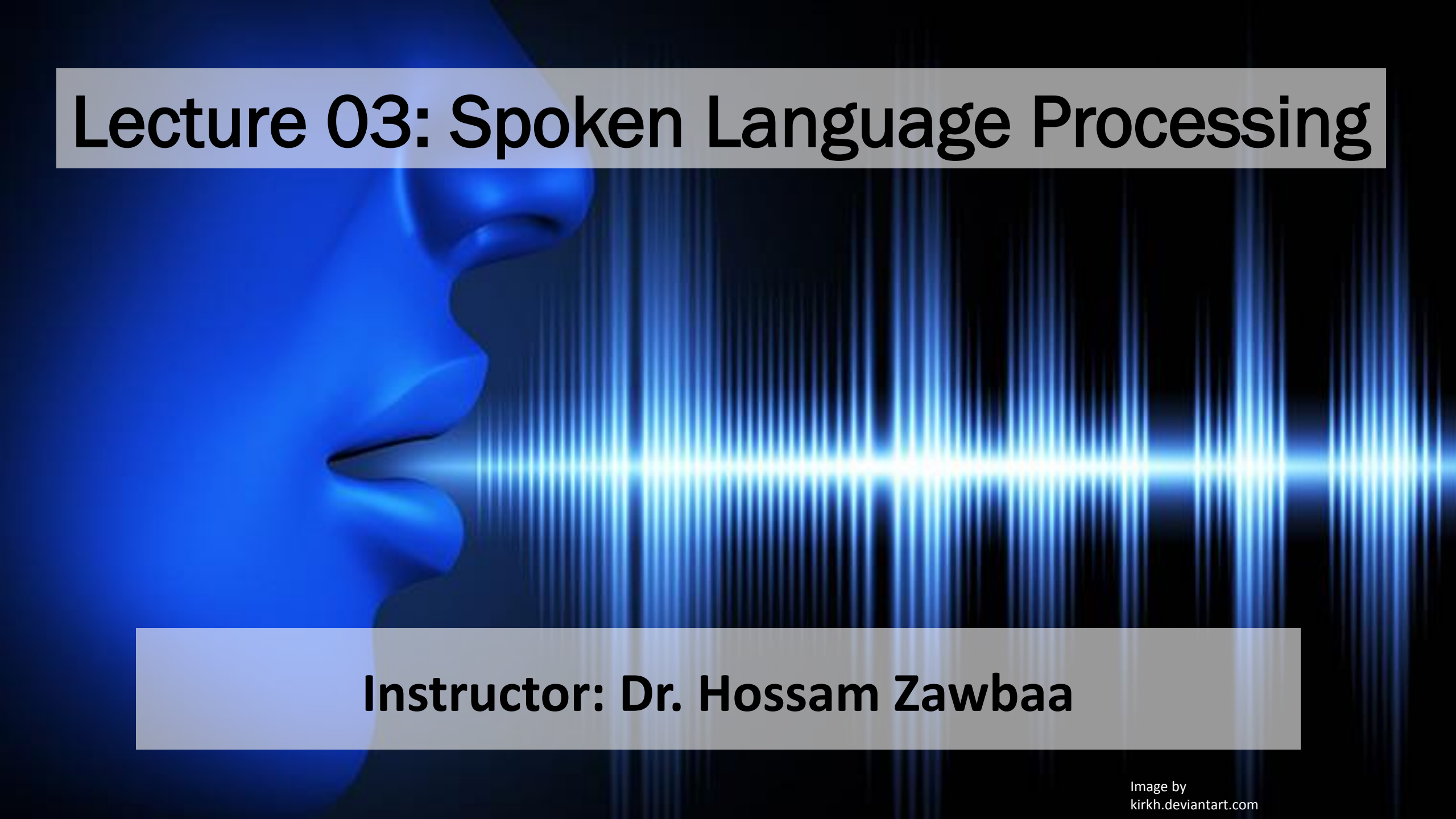
Lecture 03: Spoken Language Processing



Instructor: Dr. Hossam Zawbaa

Image by
kirkh.deviantart.com

Phonetics

- ARPAbet
  - An alphabet for transcribing American English phonetic sounds.
- Articulatory Phonetics
  - How speech sounds are made by articulators (moving organs) in mouth.
- Acoustic Phonetics
  - Acoustic properties of speech sounds.

ARPAbet

- http://www.stanford.edu/class/cs224s/arpabet.html

- The CMU Pronouncing Dictionary

http://www.speech.cs.cmu.edu/cgi-bin/cmudict

- What about other languages?
- International Phonetic Alphabet:

http://en.wikipedia.org/wiki/International_Phonetic_Alphabet

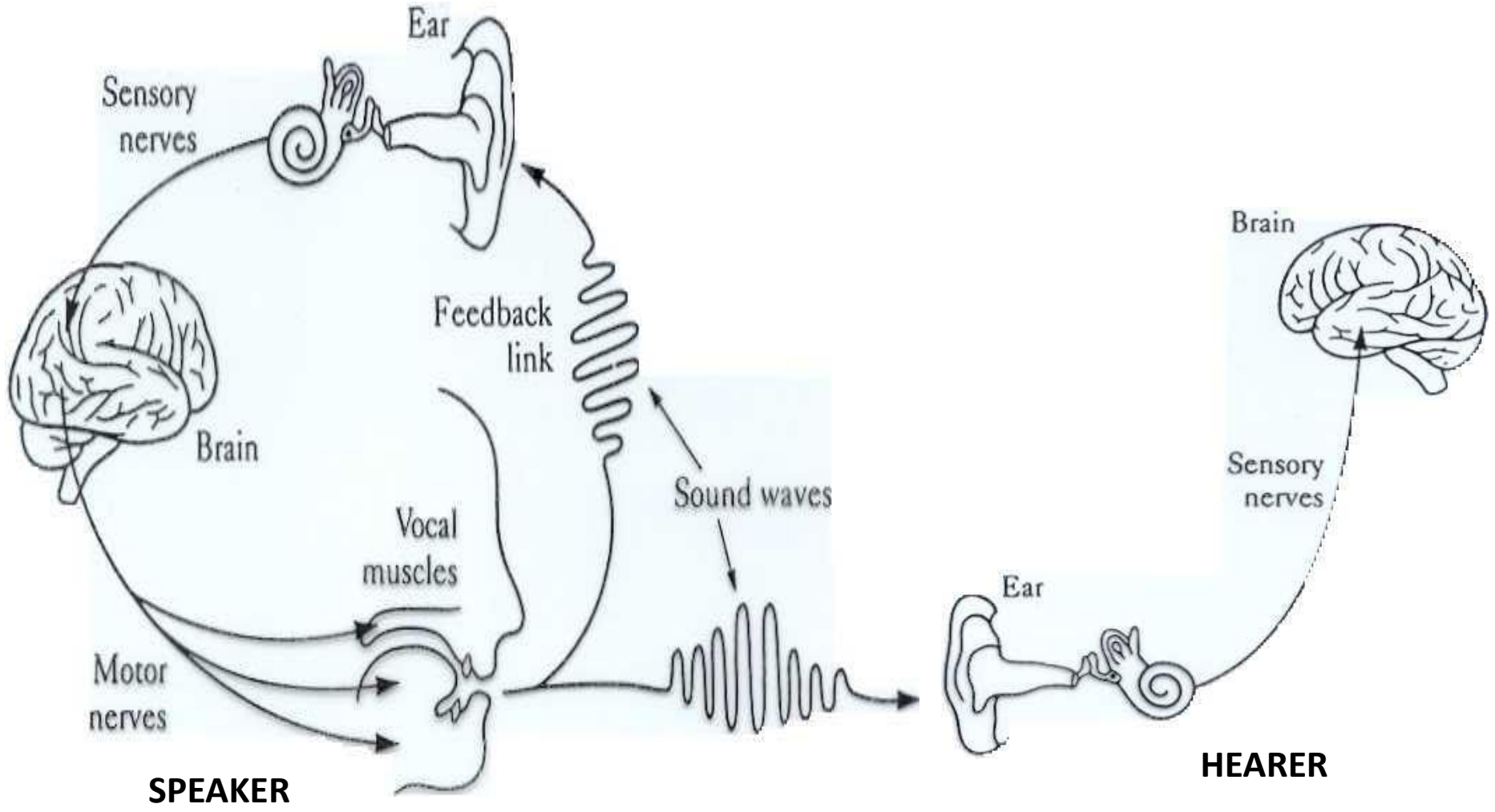
ARPAbet Vowels

	b_d	ARPA		b_d	ARPA
1	bead	iy	9	bode	ow
2	bid	ih	10	booed	uw
3	bayed	ey	11	bud	ah
4	bed	eh	12	bird	er
5	bad	ae	13	bide	ay
6	bod(y)	aa	14	bowed	aw
7	bawd	ao	15	Boyd	oy
8	Budd(hist)	uh			

