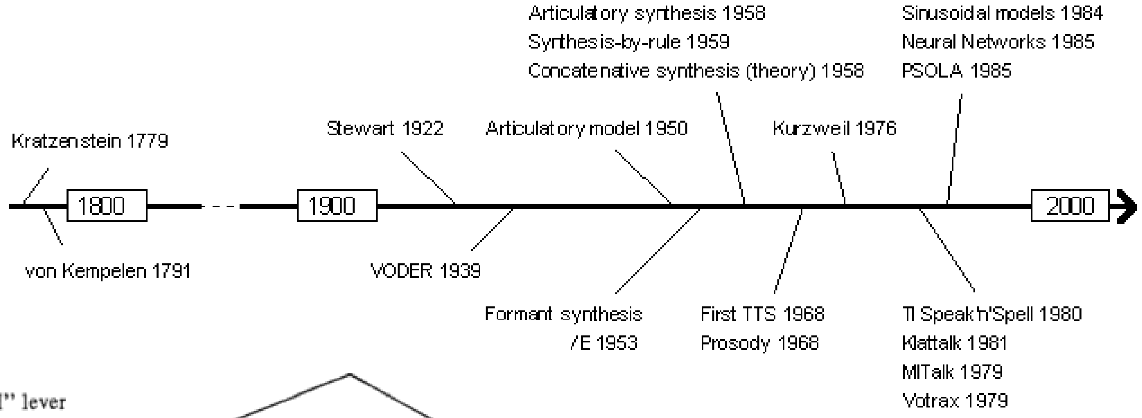


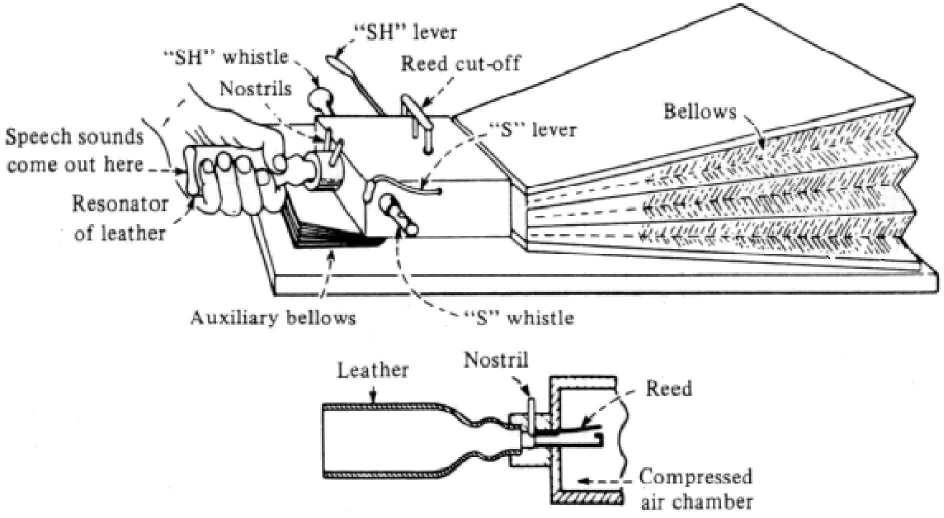


Instrumental Analysis –

History - Pattern Playback

Speech synthesis.. A long history





Perception of vowels and diphthongs

Steady state cues-

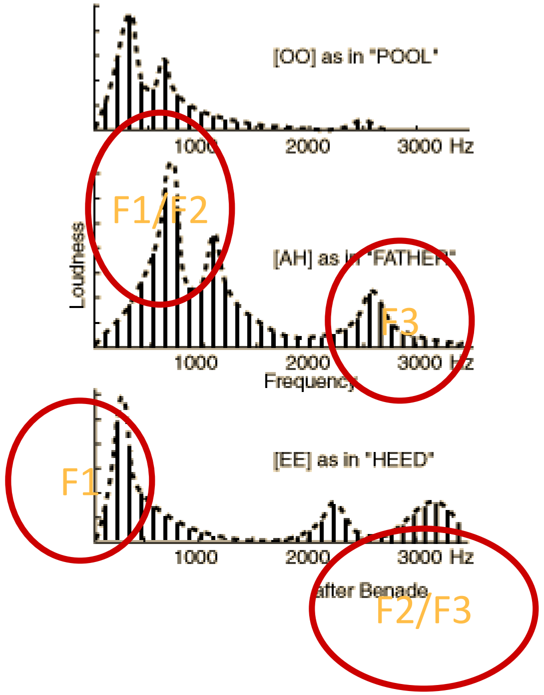
- Is it relationships between the formants that are important?
- OR Constant ratios of F2/F1?

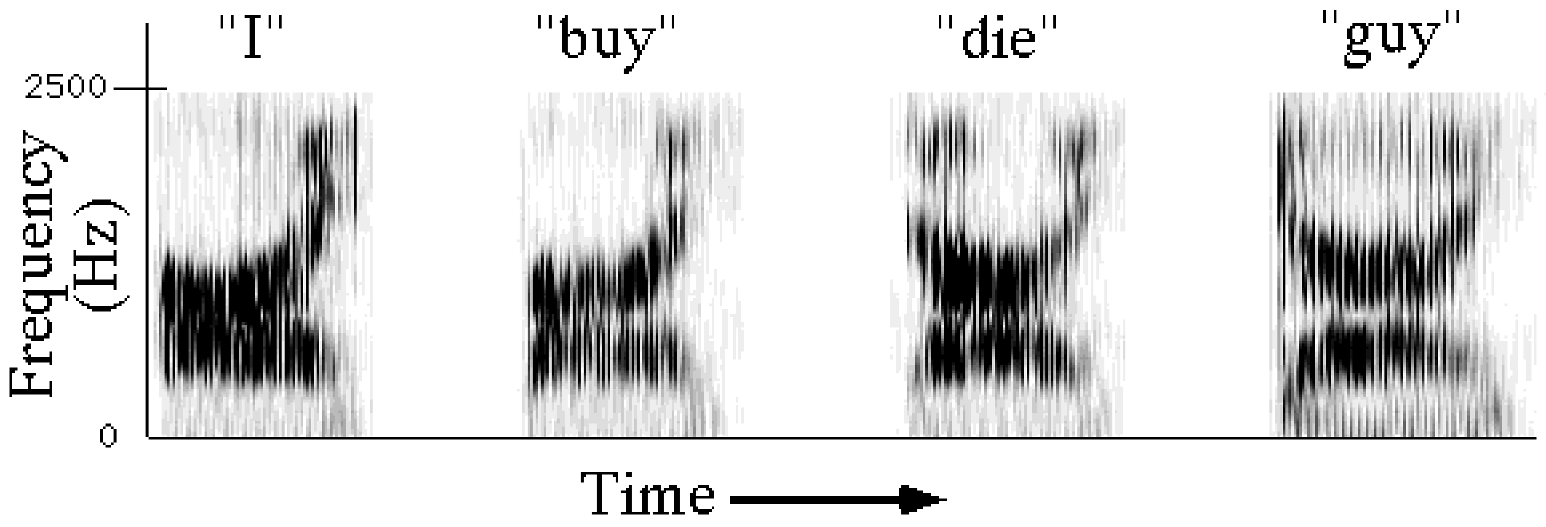
Dynamic cues –

“silent center” experiments show that even when vowel nucleus removed from syllables, listeners can still recover vowel quality

