ECE 598: The Speech Chain

Lecture 10: Auditory Physiology
Today

- Outer Ear: Sound Localization
- Middle Ear: Impedance Matching
- Basilar Membrane: Frequency Analysis
  - Mechanical Principles
  - Frequency Response of Auditory Filters
  - Nonlinearity of Basilar Membrane Response
- Mechano-Electric Transduction
  - Inner and Outer Hair Cells
  - Neuro-transmitter Uptake Models
  - Neural Activation Thresholds
Localization of Sound: Inter-Aural Time Delay (ITD)

\[ \text{ITD} = \left( \frac{r}{c} \right) \left( \frac{\pi}{2} - \theta + \cos \theta \right) \]

- Wavefronts (lines of constant pressure)
- Wave traveling direction
- Diffusion of sound around the head
Localization of Sound: Inter-Aural Amplitude Difference

- $f < \frac{c}{4r} \sim 1\text{kHz}$: head $<< \lambda = \frac{c}{f}$; sound diffuses around head
- $f > \frac{c}{2r} \sim 2\text{kHz}$: head $> \lambda$, so sound is blocked by the head

Low frequency, Long wavelength (shown: wave “troughs”, i.e., pressure minima)  
High frequency, Short wavelength (shown: wave “troughs”, i.e., pressure minima)

Sound diffuses around the head; no shadow  
Head shadow: Sound unable to diffuse around large obstacle
Localization of Sound: Echoes from the Pinna and Shoulders
Localization of Sound, Summary: Head-Related Transfer Function

- Source:
  \[ s(t) = \cos(\omega t) \]

- Received at near ear:
  \[ x_R(t) = A_R(\omega) \cos(\omega t + \phi_R(\omega)) \]

- Received at far ear:
  \[ x_L(t) = A_L(\omega) \cos(\omega t + \phi_L(\omega) - \omega \tau_{ITD}) \]

- Near ear frequency response:
  \[ H_R(\omega) = A_R(\omega)e^{j\phi_R(\omega)} \]

- Far ear frequency response:
  \[ H_L(\omega) = A_L(\omega)e^{j(\phi_L(\omega) - \omega \tau_{ITD})} \]
Middle Ear Functions

- **Impedance Matching:**
  - Sound transmission in water (inner ear) requires much higher pressure than sound transmission in air (outer ear).
  - Without middle ear, sound incident on oval window would bounce away ($\gamma = 1$)
  - Middle ear reduces $g$ so that not all sound is reflected

- **Reduce Exposure to Loud Environments**
  - Strap muscles loosen in loud environments, reducing the amplitude of sound transmitted to inner ear
  - Effect is relatively slow (hundreds of milliseconds), so not useful for adaptation to rapid sounds (gunshots)
Impedance Mis-Match Between Water and Air: Without a Middle Ear, What Would You Hear?

Continuity of pressure at the boundary:
\[(p_{a+} + p_{a-}) = (p_{w+} + p_{w-})\]

Continuity of volume velocity at the boundary:
\[\left(\frac{A}{\rho_a c_a}\right)(p_{a+} - p_{a-}) = \left(\frac{A}{\rho_w c_w}\right)(p_{w+} - p_{w-})\]

Densities:
\[\rho_a = 0.001 \text{ g/cc}, \quad \rho_w = 1 \text{ g/cc}\]

Speeds of Sound:
\[c_a = 354 \text{ m/s}, \quad c_w = 1000 \text{ m/s}\]

Suppose \(p_{w-} = 0\), meaning that the only input sound is \(p_{a+}\)

Then...