Note: Many speakers pronounce Buddhist with the vowel uw as in booed, So for them [uh] is instead the vowel in "put" or "book"

https://corpus.linguistics.berkeley.edu/acip/

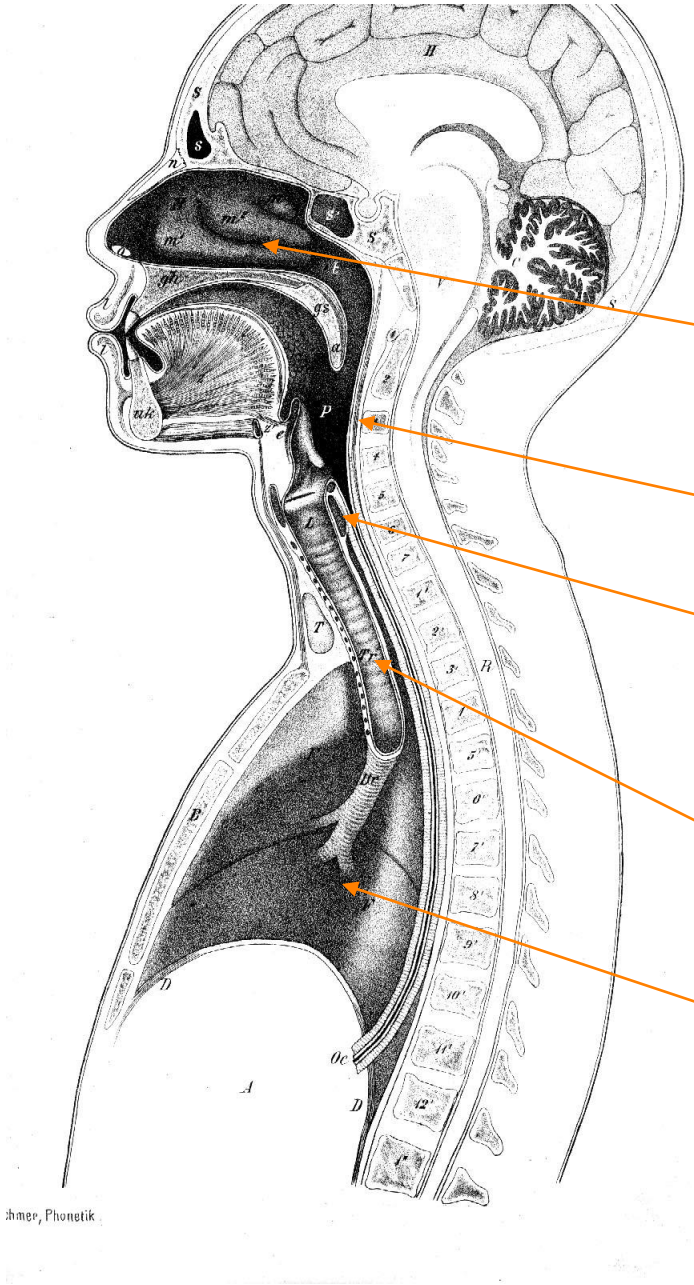
The Speech Chain (Denes and Pinson)



Speech Production Process

- Respiration
  - We (normally) speak while breathing out.
  - Respiration provides airflow.
- Phonation
  - Airstream sets vocal folds in motion.
  - Vibration of vocal folds produces sounds.
  - Sound is then modulated by:
    - Articulation and Resonance
    - Shape of vocal tract, characterized by:
      - Oral tract
        - Teeth, soft palate (velum), hard palate
        - Tongue, lips, uvula
      - Nasal tract

Text adopted from Sharon Rose



Sagittal section of the vocal tract
(Techmer 1880)