Vowel and diphthong perception – cont'd

- Back vowels (e.g. [u, a]) are perceived on the basis of the average of F1 and F2, as well as F3
- Front vowels (e.g. [i]) are perceived on basis of F1 frequency and average of F2 and F3





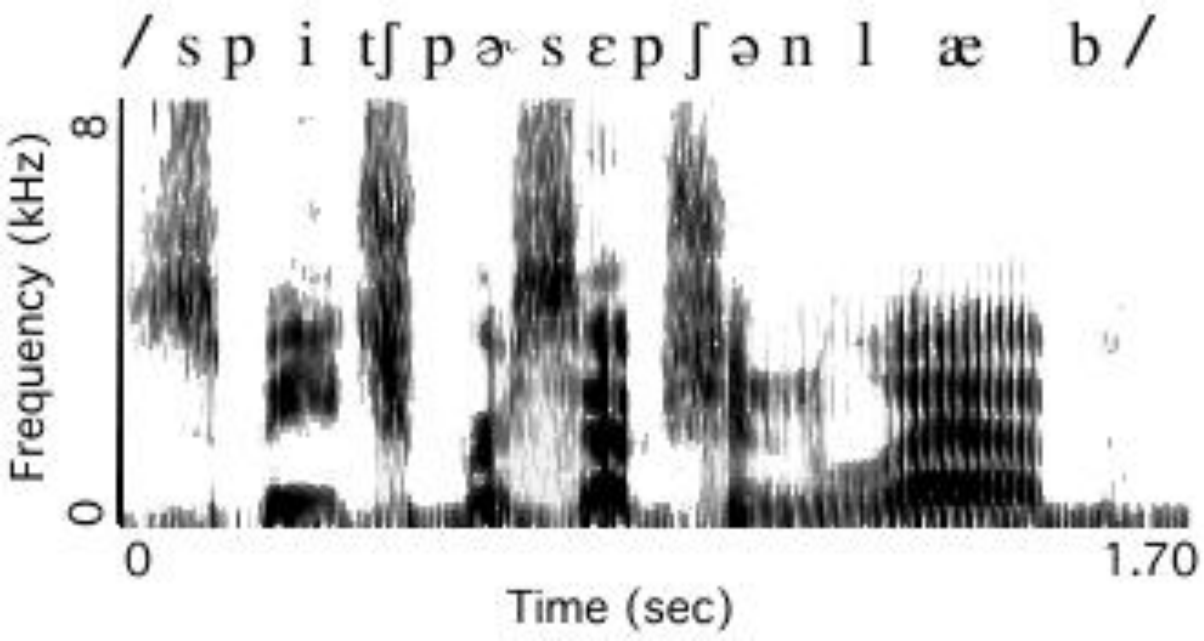
Perception of Diphthongs

Perceived on basis of their formant transitions

Salient feature: rapidity of transition

(rate of formant change)

Consonant Perception



Perception different for consonants than vowels

Greater variety of consonant types than vowels

Greater complexity for consonants

Many consonants perceived categorically

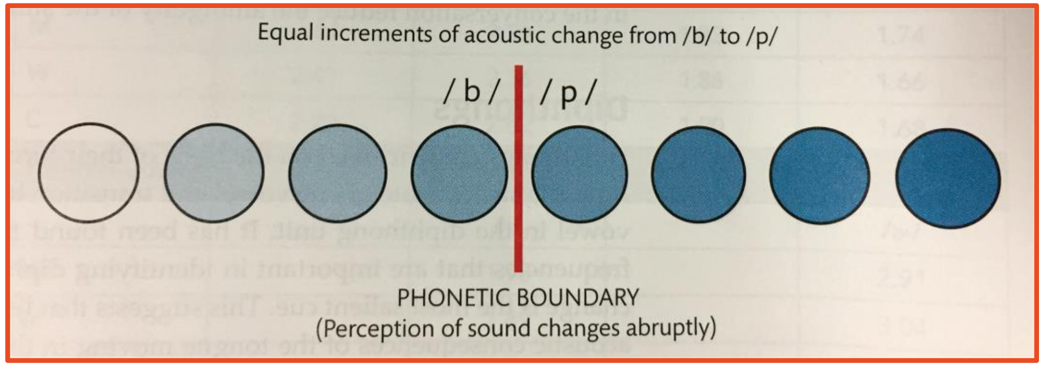
Categorical vs. Non-categorical (Graded) Perception

CATEGORICAL PERCEPTION

- "All-or-none" response.
- Listeners can identify stimuli well across categories and cannot discriminate well within categories.

