- **The source-filter model of vowel acoustics**
- **Modeling schwa**

Background reading:

- *AAP Ch 2, sec 2.1 and 2.4 (review)*

0. Today's objectives

After today's class, you should be able to:

- **Explain**, and **calculate**, how the **frequencies** and the (relative) **amplitudes** of the components are determined for a uniform-vocal-tract [ə] in the **source-filter model of vowel acoustics**
  - The contribution of the glottal source
  - The contribution of the vocal-tract filter
- Recognize and correctly use terms and abbreviations for frequencies in different contexts

1. Review: Tube resonances

- Last week in class:

  How do we get from tube length to resonance frequencies?

- Lab #3:

  How do we get from resonance frequencies to tube length?

1. Review: Tube resonances

- Last week in class:

  How do we get from tube length to resonance frequencies? $L \rightarrow \lambda_n \rightarrow f_n$, for the nth resonance

- Lab #3:

  How do we get from resonance frequencies to tube length? $f_n \rightarrow \lambda_n \rightarrow L$, for the nth resonance

2. The glottal-source wave

- Review: What is the **glottal-source wave**?
  - Also called the **voicing wave**(form) in *AAP* Ch 2
  - → The sound wave produced by ______

2. The glottal-source wave

- Review: What is the **glottal-source wave**?
  - Also called the **voicing wave**(form) in *AAP* Ch 2
  - → The sound wave produced by the vibration of the vocal folds

- To actually hear this sound wave, you would have to put a microphone right above the glottis
  - The sound waves of any speech we normally hear are **further modified** by passing through the vocal tract
    - That is the **filter** part of the source/filter model!

2. The glottal-source wave

- The **shape** of this wave is determined by the way the vocal folds open quickly and close gradually

  → More about this later in the semester

- Computer-synthesized glottal-source wave: [lsrc.wav] (U Delaware Speech Research Tutorials)

- Look at the **waveform** in Praat
  - Periodic? Simple or complex? What is f_0?

2. The glottal-source wave

- Computer-synthesized glottal-source wave:
  [lsrc.wav] (U Delaware Speech Research Tutorials)

- Look at the **waveform** in Praat
  - Periodic? Simple or complex? What is f_0?

- Look at the **spectrum** (window length = 0.5 sec)
  - What is the lowest-frequency component?
  - How do the components' **frequencies** relate?
  - How do the components' **amplitudes** relate?
  - What is f_0? (should get the same answer as above!)

2. The glottal-source wave

- Characteristics of the glottal-source wave
  - **Amplitudes** of the components gradually **decrease** (assuming modal or ‘ordinary’ voicing)
  - **Frequencies** of the components are **whole-number multiples** of the lowest component

- The component frequencies thus have the *same relationship* as in a node/node system
  - Careful here: The vocal folds do not actually vibrate like a string of fixed length (they open and close to emit puffs of air)

2. The glottal-source wave

- Can the **frequencies** of the **components** of the glottal-source wave be manipulated by the speaker?

2. The glottal-source wave

- Can the **frequencies** of the **components** of the glottal-source wave be manipulated by the speaker?

  Yes!

  - Each human's voice can produce a range of f_0 values (or else singing would be impossible)
  - Manipulating f_0 means manipulating the component frequencies
    - f_0 is the GCD of the component frequencies
    - So in the glottal-source wave, f_0 is the same as the first component frequency

2. The glottal-source wave

- If someone produces a vowel with $\boldsymbol{f_0}$ **= 250 Hz**, what are the frequencies of the **first five components** of the **glottal-source wave**?

2. The glottal-source wave

- If someone produces a vowel with $\boldsymbol{f_0}$ **= 250 Hz**, what are the frequencies of the **first five components** of the **glottal-source wave**?
  - f_1 (H1) **=** $\boldsymbol{f_0}$ **= 250 Hz** | *don't forget the first one!*
  - f_2 (H2) = 2* f_1 = 500 Hz | *whole-number multiples*
  - f_3 (H3) = 3* f_1 = 750 Hz
  - f_4 (H4) = 4* f_1 = 1000 Hz
  - f_5 (H5) = 5* f_1 = 1250 Hz

- The components of the glottal-source wave are often called **harmonics**, abbreviated **H**(n)

3. Overview of the source-filter model

- How do you get a trumpet to play a note?

3. Overview of the source-filter model

- To play a trumpet, you “buzz” your lips
  - This creates a **sound source** at one end of the trumpet (disturbance in the medium!)

