Speech Science
COMD 6305
Weds/Fri 1:00pm - 2:15pm

William F. Katz, Ph.D.
UTD
Chapter 1.
The Nature of Sound

INCLUDING

• Definition of Speech Science
• Basic physics concepts
• Overview of sound
• Resonance
What is ‘speech science’?

• “Speech Science is the experimental study of speech communication, involving speech production and speech perception as well as the analysis and processing of the acoustic speech signal.”

• Speech Science asks questions like:
  – How is speech planned and executed by the vocal system?
  – How do the acoustic properties of sounds relate to their articulation?
  – How and why do speech sounds vary from one context to another?
  – How do listeners recover the linguistic code from auditory sensations?
  – How do infants learn to produce and perceive speech?
  – How and why do speech sounds vary between speakers?
  – How and why do speech sounds vary across speaking styles or emotions?

(-M. Huckvale, UCL)
Physics Principles

**Energy**

Ability to do work*

*(force to move a mass some distance in some amount of time)*

Common Units:

- Mass = kilograms (kg)
- Distance = meters (m)
- Time = seconds (sec)
Physics Principles - cont’

Opposing Forces
- Affect objects in motion
1. Friction
2. Inertia
3. Elasticity
Friction

• Impedes or opposes movement

(not much friction here....)
In order to vibrate, an object must possess

1) Inertia
2) Elasticity

*Any object that can vibrate can produce sound
Inertia

“Tendency for mass at rest to remain at mass”
or
“Object in motion to remain in motion”

(Note: Amount of inertia is related to mass)
Elasticity:

- Restoring force
- Ability of a substance to recover its original shape and size after distortion
- All solids are elastic; gases also behave as if elastic
**Hooke’s law** – describes **elasticity** - “restoring force is proportional to the distance of the displacement and acts in the opposite direction” 

(if not stretched beyond its elastic limit)
Density

(= mass per unit of volume)
Pressure

• Pressure = Pa
• Pa = N/m², or dyne (older literature)
• N = kg \times \text{meter} \times \text{s}^2

(where N = Newton, a measure of force)

See later slide for other measures of pressure!
Measurement of air pressure

• Dynes* per sq centimeter  dyne/cm²
• Pounds per square inch    psi
• Microbar                   μbar
• Pascal                     μPa
• Centimeters of water      cm H₂O
• Millimeters of mercury    mm Hg

*The force required to accelerate a mass of one gram at a rate of one cm/sec²
Air pressure can also be measured in \( \text{cm of H}_2\text{O} \) or \( \text{mm of Mercury (Hg)} \).
Two Units of Measure (pressure)

dynes/cm$^2$  “microbar”  (CGS system)
N/m$^2$  “Pascal”*  (MKS system)

*(For speech, micropascal is more commonly used)
Pressure at different locations may vary

- \( P_{\text{atmos}} \)
- \( P_{\text{oral}} \)
- \( P_{\text{trach}} \)
- \( P_{\text{alveolar}} \)
Air

• Composed of gas molecules
• Particles travel in Brownian motion
• Has pressure (force per unit area)
  – Pressure related to speed of Brownian motion
  “rapid changes in relatively static atmospheric pressure we call sound pressure or sound energy”
Air Movement

• **driving pressure**: (difference in pressure) high pressure FLOWS to low pressure

• **volume velocity**: rate of flow – e.g. `liter/sec`

• **laminar flow** - in a parallel manner

• **turbulent flow** – in a non-parallel manner (flows around an object - eddies)
Flow
Air Pressure, Volume, Density

- **Volume**: amount of space in three dimensions
- **Density**: amount of mass per unit of volume
- **Boyle’s law**: as volume decreases, pressure increases

*(assuming constant temperature)*
What is “sound?”

- Form of energy
- Waves produced by vibration of an object
- Transmission through a medium (gas, liquid, solid)
- Has no mass or weight
- Travels in longitudinal wave

text pg. 7
Transverse vs. Longitudinal Wave

Transverse

Direction of wave
Crest
Trough

Longitudinal - SOUND!

Direction of wave
Compression
Rarefaction

DIRECTION OF WAVE TRAVEL

Cool!
Air Pressure Changes from Sound

- Compression
- Rarefaction
Propagation of Sound

C = Condensation
R = Rarefaction

DEMO: Longitudinal Wave
Periodic waves

- Simple (sine; sinusoid)
- Complex (actually a composite of many overlapping simple waves)
Sinusoid waves

• Simple periodic motion from perfectly oscillating bodies
• Found in nature (e.g., swinging pendulum, sidewinder snake trail, airflow when you whistle)
• Sinusoids sound ‘cold’ (e.g. flute)
Simple waves - key properties

• **Frequency** = cycles per sec (cps) = Hz

• **Amplitude** – measured in decibels (dB), 1/10 of a bel

(Note: dB is on a log scale, increases by powers of 10)
• Frequency = cycles second
  Or “cps” or “Hertz (Hz)”
  Higher frequency = higher pitch

• Intensity
  – Magnitude (amplitude) of vibration
  – Amount of change in pressure

• Frequency and intensity are independent!
Amplitude (intensity)

Peaks (condensations) = increases in pressure

Troughs (rarefactions) = decreases in pressure

More force is applied and this increases pressure at peaks
Frequency - Tones
Quickie Quiz!