GRADED PERCEPTION

- No sharp differences in identification or discrimination functions



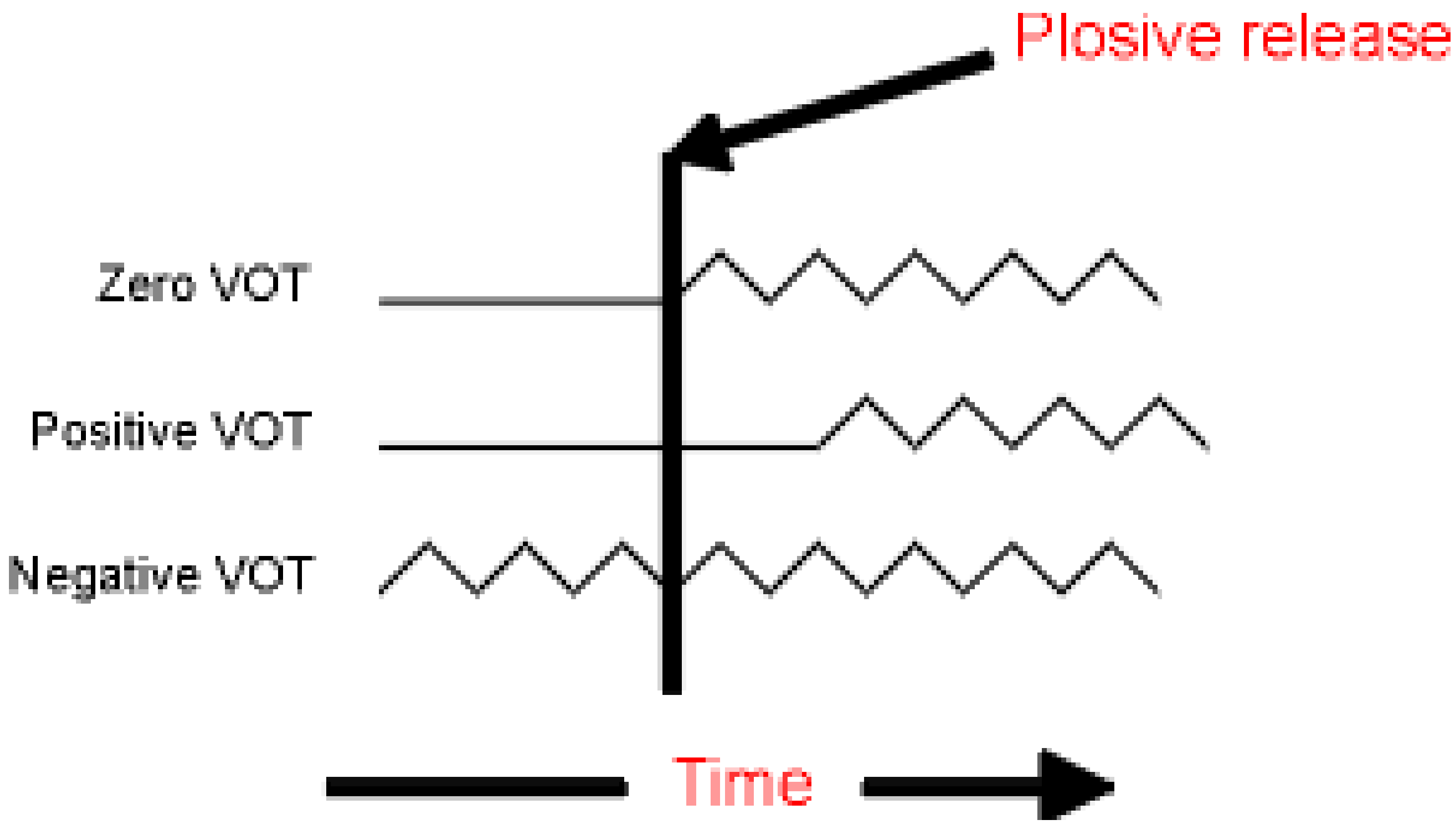
Examples of Categorical Perception in Speech

CONTRAST	MAJOR ACOUSTIC CUE
Word initial voicing in oral stops [ba-pa]	Voice Onset Time
Word final voicing in oral stops [ab-ap]	Duration of preceding vowel
Place in oral stops [ba-da-ga]	Start and direction of the second formant
Place in nasal stops [ma-na]	Start and direction of the second formant
Voicing in final fricatives [as-az]	Duration of preceding vowel
Place in fricatives [sa-ʃa]	Frequency of the turbulent noise
Liquids [la-ra]	Frequency of the third formant

Two important tasks in psychology

Identification – *"Label the sound you hear"* (arguably more complex or "higher-level" than discrimination?)

Discrimination – *"Indicate whether the two sounds you hear are same or different"*