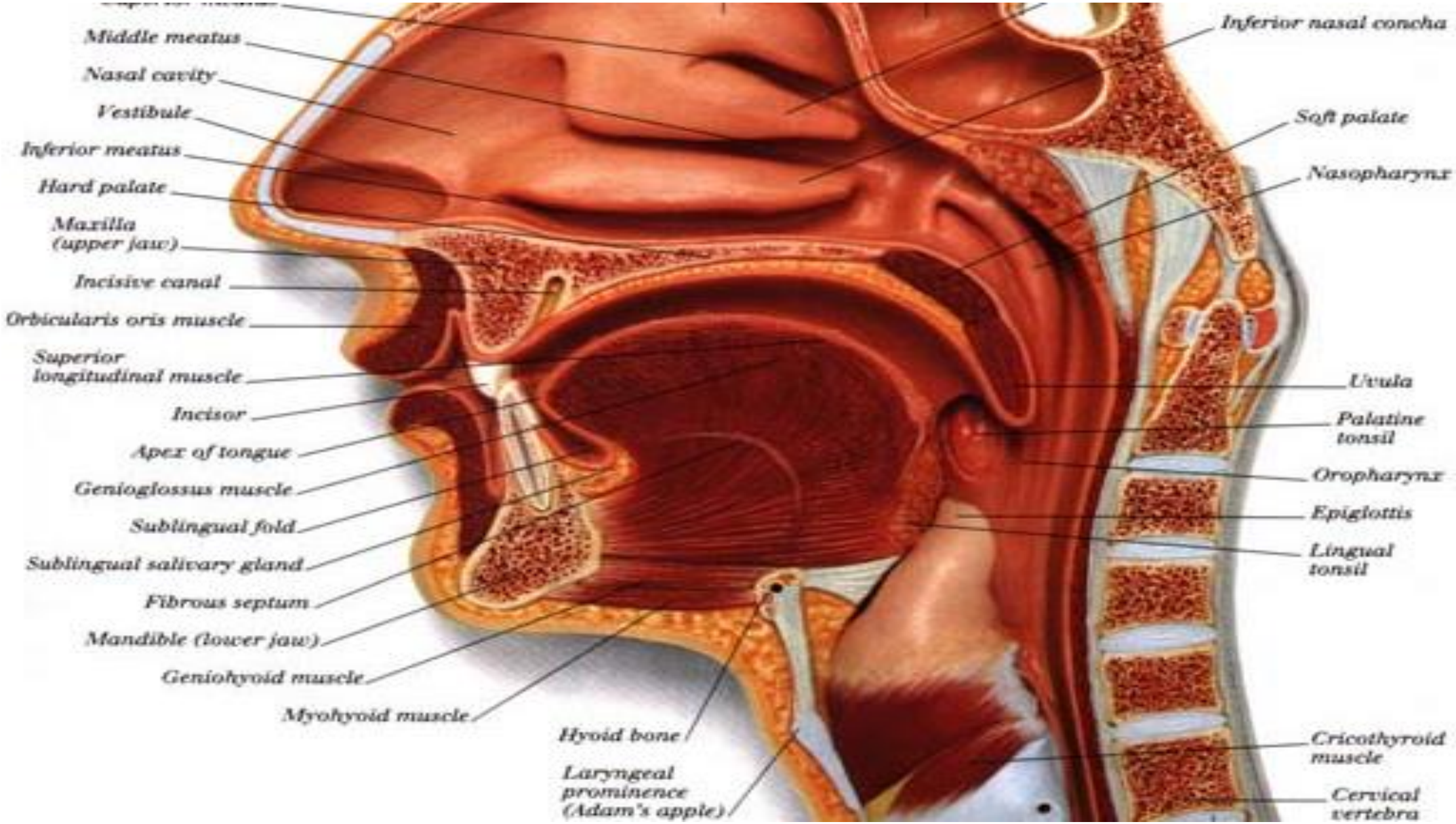
Nasal Cavity

Pharynx

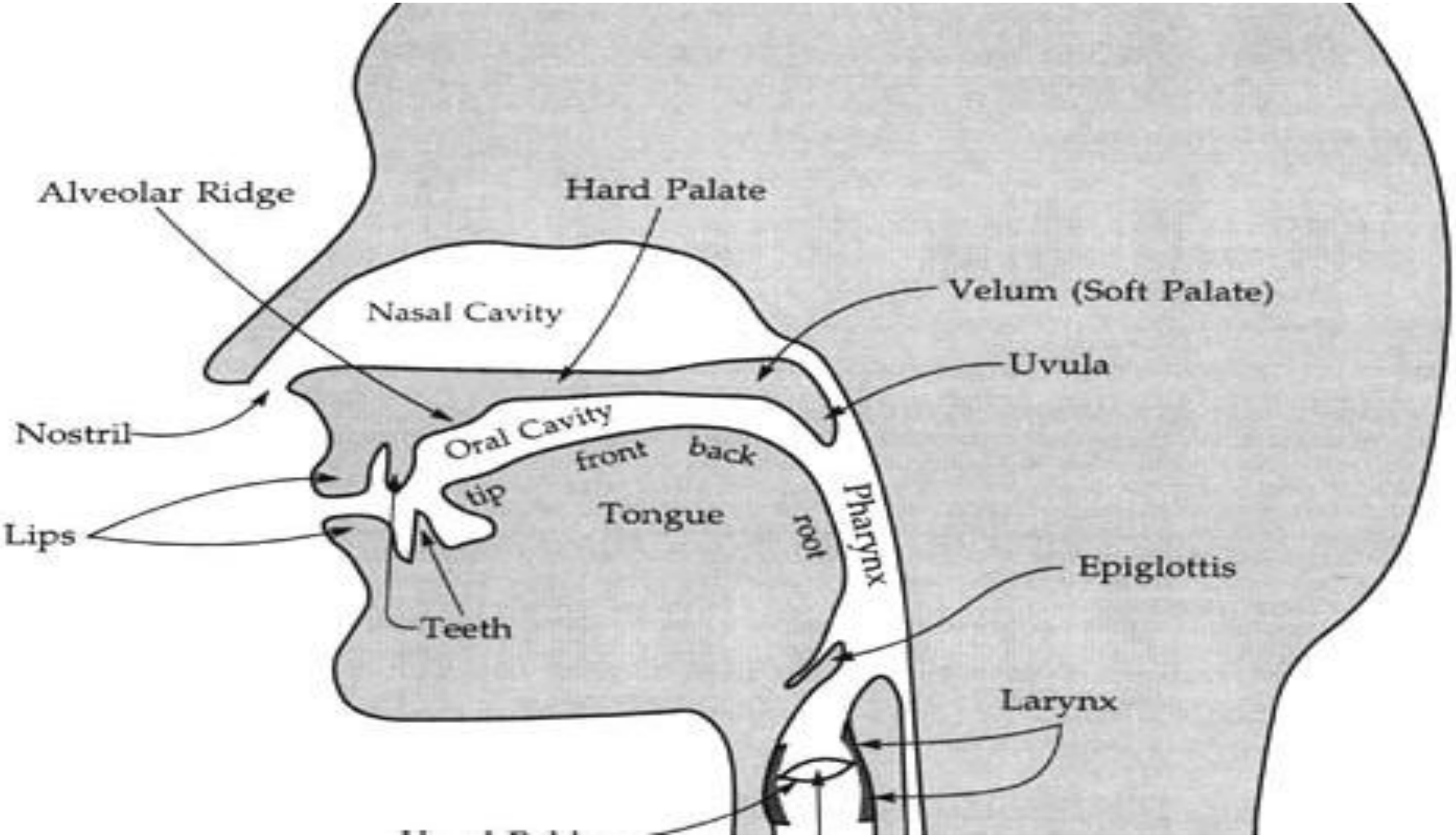
Vocal Folds (within the Larynx)

Trachea

Lungs



From Mark Liberman's website, from Ultimate Visual Dictionary



From Mark Liberman's Web Site, from Language Files (7th ed)

Larynx and Vocal Folds

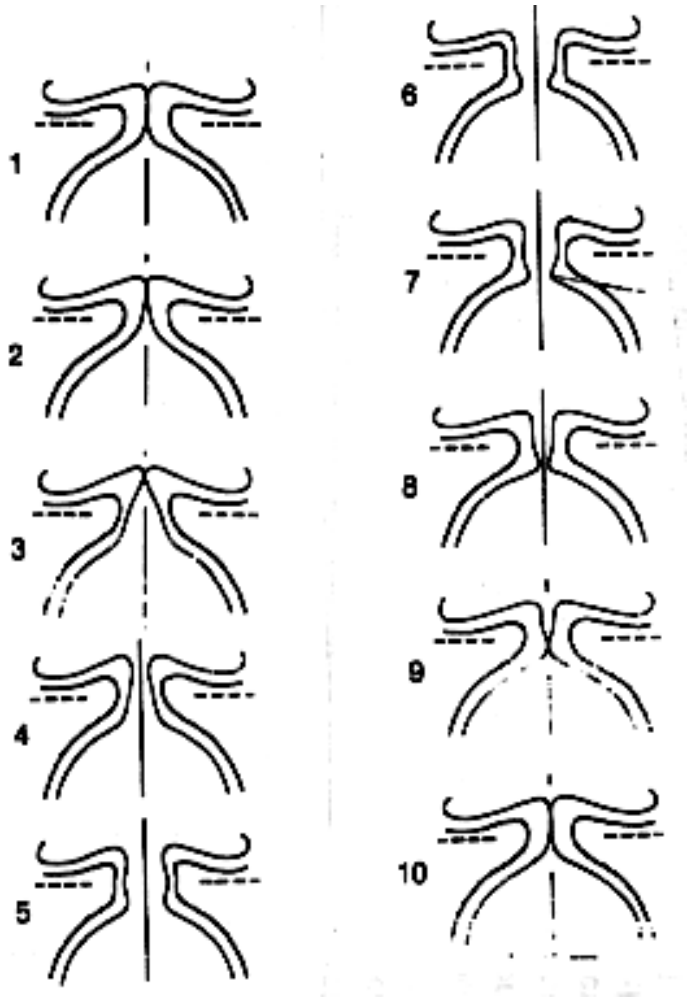
- The Larynx (voice box)
  - A structure made of cartilage and muscle
  - Located above the trachea (windpipe) and below the pharynx (throat)
  - Contains the vocal folds
  - (adjective for larynx: laryngeal)
- Vocal Folds (older term: vocal cords)
  - Two bands of muscle and tissue in the larynx
  - Can be set in motion to produce sound (voicing)

Text from slides by Sharon Rose UCSD LING 111 handout

Vertical slice through larynx, as seen from back

Figure thnx to John Coleman!!

Voicing:



- Air comes up from lungs
- Forces its way through vocal cords, pushing open (2,3,4)
- This causes air pressure in glottis to fall, since:
  - when gas runs through constricted passage, its velocity increases (**Venturi tube effect**)
  - this increase in velocity results in a drop in pressure (**Bernoulli principle**)
- Because of drop in pressure, vocal cords snap together again (6-10)
- Single cycle: ~1/100 of a second.

Figure & text from John Coleman's web site

Voiceless

- When vocal cords are open, air passes through unobstructed
- Voiceless sounds: p/t/k/s/f/sh/th/ch
- If the air moves very quickly, the turbulence causes a different kind of phonation: whisper

Consonants and Vowels

- Consonants: phonetically, sounds with audible noise produced by a constriction
- Vowels: phonetically, sounds with no audible noise produced by a constriction

- (it's more complicated than this, since we have to consider syllabic function, but this will do for now)

Text adapted from John Coleman

Articulation

- In **phonetics** and phonology, **articulation** is the movement of the tongue, lips, jaw, and other **speech** organs (the **articulators**) in ways that make **speech** sounds.
- Sound is **produced** simply by expelling air from the lungs.
- Consonants are classified according to the location where the airflow is most constricted.
- Three major kinds of place articulation:
  - **Labial** (with lips)
  - **Coronal** (using tip or blade of tongue)
  - **Dorsal** (using back of tongue)