Q: What is the frequency of this wave?

HINT: It repeats 2.5x in 10 msec
Answer:

- 250 Hz!

(2.5 cycles in .01 sec = 250 cps)
Periodic waves

A complex wave is periodic, but not sinusoidal.

Panels A and B are periodic waves, though Panel B is complex periodic.
Complex periodic waves

- Results from imperfectly oscillating bodies
- Demonstrate simple harmonic motion
- Examples - a vibrating string, the vocal folds
Complex periodic waves – cont’d

• Consists of a fundamental ($F_0$) and harmonics

• Harmonics (“overtones”) consist of energy at integer multiples of the fundamental ($x2, x3, x4$ etc…)
Harmonic series

- Imagine you pluck a guitar string and could look at it with a really precise strobe light
- Here is what its vibration will look like
From simple waves → complex (showing power spectra)

- Also known as a “line spectrum”
- At bottom is a complex wave with an $F_0$ of 100 Hz
- Note energy at 200 Hz (second harmonic) and 300 Hz (third harmonic)
Fourier’s Theorem

Fourier Synthesis

Complex waveform

Fourier Analysis

Amplitude (V)

Time - msec

100-Hz

200-Hz

300-Hz

SUM
Frequency – Tones/Adding
Frequency - Male Vowels
Frequency - Female Vowels
Periodic and aperiodic waveforms

Spectra of periodic and aperiodic sounds
Transient vs. continuous signal
Continuous spectrum

- Energy is present across a continuum of frequencies (noise)

(not necessarily equal energy at all frequencies!)
Decibels

• Unit used to express intensity or pressure of sound
  1) Ratio of 2 numbers
  2) Exponential or logarithmic scale
  3) Expressed in terms of a specified reference value
Linear vs. Logarithmic Scale

Linear
- Incremental amount between units

<table>
<thead>
<tr>
<th>Linear Scale</th>
<th>1,000</th>
<th>2,000</th>
<th>3,000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1,000</td>
</tr>
</tbody>
</table>

Logarithmic
- Increments NOT EQUAL
- Based on exponents of a given number

<table>
<thead>
<tr>
<th>Log Scale</th>
<th>$10^2$</th>
<th>$10^3$</th>
<th>$10^4$</th>
<th>$10^5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>900</td>
<td>9,000</td>
<td>90,000</td>
<td></td>
</tr>
</tbody>
</table>
Why use log scale for sound?

• Large dynamic range between pressure of smallest audible sound and largest tolerable sound

• Difficult to use with interval scale, so use a ratio scale
Decibel

• Logarithm of the ratio of 2 intensities
• Ratio of power or pressure measured, divided by a reference power or pressure (usually 20 μPA)
• Logarithm labeled a “bel”
• “bel” is still too large, so 1/10<sup>th</sup> of bel is used: decibel
dB – cont’d

- Midrange frequencies can be perceived when they are less intense.
- ..whereas a very low or high frequency sound must have more intensity to be audible.
- Why you need to boost the bass on your stereo system if you turn the volume down low -- otherwise the lows become inaudible.

The ‘speech banana’
More dB facts

- **Threshold of hearing** = “softest sound of a particular frequency a pair of human ears can hear 50% of the time under ideal listening conditions”
- Many prefer **A-weighting** – considers human hearing in the best possible way (dBA)
- Typically tested with a **sound level meter**
- **FACT:** There can be a negative dB value!

Orville Labs, Minneapolis
-9.4 dBA
Physical vs. perceptual

**PHYSICAL**
- Fundamental frequency ($F_0$)
- Amplitude/ Intensity
- Duration

**PERCEPTUAL**
- “Pitch”
- “Loudness”
- “Length”
Phon Scale

• If a given sound is perceived to be as loud as a 60 dB sound at 1000 Hz, then it is said to have a “loudness of 60 phons”

• 60 phons means "as loud as a 60 dB, 1000 Hz tone”

• Instrumentation reference

Psychoacoustic scale for intensity
Bark scale (pitch)

- A psychoacoustical scale for pitch
- Basically log-linear
- Ranges from 1 to 24 and corresponds to the first 24 critical bands of hearing.
- Similar to Mel Scale