VOT - review

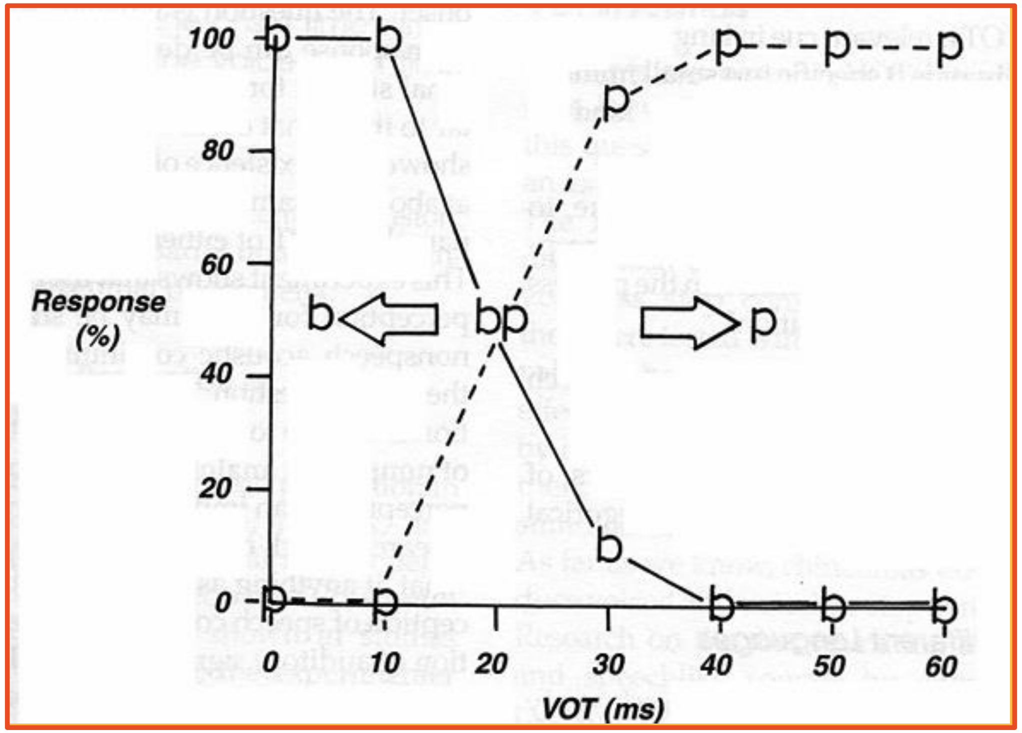
/b/ vs. /p/

/d/ vs. /t/

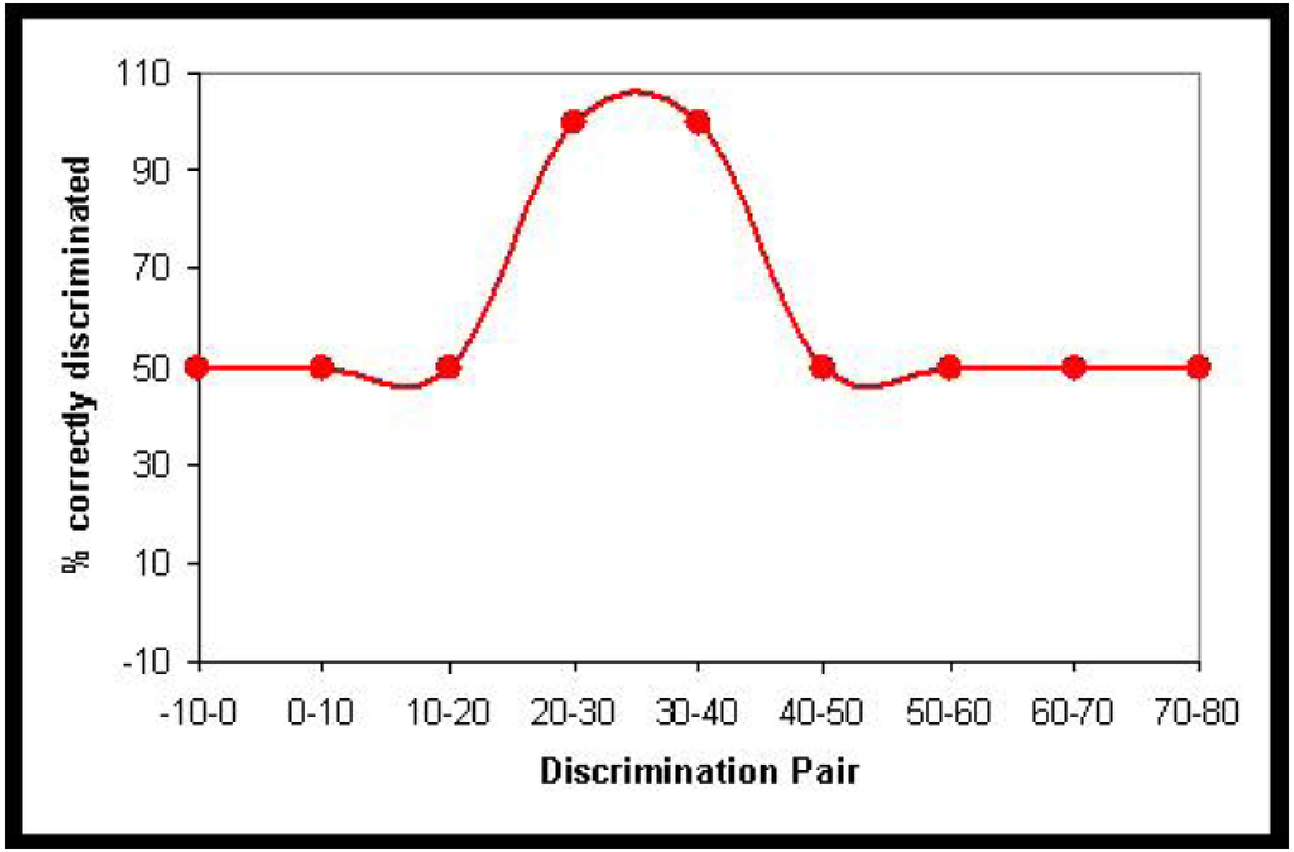
/k/ vs. /g/

VOT – an example of a categorical cue: Identification

- Important cue for voicing of syllable-initial stop consonant



Results of discrimination task: 10 msec intervals of VOT



Cross-linguistic VOT studies

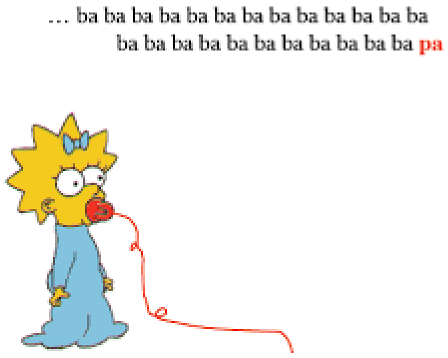
- VOT boundaries vary in different languages.
- For example, Thai has a 3-way boundary between voiceless aspirated, voiceless unaspirated, and voiced stops (*see next slide*).
- Nevertheless, different languages may show a common ground in VOT boundaries reflecting general characteristics of the auditory system
- This could relate to ‘matched’ or ‘tuned’ systems between perception and production

VOT studies in animals

- Humans talk, frogs croak
- Frogs have auditory systems tuned to their species-specific calls
- Animals appear to perceive categorically.
- Work with frogs and chinchillas shows that categorical perception is not uniquely human

VOT experiments in infants

- HASP (high-amplitude sucking) and head-turning paradigms indicate that young infants show categorical effects for VOT, much like adults [*see next slide* →]
- Infants can discriminate speech sounds of all of the worlds languages
- However, as they become experienced in their native language(s), infants ‘unlearn’ foreign sounds and hone in on only those sounds needed for their native languages



High amplitude sucking paradigm
(shortly after birth)

For more info and videos: https://www.youtube.com/watch?v=EFlxiflDk_o

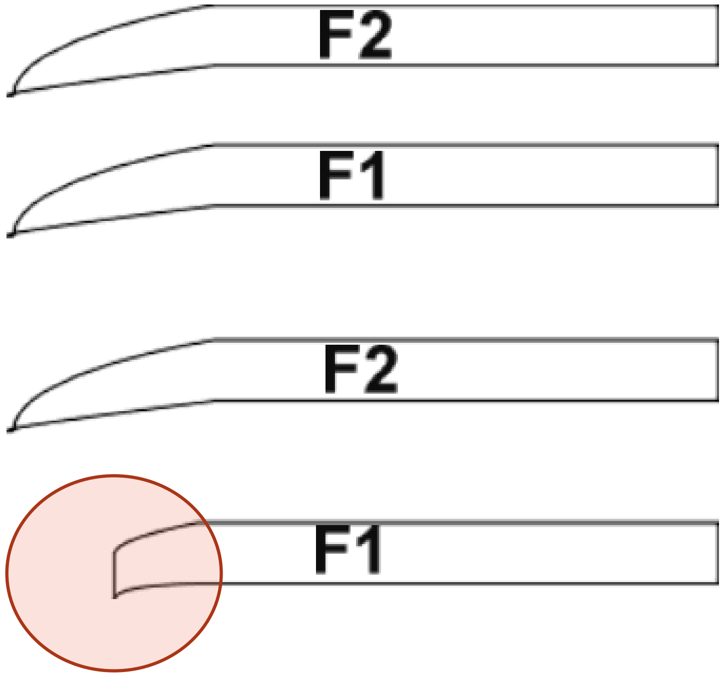
HASP (High amplitude sucking paradigm) data

Infant speech perception - summary

- Infants are born preferring the sound of their own language (through prosodic cues available *in utero*)
- They learn the phonemes of their language by ~ 9 - 12 months (by ignoring distinctions not phonemic in their language).
- Infants use stress patterns and statistical regularities to help segment words from continuous speech.