Labial place

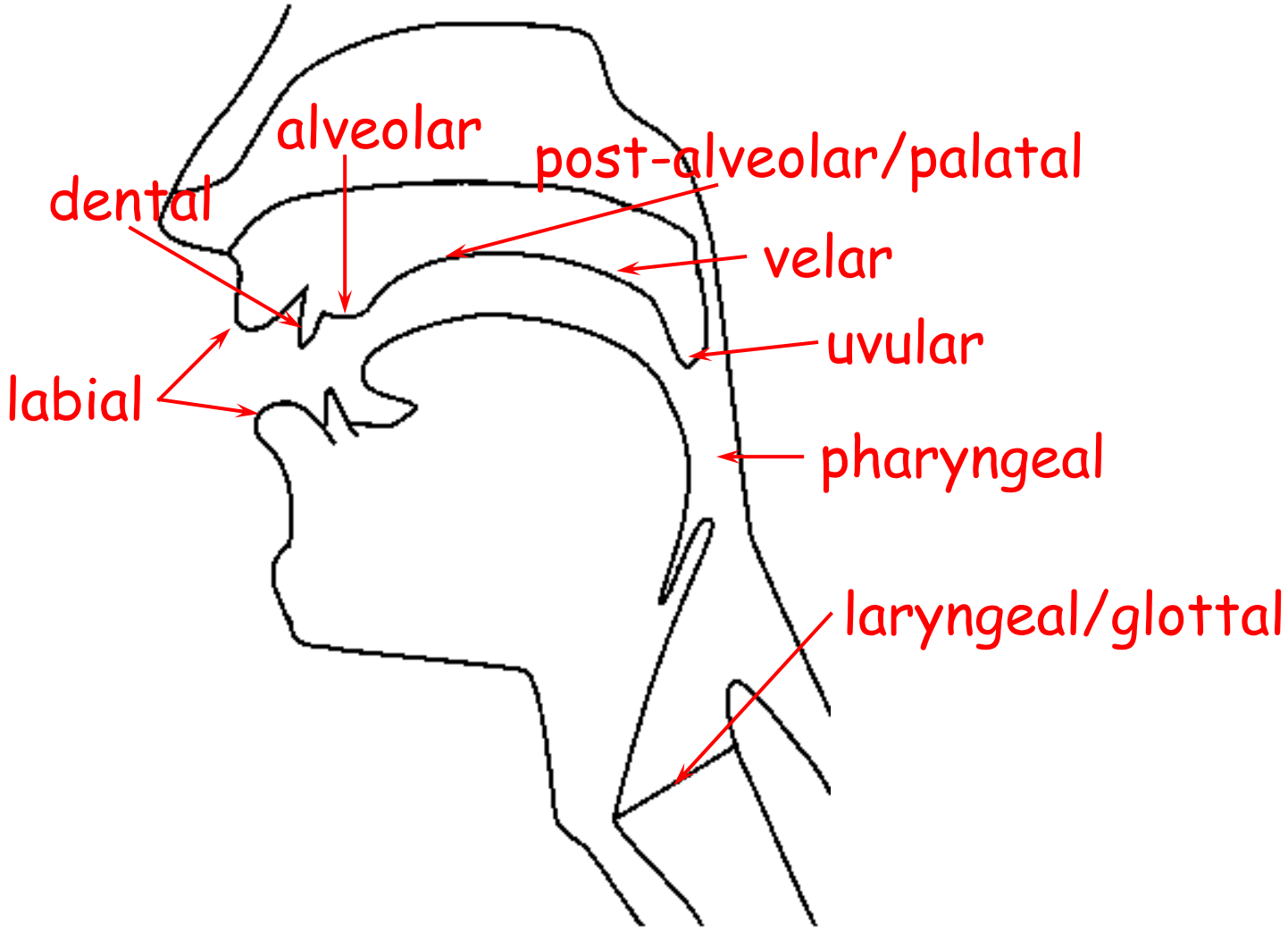
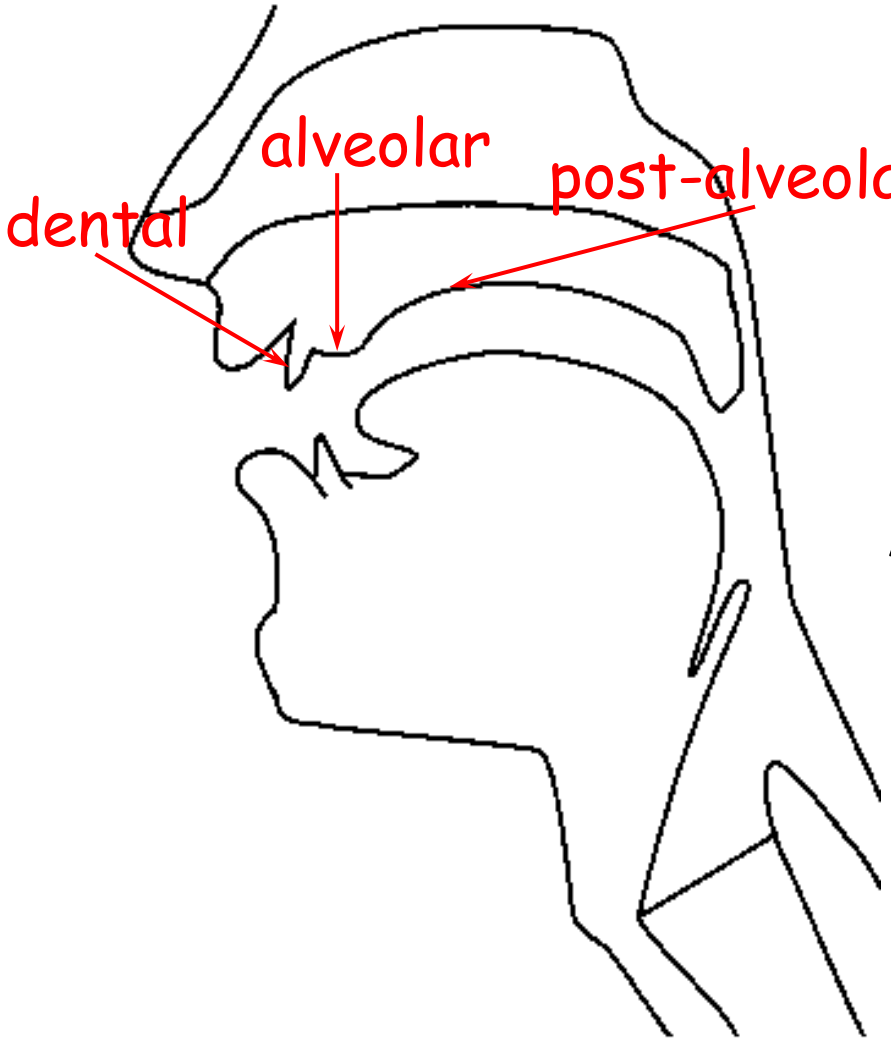


Figure thanks to Jennifer Venditti

Coronal place



Dental:
th/dh
Alveolar:
t/d/s/z/l
Post:
sh/zh/y

Figure thanks to Jennifer Venditti

Dorsal Place

Velar:

k/g/ng

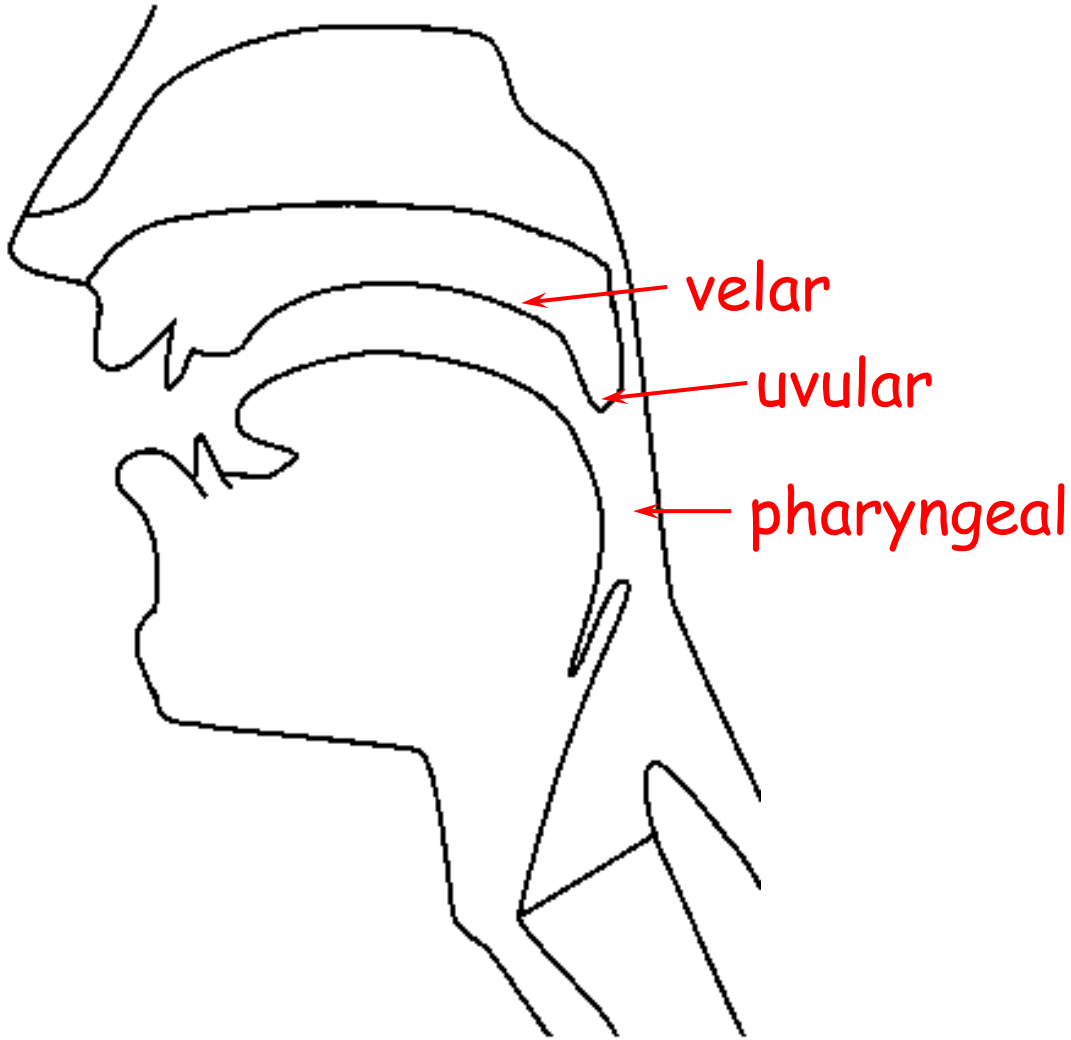
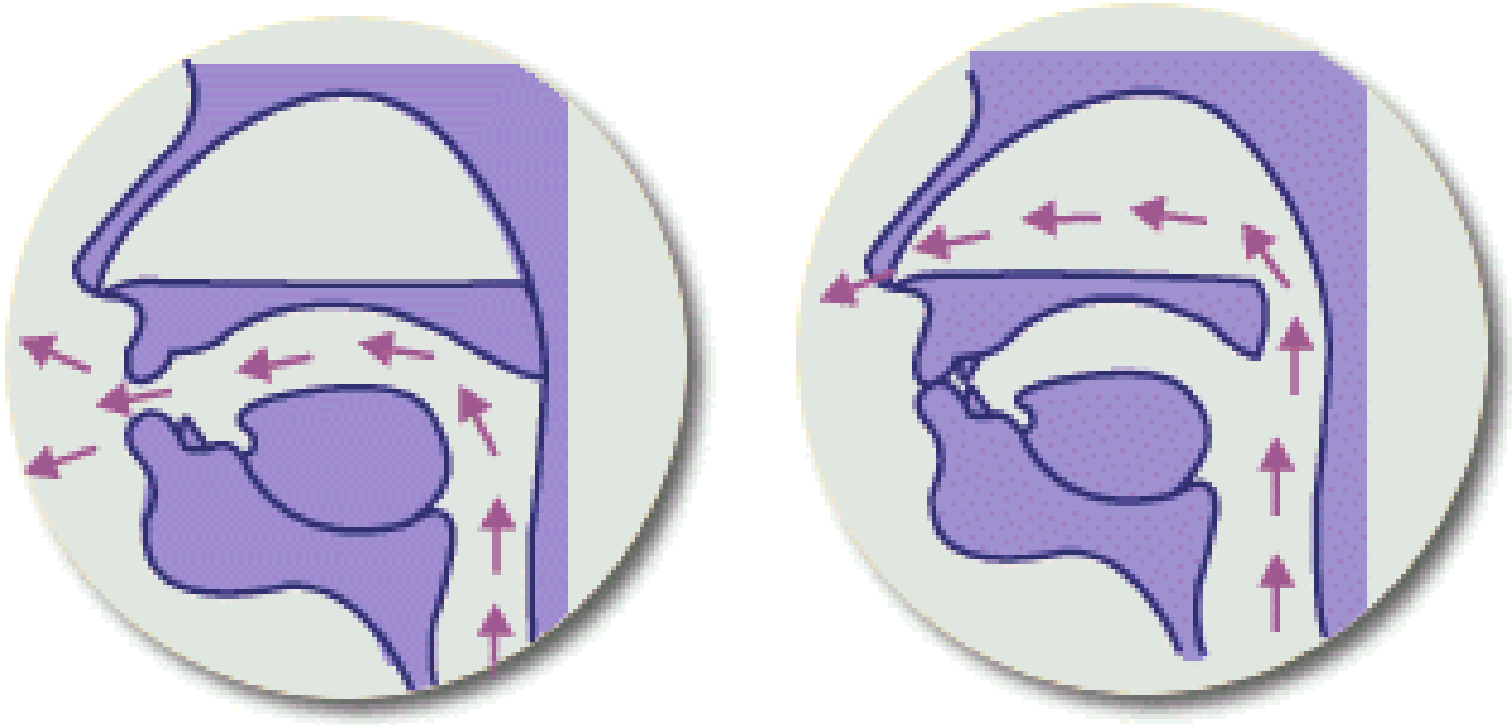


Figure thanks to Jennifer Venditti

Manner of Articulation

- Stop: complete closure of articulators, so no air escapes through mouth
- Oral stop: palate is raised, no air escapes through nose. Air pressure builds up behind closure, explodes when released
  - p, t, k, b, d, g
- Nasal stop: oral closure, but palate is lowered, air escapes through nose.
  - m, n, ng

Oral vs. Nasal Sounds



Thanks to Jong-bok Kim for this figure!

Vowels

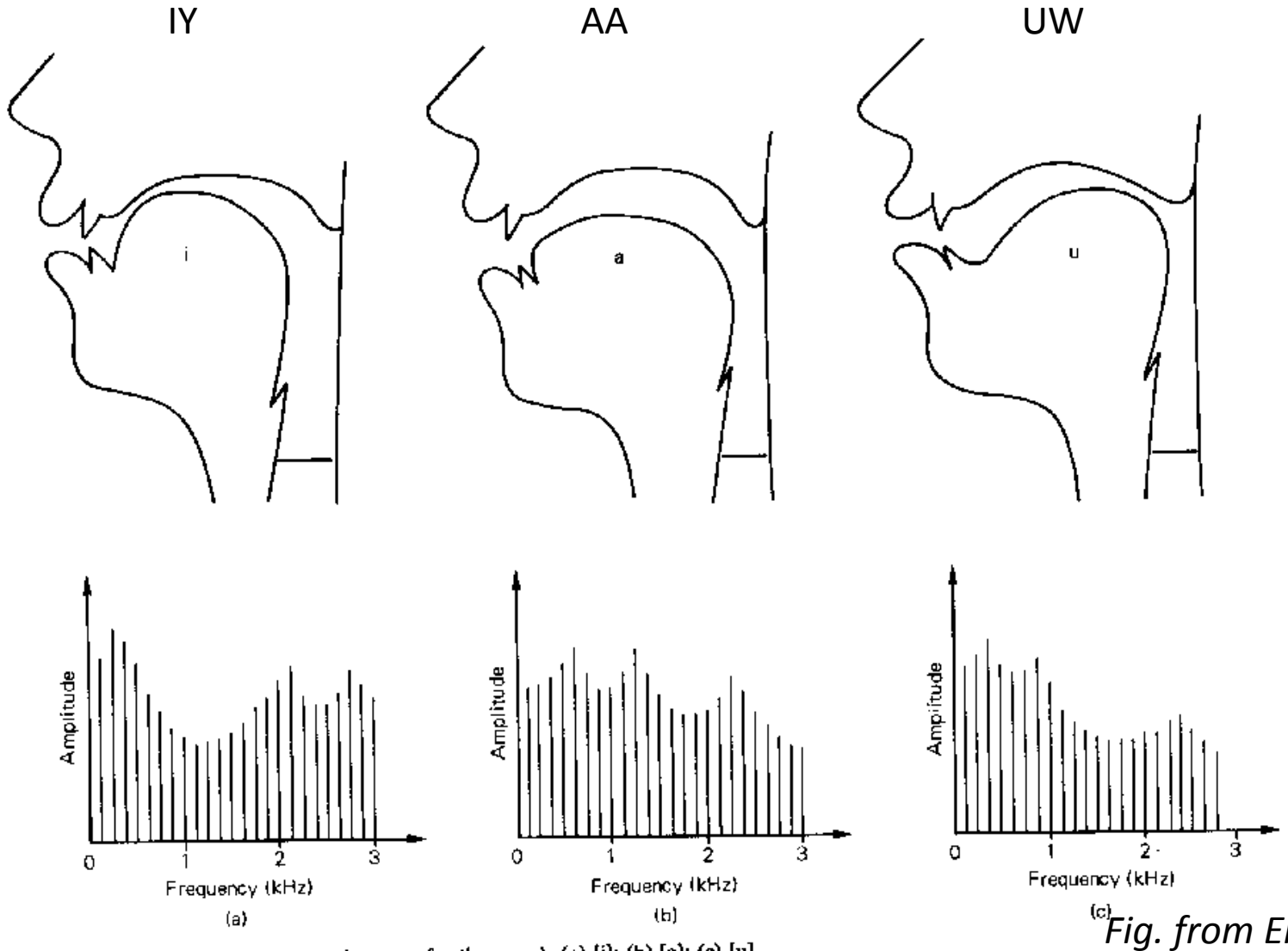
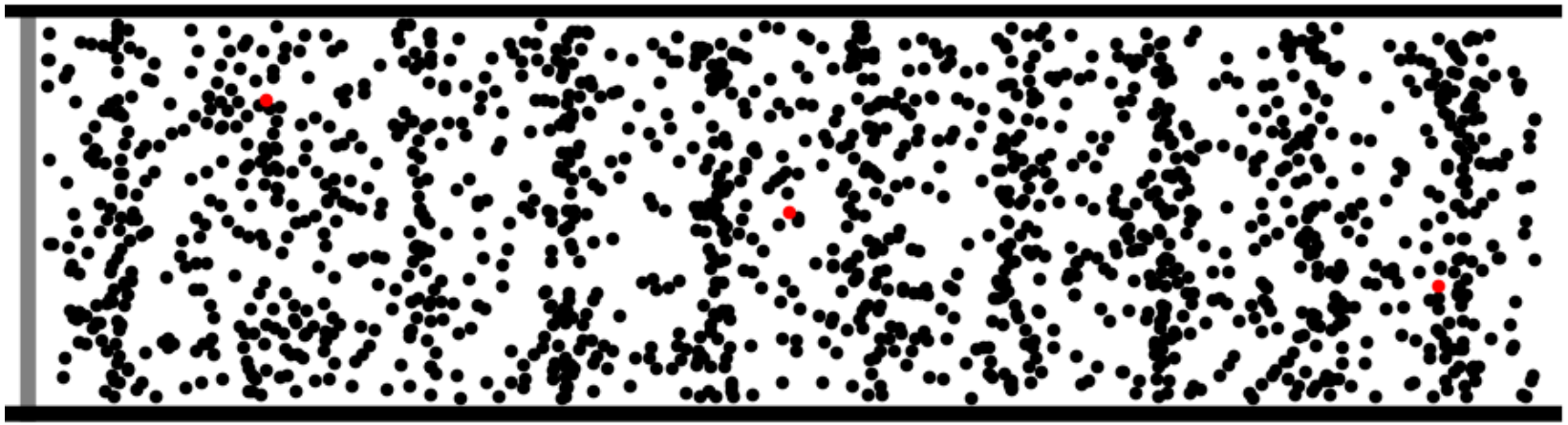