- But what we *hear* is a trumpet note, not a buzz: The trumpet acts as a **filter**
  - It **changes the shape** (→ the **sound quality**) of the **complex wave** produced by the sound source
  - It does this by **suppressing** (reducing the amplitude of) some of the source wave’s **components** and **amplifying** (increasing the amplitude of) others

3. Overview of the source-filter model

- In much the same way, when we produce a **voiced speech sound** such as a **vowel**…
  - The vocal folds **buzz** (glottal-source wave)
  - The vocal-tract tube **changes the shape** of the complex wave by **suppressing** some source **components** and **amplifying** others

- The **source-filter model** of the vocal tract
  - Sound **source:** the **vibrating vocal folds**
  - **Filter:** the vocal tract as a **tube** (or a series of tubes)

4. The vocal-tract filter

- Much like a trumpet, the vocal tract **filters** the sound energy by **suppressing** some **components** of the glottal-source wave and **amplifying** others
  - The components **amplified** by the vocal tract are those close to its **resonance frequencies**
  - The **resonance frequencies** of the vocal tract depend on its **length** and **shape**
  - The vocal tract forms different **shapes** for different **vowels** (and consonants—more on this later)
- Therefore, each different **vowel** (such as [i] vs. [a]) corresponds to a different vocal-tract **filter**

4. The vocal-tract filter

- So if each different **vowel** corresponds to a different vocal-tract **filter** …

  This means that each **vowel** corresponds to a different set of **amplified frequencies** (vocal-tract **resonances**)

  - **These are known as the vowel's formants**

4. The vocal-tract filter

- Try it out: Synthesized vowels [lii.wav] | [laa.wav]
  (U Delaware Speech Research Tutorials)

- Download and open these sound files in Praat
  - What **vowels** are they? (What do you hear?)
  - Compare them to each other and to the glottal-source wave we looked at before [lsrc.wav]:
    - How do their $\boldsymbol{f_0}$**s** compare?
    - How do their waveform **shapes** compare?
    - How do their **spectra** compare?

4. The vocal-tract filter

- Each vowel (such as [i] vs. [a]) corresponds to a different set of **resonance frequencies** (**formants**)
- The **shape of a complex wave** is determined by the frequencies **and amplitudes** of its components
  → Two different vowel categories have **differently shaped** sound waves *even when they are sung on the same pitch* (=have the same glottal-source wave)
- Waves of different shapes correspond to sounds with different **sound qualities (timbres)** — this is how different vowels differ perceptually

5. Deriving the vocal-tract filter for [ə]

- To analyze the acoustics of vowels, we need to understand **why** they have the resonance frequencies (formants) that they have

- The first step: What are the resonance frequencies of the **neutral vocal-tract vowel, [ə]**?
  - The vocal tract is approximately a **uniform tube** that is **closed** at one end (the glottis) and **open** at the other end (the lips)
  - The neutral vocal-tract vowel is typically represented as [ə] (schwa), but note that it doesn't sound exactly the same as English [ə] as in *sof<u>a</u>*

5. Deriving the vocal-tract filter for [ə]

- The first step: What are the resonance frequencies of the **neutral vocal-tract vowel, [ə]**?
  - The vocal tract is approximately a **uniform tube** that is **closed** at one end (the glottis) and **open** at the other end (the lips)
  - What do we need to know in order to find the resonance frequencies of this vowel?

5. Deriving the vocal-tract filter for [ə]

- The first step: What are the resonance frequencies of the **neutral vocal-tract vowel, [ə]**?
  - The vocal tract is approximately a **uniform tube** that is **closed** at one end (the glottis) and **open** at the other end (the lips)
  - What do we need to know in order to find the resonance frequencies of this vowel?
    - The length of the tube (vocal tract)
    - The boundary conditions of the tube
    - The speed of sound in air (use 350 m/s)

5. Deriving the vocal-tract filter for [ə]

- We know how to calculate the resonance frequencies for a tube of this type: it is a **node/antinode** or **quarter-wavelength** system
  - What is the distance between each pair of resonance frequencies in a system of this type?

5. Deriving the vocal-tract filter for [ə]

- We know how to calculate the resonance frequencies for a tube of this type: it is a **node/antinode** or **quarter-wavelength** system
  - What is the distance between each pair of resonance frequencies in a system of this type?
    - Each resonance is an odd-numbered multiple of the first ($1*f_1$, $3*f_1$, $5*f_1$, …)
    - So the distance between each pair of resonances is 2× the lowest resonance's frequency

5. Deriving the vocal-tract filter for [ə]

- The **precise values** for the formants of [ə], or any vowel, will differ from person to person, since people have vocal tracts of different lengths
  - However, the **relative distance between** the formants for each vowel in a language is (reasonably) consistent from speaker to speaker

5. Deriving the vocal-tract filter for [ə]

- For a vocal tract that is 17.5 cm (0.175 m) long, calculate the first three resonance frequencies
  - These values are the **first three formants** of [ə] for this vocal tract (use "F" for formant frequencies)

5. Deriving the vocal-tract filter for [ə]

- For a vocal tract that is 17.5cm (0.175m) long, calculate the first three resonance frequencies
  - These values are the **first three formants** of [ə] for this vocal tract (use “F” for formant frequencies)