Multiple acoustic cues in consonant perception

Example: Voicing cue in syllable-initial stops 'F1 cutback'



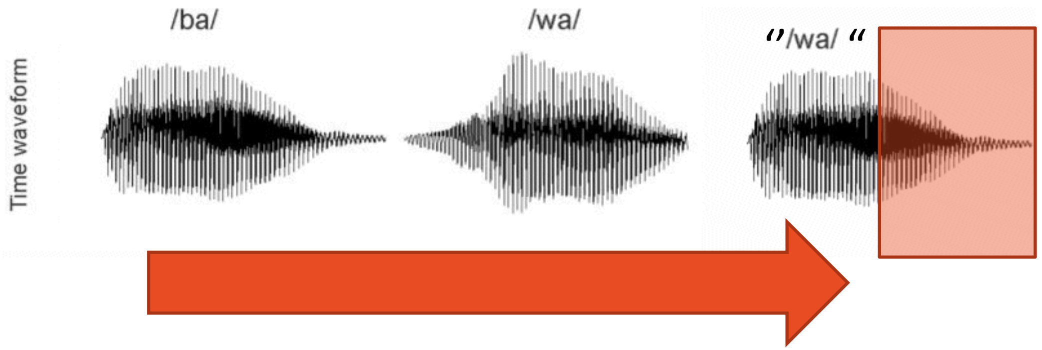
/ba/

/pa/ with F1 cutback

Another example – manner of articulation

<u>Example</u>:

- Stop vs. glide distinction
- “ba” vs. “wa”
- If the vowel portion of the syllable /ba/ is truncated, the consonant will sound more like ‘wa’



(← a “/ba/” with end of syllable cut off)

Chapter 9 Clinical Applications

- HEARING LOSS
- AUDIOMETRY, OTOACOUSTIC EMISSIONS, BRAINSTEM TESTING
- COCHLEAR IMPLANTS
- OM AND EFFECTS ON SPEECH PERCEPTION
- LANGUAGE, READING DISABILITIES AND SPEECH PERCEPTION
- RELATION BETWEEN SPEECH PERCEPTION AND ARTICULATION

Types of hearing loss

- <u>TYPE</u>: Outer ear (conductive), sensorineural*, or mixed
- <u>DEGREE</u>:

mild	moderate	moderately severe	severe	profound
25 – 40 dB				> 90 dB

- <u>CONFIGURATION</u>:
- flat
- <u>rising</u> (lower frequencies more affected than higher)
- gradually or precipitously <u>sloping</u> (higher more affected than lower)