The reflected sound is:
\[p_{a-} = \frac{(\rho_w c_w - \rho_a c_a)}{(\rho_w c_w + \rho_a c_a)} p_{a+} = 0.9994 p_{a+}\]

The transmitted sound is:
\[p_{w+} = \frac{2\rho_a c_a}{(\rho_w c_w + \rho_a c_a)} p_{a+} = 0.0006 p_{a+}\]
Lever system reduces the effective input impedance \((z = \frac{p}{v})\) of water by a factor of \(L^2\)

Resulting reflection coefficient

\[
\gamma = \frac{(\rho_w c_w - L^2 \rho_a c_a)}{(\rho_w c_w + L^2 \rho_a c_a)} \sim 0.98 - 0.99 < 1
\]
Acoustic Impedance of Ear Canal Informative About Middle Ear Function

- Remember how to calculate impedance?
  1. Impose a Boundary Condition at Far End
     \[ p_a^- = \gamma p_a^+ \]
  2. Calculate \( z = p/v \) at Near End
     \[
     z = \frac{p}{v} = \frac{\rho c (p_a^+ e^{-jkx} + p_a^- e^{jkx})}{(p_a^+ e^{-jkx} - p_a^- e^{jkx})} = \frac{\rho c (e^{2jkL} + \gamma)}{(e^{-2jkL} - \gamma)}
     \]

- So by measuring the acoustic input impedance of the auditory canal very precisely, it’s possible to deduce \( \gamma \) at different frequencies, and thus to learn something about health of the middle ear (product: Mimosa Acoustics)
Inner Ear Anatomy
(image courtesy Alec Salt, Otolaryngology, Washington University)
Inner Ear Anatomy: Charged Fluids
(image courtesy of Alec Salt, Otolaryngology, Washington University)
Cross-Section of the Basilar Membrane
(image courtesy wikipedia)
Frequency Selectivity of Places on the Basilar Membrane

Basilar Membrane (separates scala media & scala tympani)

Oval Window

Unroll

Base, x=0mm:
High Stiffness
Low Mass
\( f_c = \frac{\sqrt{k/m}}{2\pi} \)
\( \sim 16000\text{Hz} \)

Apex, x~3cm:
Low Stiffness
High Mass
\( f_c = \frac{\sqrt{k/m}}{2\pi} \)
\( \sim 40\text{Hz} \)

In between: Each position, x, is tuned to a different mechanical resonance
\( x(f_c) \sim 30\text{mm} - (11\text{mm}) \ln(1 + 46f_c/(f_c+14700)) \)
Frequency Selectivity of Places on the Basilar Membrane

Wave $p_w e^{-j\omega x/c}$ propagates forward at $c=1000\text{m/s}$ until...

Wave energy is absorbed by oscillation of the basilar membrane at $x(f_c=\omega/2\pi)$
Frequency Selectivity of Places on the Basilar Membrane

Traveling waves in the cochlea.

"Concerning the pleasures of observing, and the mechanics of the inner ear,"

Nobel Lecture, 1961,
Georg von Békésy

(courtesy of Pacific Biosciences Research Center Hawaii)
Bandwidth of the Auditory Filters:

100Hz at $f_c < 500$Hz, 0.2$f_c$ at $f_c > 500$Hz

(image courtesy Julius Smith and Jonathan Abel, CCRMA, Stanford)

Equivalent Rectangular Bandwidth (ERB) =

Bandwidth of an ideal BPF that passes the same total energy as the basilar membrane section at the same $f_c$
Velocity of Basilar Membrane Causes Inner Hair Cell Follicles to Bend
Velocity of Basilar Membrane Causes Inner Hair Cell Follicles to Bend
Bending of Follicles Causes Depolarization of IHC

 Scala Media (+80mV)

 Cations enter through follicle tips when follicles bend

 Organ of Corti (0mV)

50mV
Depolarization of IHC Causes Release of Neurotransmitter

Cations enter through follicle tips when follicles bend

Neurotransmitter released

Synapse, afferent neuron
Neurotransmitter Dynamics
(Three-Store Model: Meddis, JASA 1986)

Result: probability of neuron firing is a smoothed (lowpass filtered) version of the IHC voltage

Neurotransmitter release (instant)
Neurotransmitter binding (several ms)

Neurotransmitter Re-uptake (several ms)
Signal Processing in the Inner Ear (Simulated)

Acoustic Signal: "Three"

Basilar Membrane Velocity ($f_c=500\text{Hz}$)

IHC Follicle Bending

Neuron Probability of Firing
Neural Response to a Synthetic Vowel
(Cariani, 2000)