(1961)
Phase

- A measure of the position along the sinusoidal vibration
- These two waveforms are slightly out of phase (approx. 90° difference)
- Important for sound localization
Phase examples

Fig. 1. Illustration of the in-phase, anti-phase and 90° bimanual coordination tasks.
Damping - example

(Loss of vibration due to friction)
Review of source characteristics

- **Simple waves** are a good way to learn about basic properties of frequency, amplitude, and phase.
- Examples include whistling; not really found much in speech

- **Complex waves** are found in nature for oscillating bodies that show simple harmonic motion (e.g., the vocal folds)
Source-filter theory

- Lungs give power
- Larynx gives buzzing

- Moving cavities shape sound

- Final speech output

- Vocal tract model demo
- https://www.youtube.com/watch?v=wR41CRbIjV4
- See chapter 6 in book for more info
Let’s look at the filter

• In speech, the filter is the supralaryngeal vocal tract (SLVT)

• The shape of the oral/pharyngeal cavity determines vowel quality (via resonance)

• SLVT shape is chiefly determined by tongue movement, but lips, velum and (indirectly) jaw also play a role
Resonance

• Resonance = reinforcement or shaping of frequencies as a function of the boundary conditions through which sound is passed.

• To get a basic idea of resonance, try producing a vowel with and without a paper towel roll placed over your mouth! The ‘extra tube’ changes the resonance properties.
Resonance – cont’d

• The SLVT can be modeled as a kind of bottle with different shapes… as sound passes through this chamber it achieves different sound qualities

• The resonances of speech that relate to vowel quality are called **formants**. Thus, R1 = F1 (‘first formant). R2 = F2, etc.

• F1 and F2 are critical determinants of vowel quality
Resonators
Resonance - concepts

• Natural frequency (Resonant frequency RF)

• Mechanical (vibrating body, e.g. vocal folds)

• Acoustic (air-filled space)
Resonance

- Child’s swing behaves like a pendulum and can help us understand resonance
- Restorative & displacement forces, inertia & equilibrium
- Two ways to inject energy into the oscillation:
Resonance

• Resonance: large increase in vibration when a force is applied at a natural frequency of the medium

• Set an object into vibration (hit, drop, etc) and the rate at which it vibrates ~ natural frequency (or set of frequencies).
Vibration

Free vibration – no additional force is applied after an object is set into motion.

Forced vibration – when forces “drive” the motion to vibrate at its natural frequency. (e.g., guitar string).
Forced Vibration

The Tacoma-Narrows Bridge twisting in sympathetic resonance with the wind (left) and stretching to the point of breaking (right).

http://www.youtube.com/watch?v=3mclp9QmCGs
Resonant Freq of Glass

- Breaking glass with sinusoid sound
- Freaky deaky guy actually breaks glass with voice (Mythbusters)
- Typical case of faking it
Why is resonance important?

- In vocal tract, allows some frequencies to be magnified over others.
- In the auditory system (ear canal) allows for certain sounds to be amplified as compared to others.
Acoustic Resonators / Bandwidth

Regularly shaped acoustic resonator will be narrowly tuned and lightly damped

Irregularly shaped acoustic resonator will be broadly tuned and heavily damped
Narrow vs. Broad Filter

(a) narrow band filter

(b) broad band filter

(regular resonator)

(irregular resonator)

Source: Clark & Yallop (1995)
Passband Resonator

![Diagram showing passband resonator characteristics](image)

- Center Frequency (Fc)
- Lower Cutoff Frequency (Fl)
- Upper Cutoff Frequency (Fu)
- Pass Band
- 3 dB Down Point

**Parameters:**
- Frequency (F)
- Relative Amplitude (y-axis)

**Graph Range:**
- Frequency: 100 to 1000
- Relative Amplitude: Scale not specified

**Graph Notes:**
- The diagram illustrates the passband characteristics, showing the frequency range and amplitude behavior of a passband resonator.
Input \rightarrow \text{SLVT} \rightarrow \text{speech output}

Application: Sound spectrograms, speech analysis
F1 through F3

• Closed-tube model of the VT, showing first three resonances (formants)
Formants for three GAE vowels
Formants – ‘HAR’

✓ **H**: *Height* relates inversely to F1.
✓ **A**: *Advancement* relates to F2.
✓ **R**: *Rounding* is a function of lip protrusion and lowers all formants - through lengthening of the vocal tract by approximately 2 to 2.5 cm.
Figure 12-8: F2 x F1 plot — American English vowels.

Peterson & Barney, 1952
To be continued…

• We have described **resonance** and **formant frequencies with respect to vowels**
• Also important for **consonants**
• Because consonants require more spatial and temporal precision, these details will be covered later…