Fig. from Eric Keller

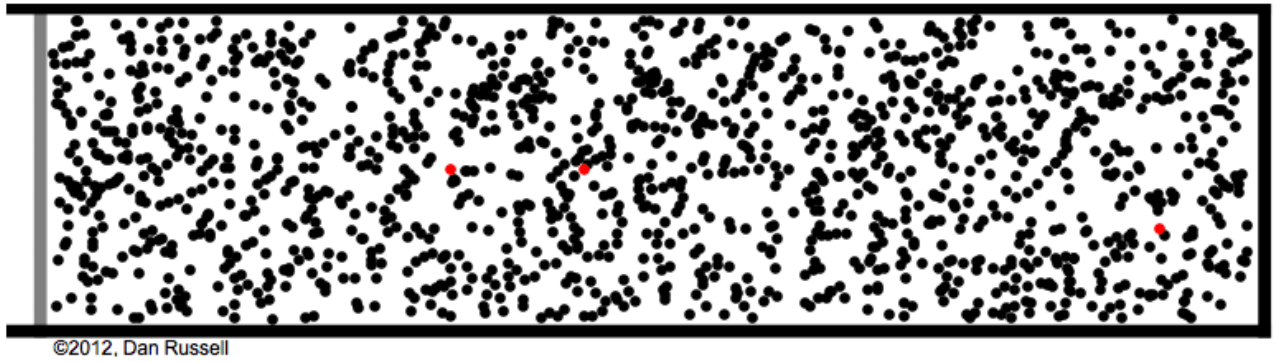
Where to go for more info

- Ladefoged, Peter. 1993. A Course in Phonetics
- Mark Liberman's site
  - http://www.ling.upenn.edu/courses/Spring_2001/ling001/phonetics.html
- John Coleman's site
  - http://www.phon.ox.ac.uk/%7Ejcoleman/mst_mphil_phonetics_course_index.html
- Jennifer Smith's resource page
  - http://www.unc.edu/~jlsmith/pht-url.html

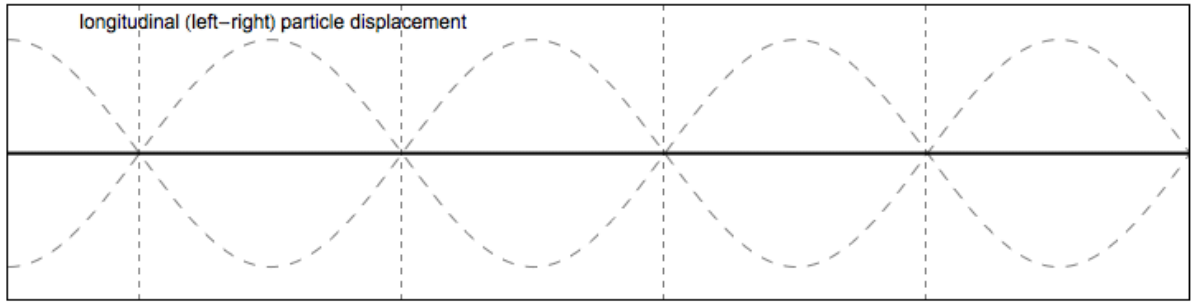
Sound waves are longitudinal waves



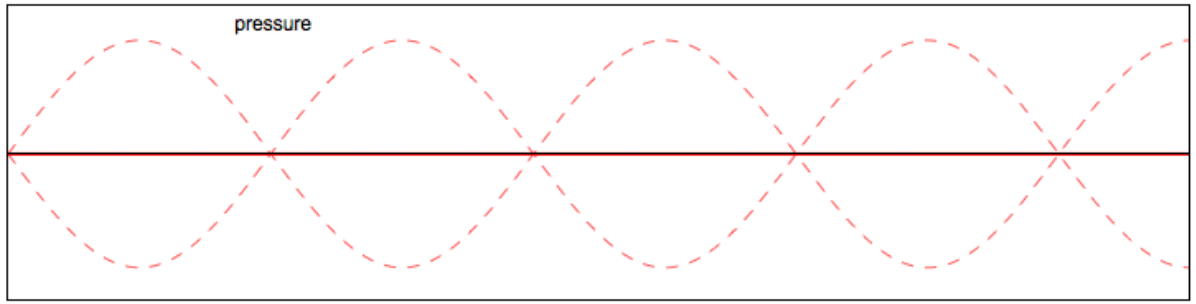
Dan Rusell Figure



particle displacement



pressure



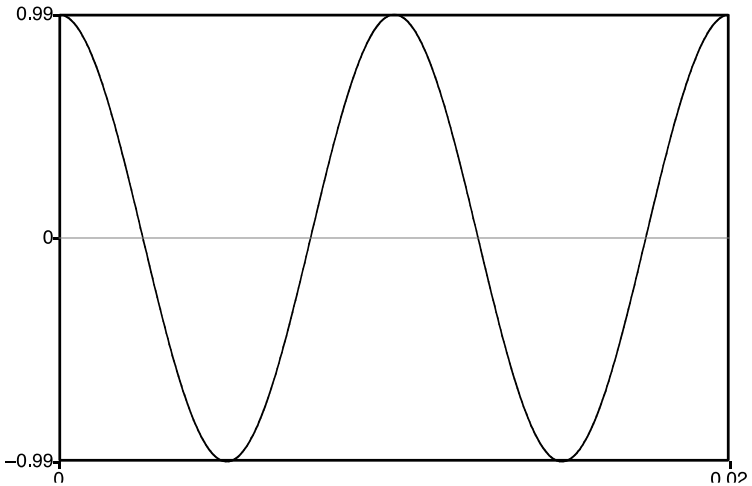
Dan Rusell Figure

Remember High School Physics
Simple Period Waves (sine waves)