**e.g., noise exposure, aging, genetics*

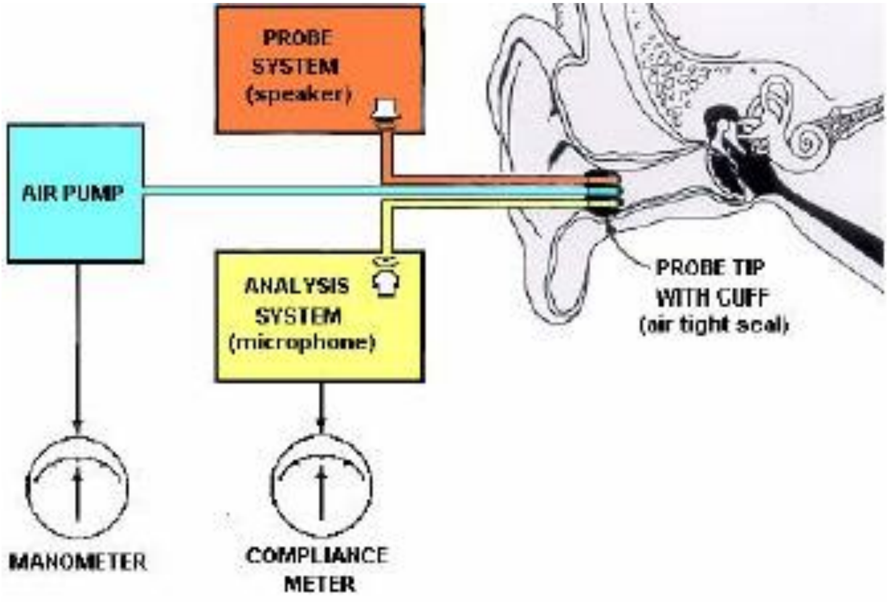
Immitance audiometry

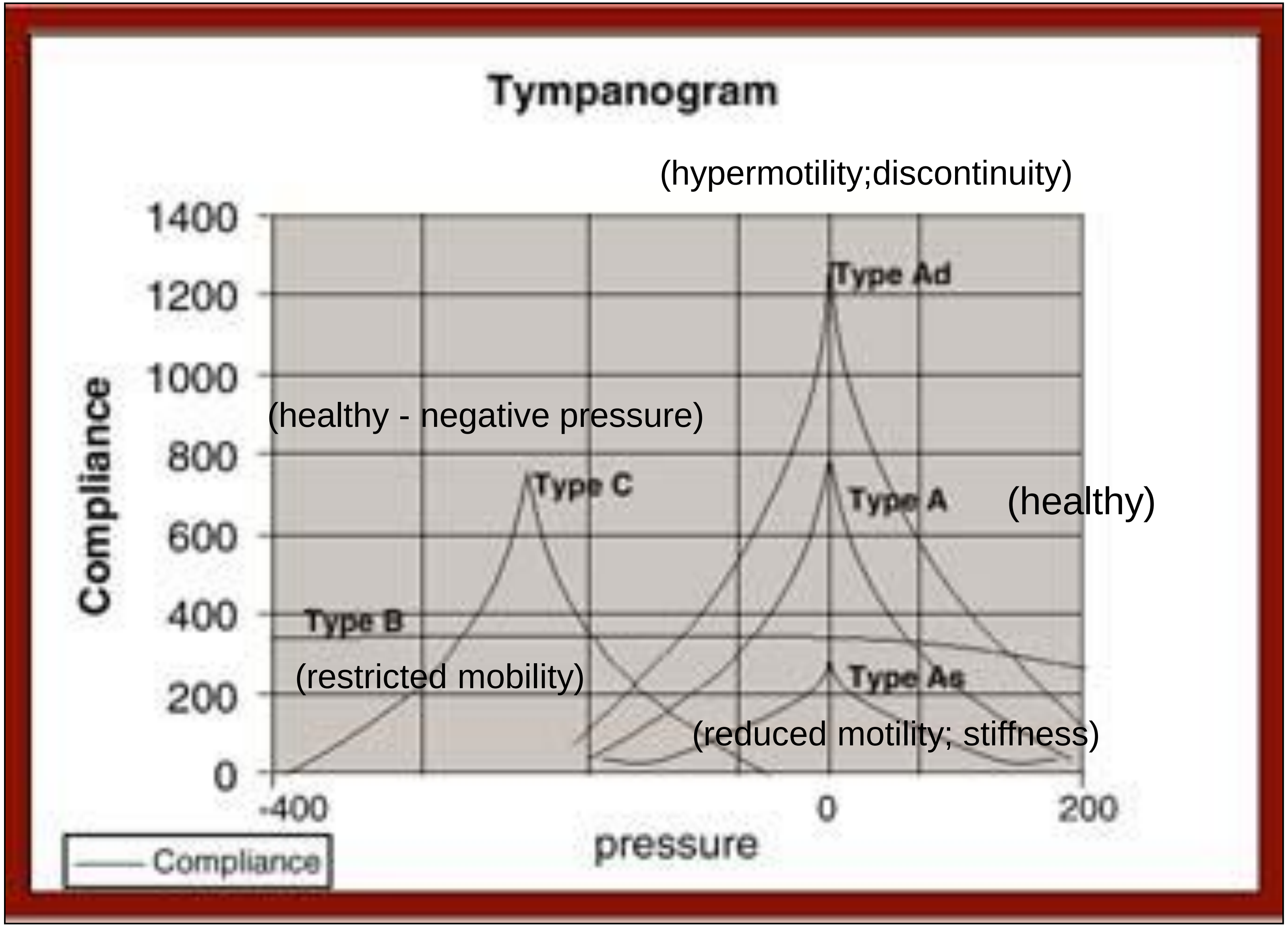
Im(pedence) + Ad(mittance)

Tympanometer

For ME problems

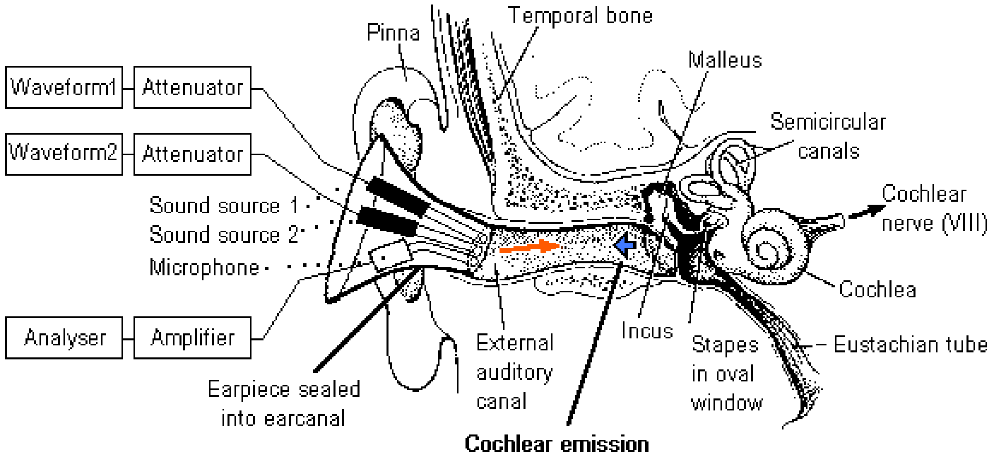
High admittance (*milliohms*) = low impedence



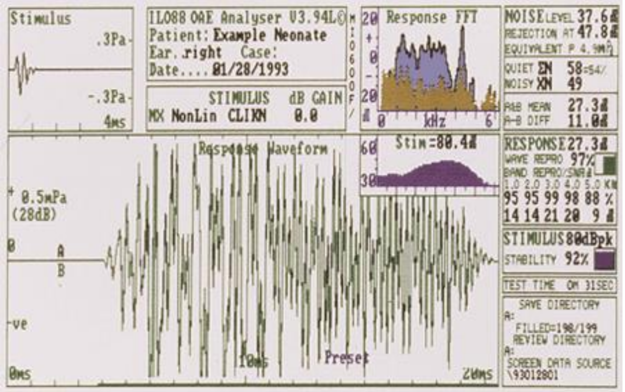

Tympanogram
(hypermotility;discontinuity)
1400
1200
1000
800
600
400
200
0
Compliance
Type Ad
(healthy - negative pressure)
Type C
Type A
(healthy)
Type B
(restricted mobility)
Type As
(reduced motility; stiffness)
-400
0
200
pressure
Compliance

Otoacoustic emissions testing

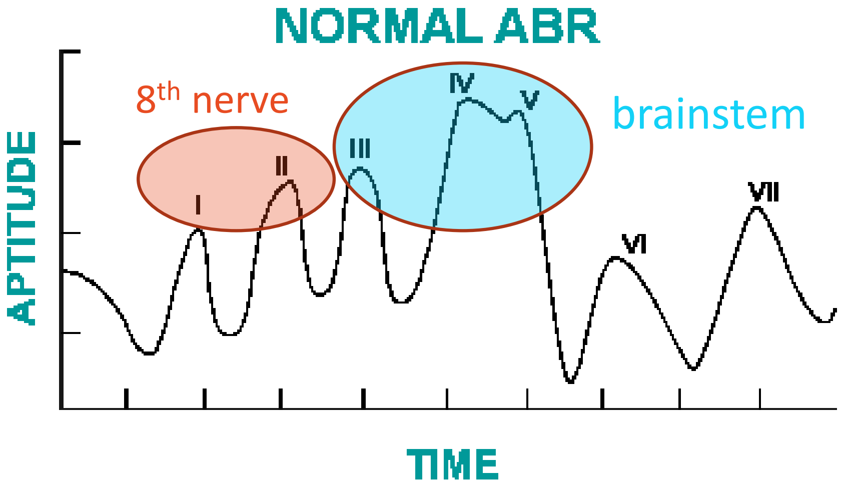
- Low intensity sounds – present in healthy ears
- Can be spontaneous (SOAEs), or evoked (EOAEs), e.g. with clicks
- Diagnostic of sensory hearing loss
- Easily tested, even in babies



https://www.youtube.com/watch?v=QvrBogzziXA



Auditory brainstem response (ABR) testing



- Used to identify hearing loss type and degree
- Can be done during sleep
- Brief video:
- https://www.youtube.com/watch?v=QTKvtKYLlQ8

HI and speech perception

- Vowel perception does not pose as great a problem as consonant impairment (esp. fricatives!)
- Place of articulation most common perceptual errors (because rapid and high frequency)
- Task of identification esp. difficult (e.g. compared to discrimination)
- For adults, likely due to redundancy of acoustic cues in the signal
- For children, Vs with high frequency components (high vowels, e.g. /i/ <-> /u/ confusion) may pose more problems

Consonant perception in HI – factors

- Audibility – whether a specific speech cue is presented at a level that the person can hear (=Suprathreshold level)

- Hi lack of recognizing speech cues as audible suggests possible lack of ability to use dynamic information(?)