1st resonance frequency (F_1):

$F_1 = c/\lambda_1$ $\lambda_1 = 4L = 4*0.175\text{m} = 0.70\text{m}$

$c = 350\text{m/s}$

$F_1 = (350\text{m/s})/(0.70\text{m}) = 500/\text{s} = \mathbf{500Hz}$

F_2 and F_3 are odd-numbered multiples of F_1:

$F_2 = 3*500\text{Hz} =$ **1500Hz** | $F_3 = 5*500\text{Hz} =$ **2500Hz**

6. Putting source and filter together: [ə]

- Synthesized [ə] [vs_x.wav] (using same glottal source)
  (U Delaware Speech Research Tutorials)
  - Open this sound file in Praat
  - View the **waveform**
  - View the **spectrum** (window length=0.5 sec)

- Understanding this [ə] spectrum:
  - What determines the components' **frequencies**?
  - What determines their **amplitudes**?

6. Putting source and filter together: [ə]

Discussion

- Will [ə] produced by someone with a 17.5cm vocal tract always have **components** at 500Hz, 1500Hz, and 2500Hz?

6. Putting source and filter together: [ə]

- Will [ə] produced by someone with a 17.5cm vocal tract always have **components** at 500Hz, 1500Hz, and 2500Hz?

NO!

6. Putting source and filter together: [ə]

- **VERY IMPORTANT**
  - The vocal-tract **filter** is ***not*** a sound **source**
  - The filter is ***not*** adding components to the complex wave at its own resonance frequencies
  - All component **frequencies** in the source-filter model are coming from the **source**
  - The **filter**'s job is only to modify **amplitudes** of the components contributed by the source

6. Putting source and filter together: [ə]

- For next time:
  See web demo "The source/filter model"
  (Kevin Russell, U. Manitoba)
  - The **filter** associated with [ə] (for a vocal tract of a particular length)
  - The **spectrum** of the **glottal source** (for voicing produced with a particular f_0)
  - The **spectrum** of the **resulting [ə]**—which is the result of **modifying** the given glottal source with the given [ə] filter

6. Putting source and filter together: [ə]

- When we put source and filter together...
  - In the spectrum of the [ə], the **components** (also called *harmonics*) of the **source** wave near the **formants** (the frequencies corresponding to amplitude peaks in the **filter**) are **amplified** the most
  - Note that the amplification effect falls off gradually from a peak; this peak marks the actual formant/resonance frequency

6. Putting source and filter together: [ə]

- When we put source and filter together...
  - The **frequencies** of the components of the resulting sound are determined by the _____
  - The **amplitudes** of the components of the resulting sound are determined by the _____

6. Putting source and filter together: [ə]

- When we put source and filter together…
  - The **frequencies** of the components of the resulting sound are determined by the **source**
  - The **amplitudes** of the components of the resulting sound are determined by the **filter**

6. Putting source and filter together: [ə]

- ALWAYS REMEMBER:
  Formants are related to **vocal-tract** size and shape,
  NOT to f_0 of phonation!

  (a) The same vowel can be sung on different pitches

  (b) Different vowel categories can be sung on the same pitch

  - If the glottal source determined the formant frequencies, or vice-versa,
    (a) and (b) would be impossible

7. Terms and abbreviations for frequencies

When we are talking about **speech acoustics**:

- The abbreviation $\boldsymbol{f_0}$ (or F0, F_0) is used for the **fundamental frequency** of the **glottal source**
- The abbreviation **H(n)** is often used for the *n*th **resonance** (*n*th "**harmonic**"=whole-number multiple) of the **glottal source**
    - The glottal source has components at whole number multiples of the lowest resonance, so the lowest resonance = f_0 (remember GCD?)
    - Thus, for a voiced speech sound, $\boldsymbol{f_0}$ **= H1**

7. Terms and abbreviations for frequencies

When we are talking about **speech acoustics**:

- The abbreviation **F1** is used for the **first formant**, that is, the **first resonance frequency** of the vocal-tract **filter**
- **F2, F3,** etc. are used for the 2nd, 3rd, etc, **formants**, that is, **subsequent resonance frequencies** of the vocal-tract **filter**
- *Comprehension check:* What is the relationship between f_0 (F0) and F1 in a vowel?

7. Terms and abbreviations for frequencies

When we are talking about waves **in general**:

- $\boldsymbol{f_0}$ (sometimes F0, F_0) is used for the **fundamental frequency** of any wave, simple or complex
- $\boldsymbol{f_n}$ is used for the *n*th resonance frequency of a standing-wave system

So:

- *IF* the object under analysis is a sound **source**,
- and *IF* higher resonances are all some **multiple** of the first one,

THEN $f_0 = f_1$