- Characterized by:
  - period: T
  - amplitude A
  - phase ϕ
- Fundamental frequency

in cycles per second, or Hz

  - $F_0=1/T$

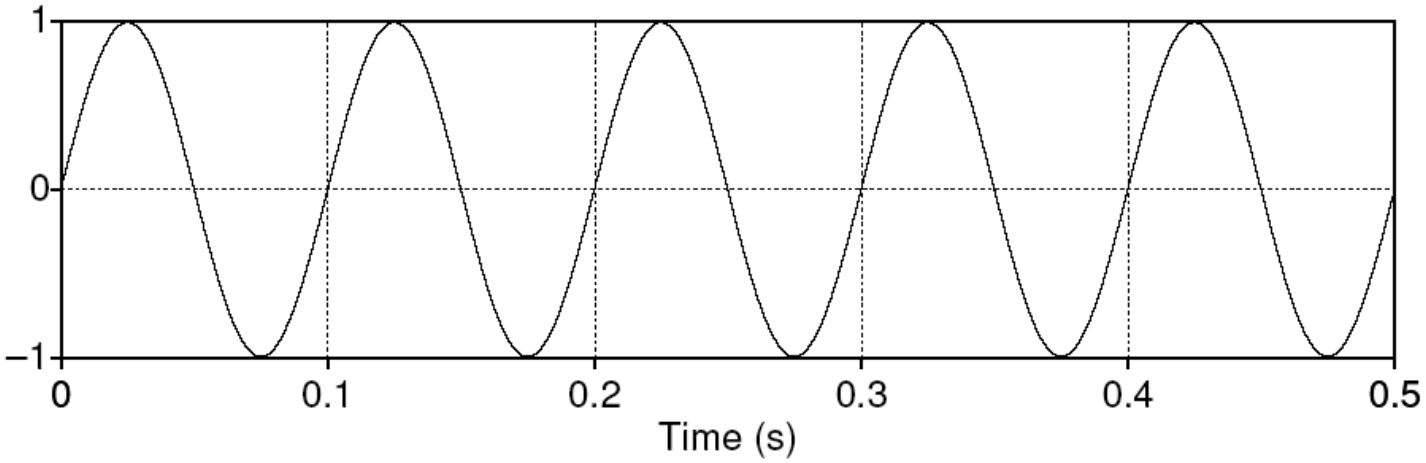


1 cycle

To listen to sine waves:
http://www.szynalski.com/tone-generator/

Simple periodic waves

- Computing the frequency of a wave:

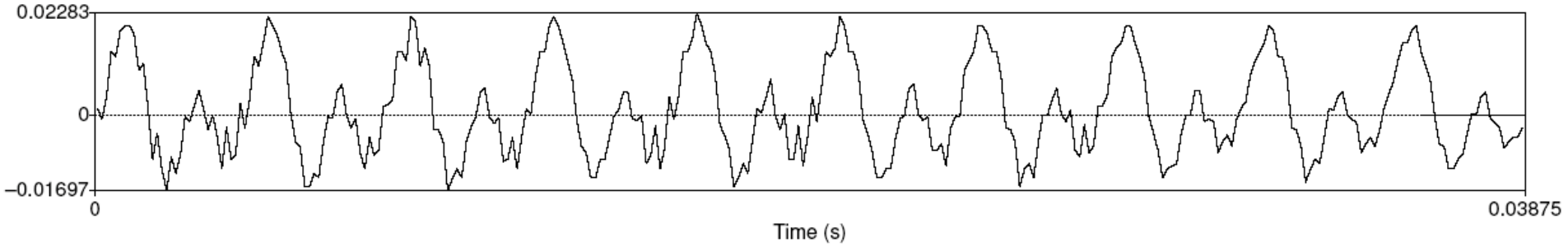


The frequency of a wave:

5 cycles in 0.5 seconds = 10 cycles/second = 10 Hz

Amplitude: 1

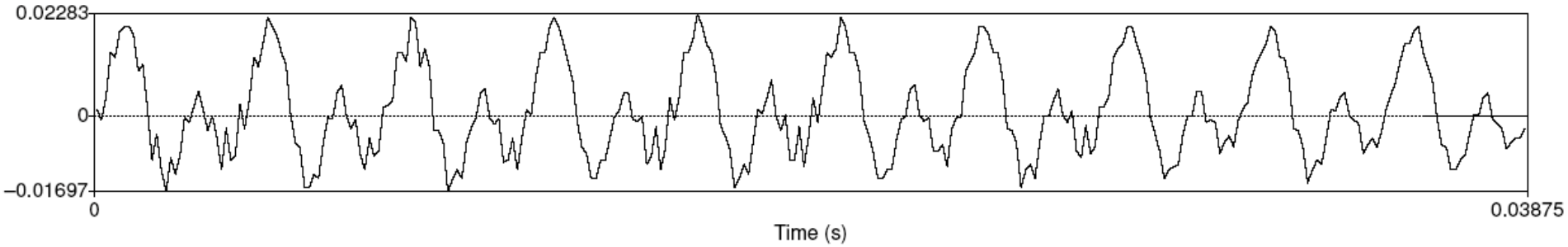
Speech sound waves



- X axis: time.
- Y axis:
  - Amplitude = air pressure at that time
    - +: compression
    - 0: normal air pressure,
    - -: rarefaction

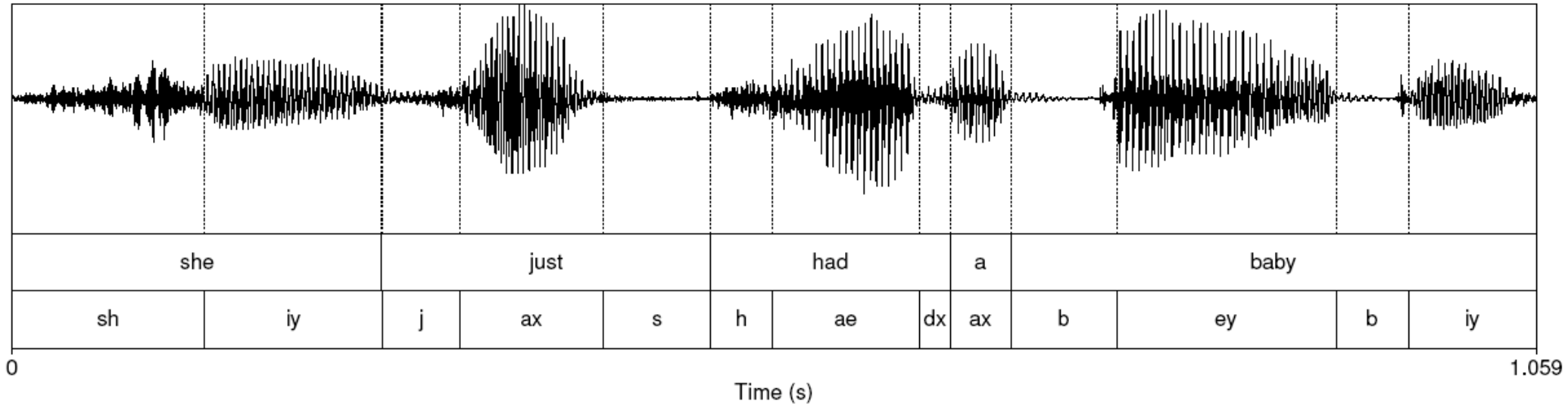
Back to waves: Fundamental frequency

- Waveform of the vowel [**iy**]



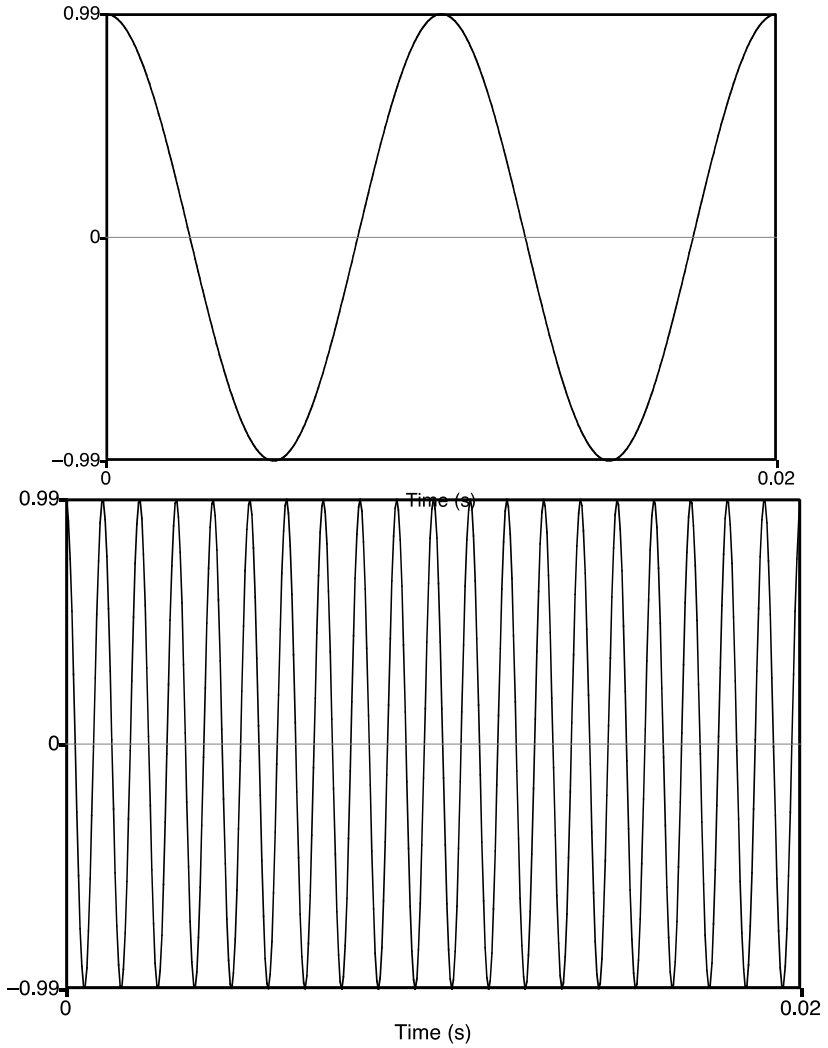
- Frequency: 10 repetitions / 0.03875 seconds = 258 Hz
- **This is speed that vocal folds move, hence voicing**
- **Each peak corresponds to an opening of the vocal folds**
- The **low frequency** of the complex wave is called the **fundamental frequency** of the wave or **F0**