Evaluation of speech perception

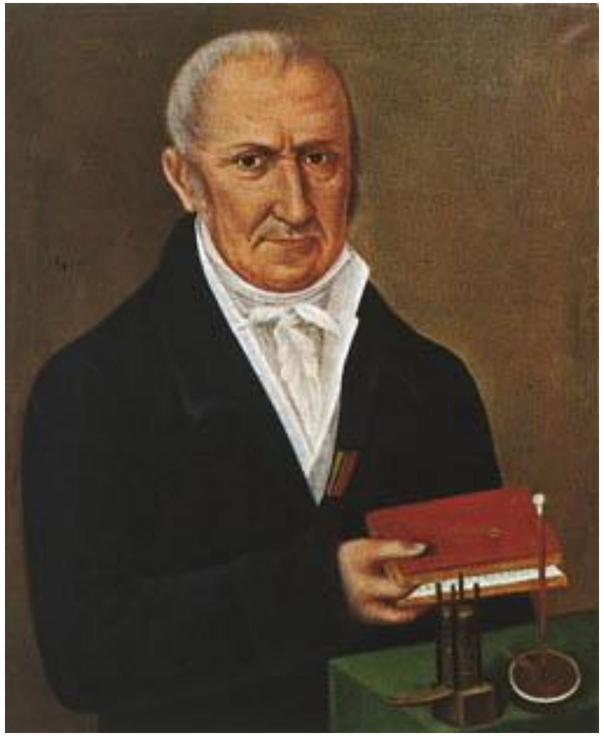
- Articulation index (AI) used to predict amount of speech audible to patients with HI
- Range: 0-1
- Some suggest counting 100 dots on an audiogram (between 1-3 kHz)

CI/Amplification/ New directions

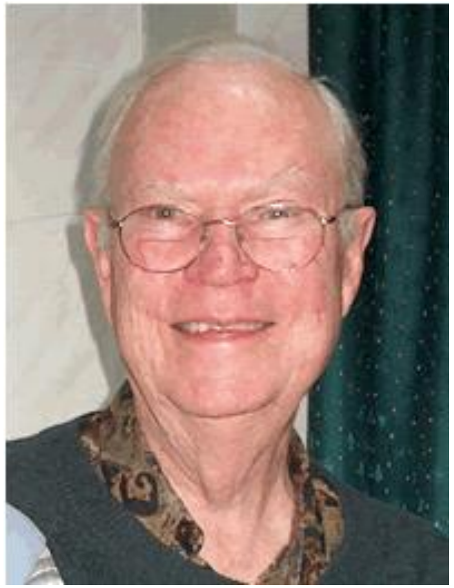
- Amplification with FM (wireless), directional microphones
- Cued Speech
- Microphone arrays for directional hearing(?)
- CIs offer one treatment option – research is expanding at a rapid pace

Cochlear implants - origins

- Count Volta (Italian physicist) who developed the electric battery, connected batteries to two metal rods ***that he inserted in his ears.***
- In 1800 he described that when the circuit was completed he received a *'jolt in the head'* and then a sound *'a kind of crackling, jerking or bubbling as if some dough or thick stuff was boiling'*



Cochlear implants - developments



House Ear Institute
Advancing Hearing Science

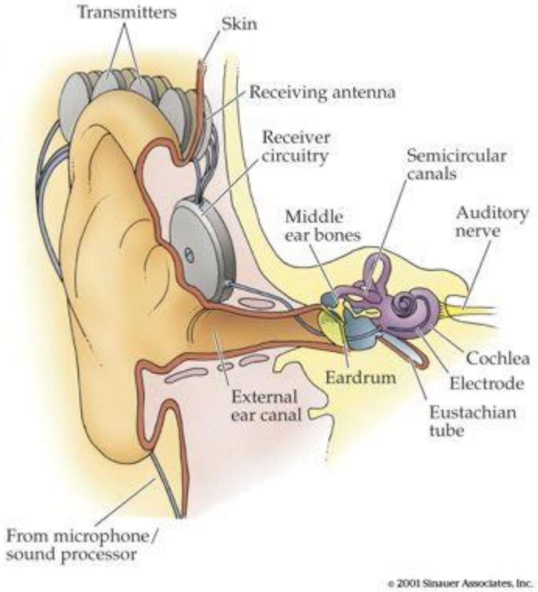
Developed and implanted in 1960s by Dr. William House (USA) and expanded by Jack Urban