She just had a baby



- Note that vowels all have regular amplitude peaks
- Stop consonant
  - Closure followed by release
  - Notice the silence followed by slight bursts of emphasis: very clear for [b] of “baby”
- Fricative: noisy. [sh] of “she” at beginning

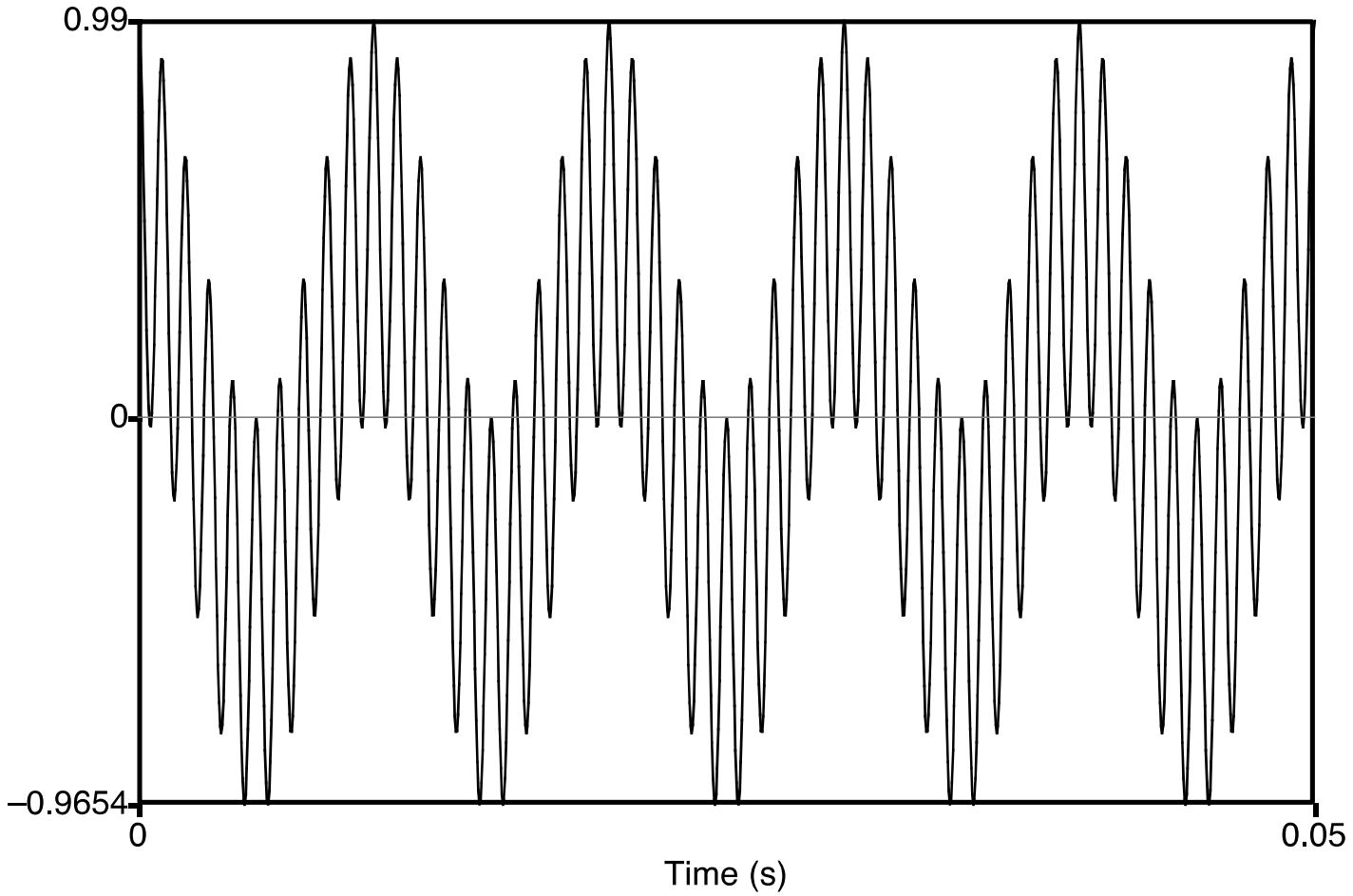
Back to freshman physics:
Waves have different frequencies



100 Hz

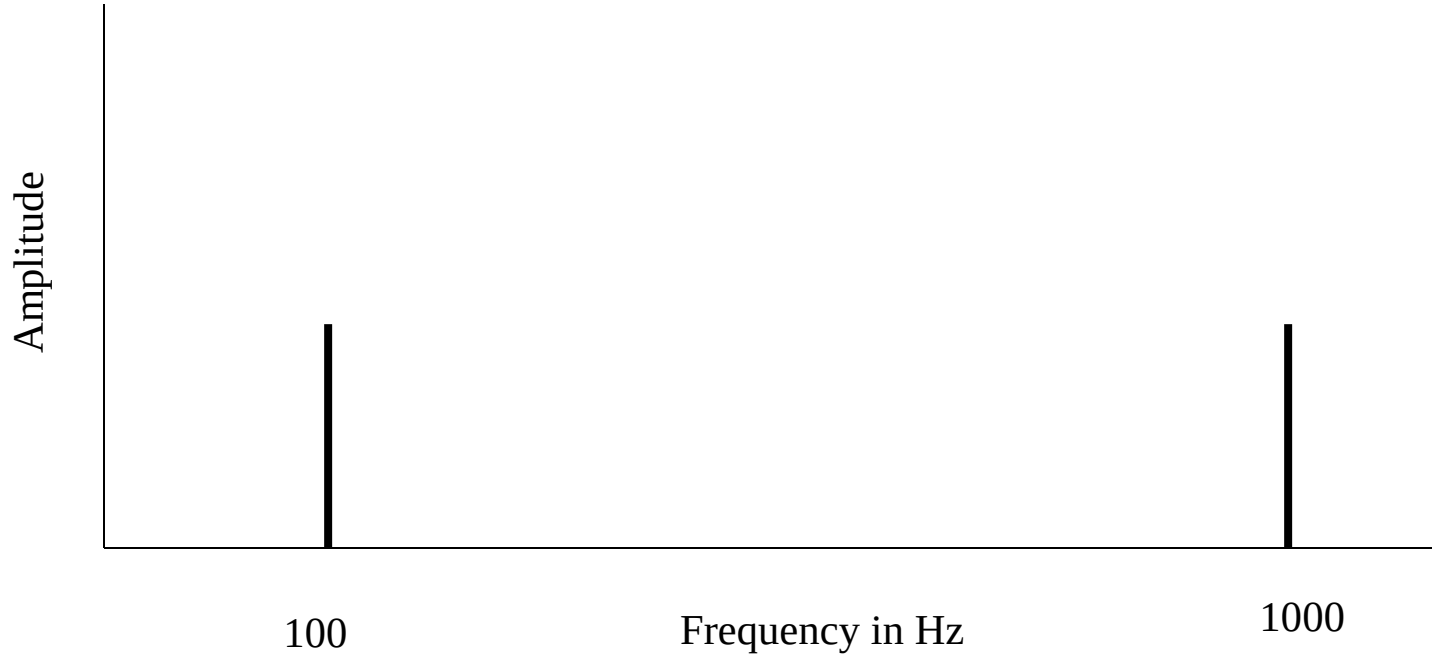
1000 Hz

Complex waves: Adding a 100 Hz and 1000 Hz wave together



Spectrum

- Fourier analysis: any wave can be represented as the (infinite) sum of sine waves of different frequencies (amplitude, phase)



Frequency components (100 and 1000 Hz) on x-axis