Cochlear

Additional development/marketing by Graeme Clark of Australia in 1970sm

Modern CIs

Cochlear implant



Cochlear implants - demo

https://www.youtube.com/watch?v=00WOao4kpwM

- Michael Dorman, Ph.D. (ASU)
- Tones
- Speech

CI Speech Coding Strategies

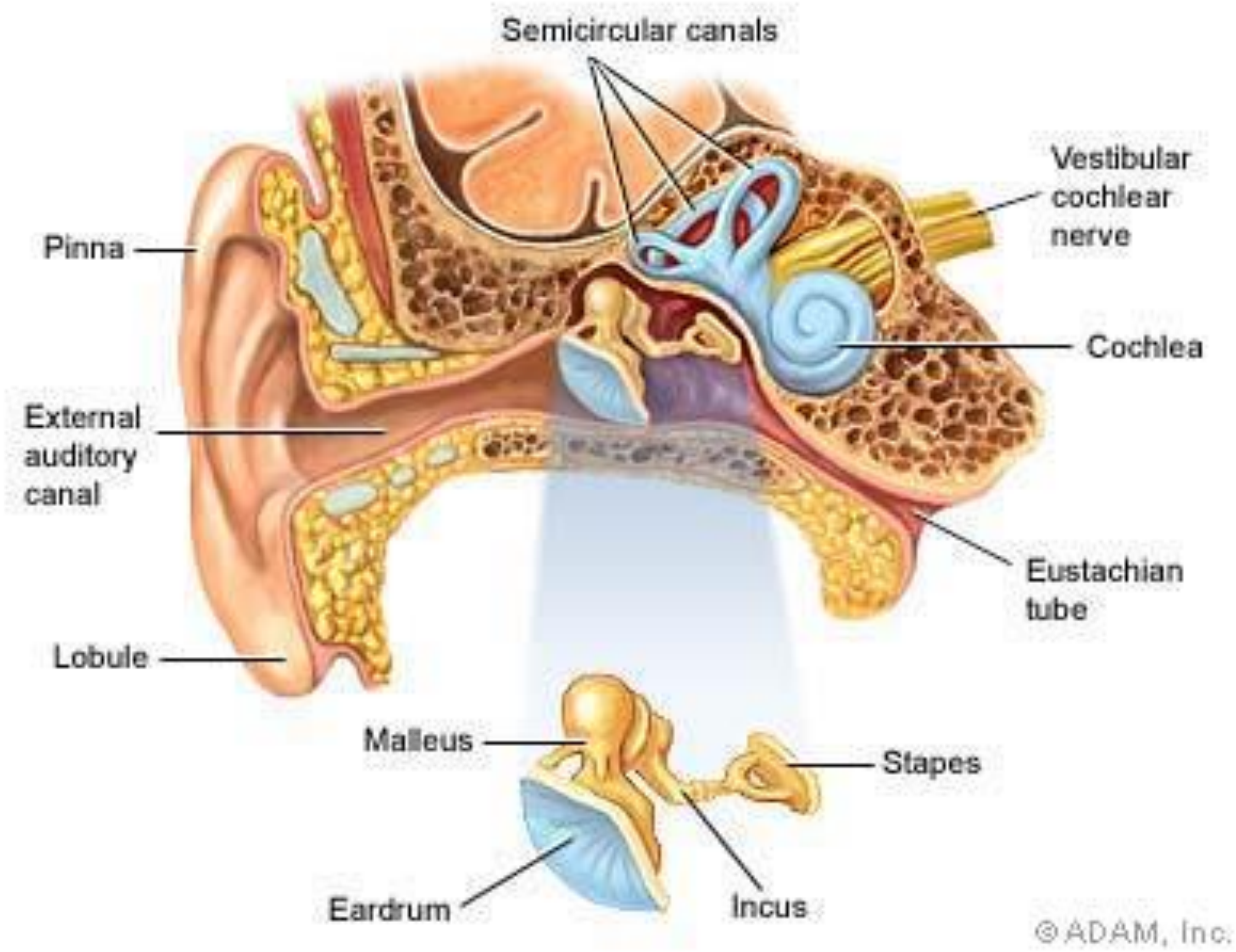
- **ACE™:** Unique to Cochlear's Nucleus 24 CI system. ACE optimizes detailed pitch and timing information.
- **SPEAK:** (spectral peak or 'M-of-N') Increases richness of important pitch information by stimulating electrodes across the entire electrode array.
- **MPEAK:** multipeak
- **CIS** : (Continuous-Interleaved Sampling) This high rate strategy uses a fixed set of electrodes. Emphasizes detailed timing information of speech.

A challenge for CI technology

Fundamental frequency resolution:

Q: What about speakers of tone languages?

A: Current practice: use CIs that minimally damage cochlea, then boost low freq auditory amplification



Otitis Media

- OME
- ME that cannot be ventilated, causing retraced tympanic membrane →
- Approx 20 – 40 dB loss
- Phonetic processing difficulties

Chronic OME in children

SOUND DEMOS: https://acoustics.org/pressroom/httpdocs/133rd/2paaa2.html

Small group studies in the 1990s suggest subtle disturbances of speech perception:

-Perceptual weighting strategies

- Phonemic awareness

Subsequent large-scale NIH funded, multi-site studies have been less conclusive

Language and reading disability

“Specific Language Impairment (SLI)” →

Or, more currently:

DLD = Developmental Language Disorder

-Temporal processing disorder?

-Difficulty with low phonetic substance?

OR:

- Working memory limitations?

- Syntactic difficulties?

ALSO: How “specific” is it, really?

(related deficits also found in Developmental dyslexia)

(Bruce Tomblin, Univ. Iowa)

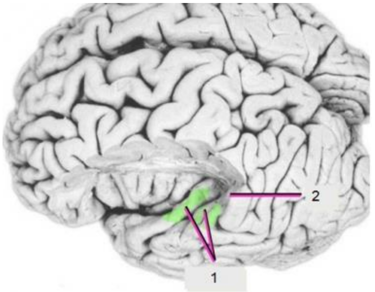
(Larry Leonard, Purdue)

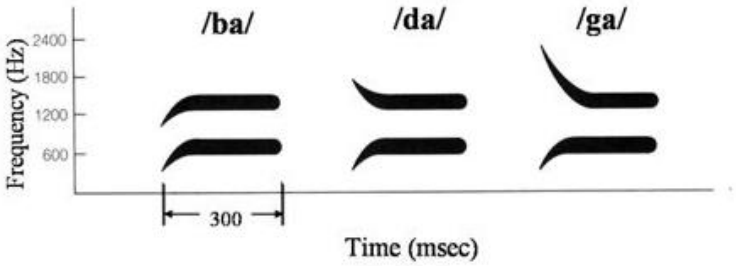
Photography by John Emerson

(Paula Tallal, Rutgers Med. School)

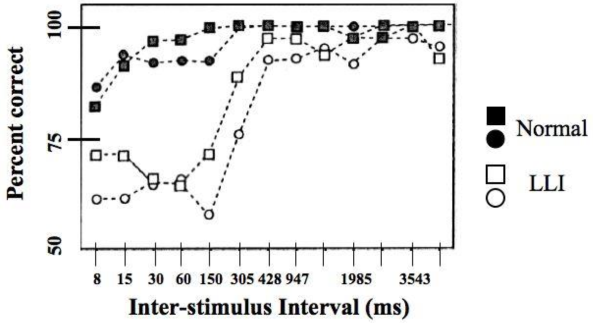
Temporal Processing Disorder/ Details

/ba/ /ba/ /ba/ … /da/…/ga/





Rapid auditory discrimination and language learning impairment



Normal Development

- Younger children rely more on auditory discrimination of small acoustic differences than older children
- A “perceptual weighting shift” may take place

Links between perception and production

Q: Do children with articulation problems have problems with the same processes in perception?

Synthetic speech experiments have been used to address this question

A: Rather mixed findings to date. There may be subgroups.

Tinnitus

https://www.youtube.com/watch?v=5bTfb3JxioU

Current UTD research