Modeling vowel formants —
• Multiple-tube model
• Perturbation theory

Background reading:

- *V&C, Ch 4, "The sounds of vowels"*
- *V&C, Ch 12, sec 12.1, "Movements of the tongue and lips for vowels"*

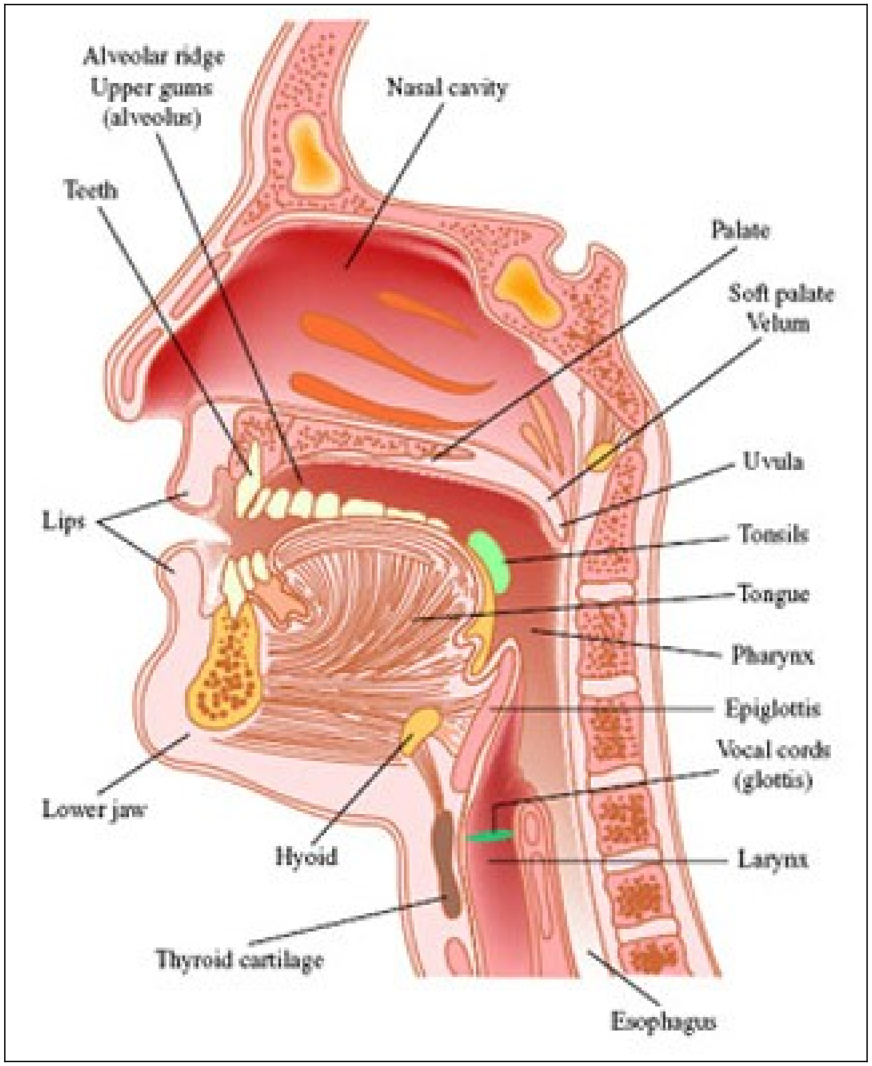
0. Today's objectives

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

- Use the multiple-tube model to predict numerical formant frequencies for [ɑ] and [i]
- Use perturbation theory to predict relative formant frequencies for [i ɑ u]
  - Vowel constrictions as tube perturbations
  - Nodes and antinodes in the vocal-tract tube
  - Perturbation "rules of thumb"
  - *Try at home:* Predict formants for [i ɑ u]

1. Background: The vocal tract

- Diagram of the human vocal tract (MIT Open CourseWare)



Pay particular attention to:

- Lips
- (Hard) palate
- Velum
- Uvula
- Pharynx

See also Fig. 11.2 in *V&C* (p 116)

2. Review: Vowels as vocal-tract tubes

[ɑ]

[i]

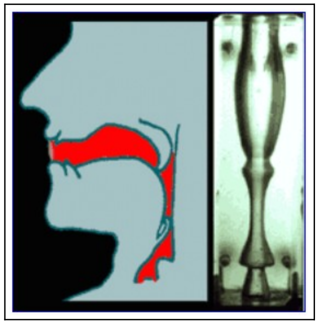
[u]

- How are these vowels described (in terms of height, etc.)?

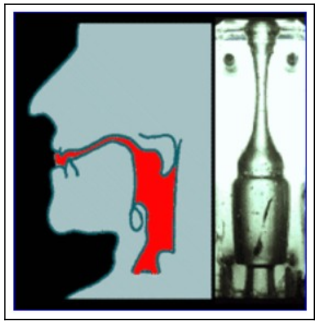
2. Review: Vowels as vocal-tract tubes

(images from Exploratorium)

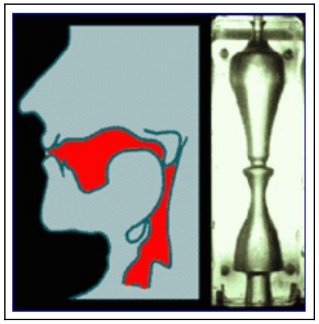
[ɑ] low back unrounded



[i] high front unrounded



[u] high back round



- Where does each vowel have a **constriction** (narrowing)? How is each **vocal-tract shape** as a series of **tubes**?

2. Review: Vowels as vocal-tract tubes

[ɑ]
- Narrowing in the **pharynx** (vertical part of vocal tract downstream of velum/uvula)
- **Wide tube** in front, **narrow tube** in back

[i]
- Narrowing at the **palate**
- **Wide tube** in front, **small narrow tube** in middle, **wide tube** in back

[u]
- Narrowing at the **velum** and **lips**
- (**Longer) wide tube** in front, **small narrow tube** in middle, **wide tube** in back + **lip rounding**

- We can use our understanding of **resonance frequencies in tubes** to model **vowel formants**

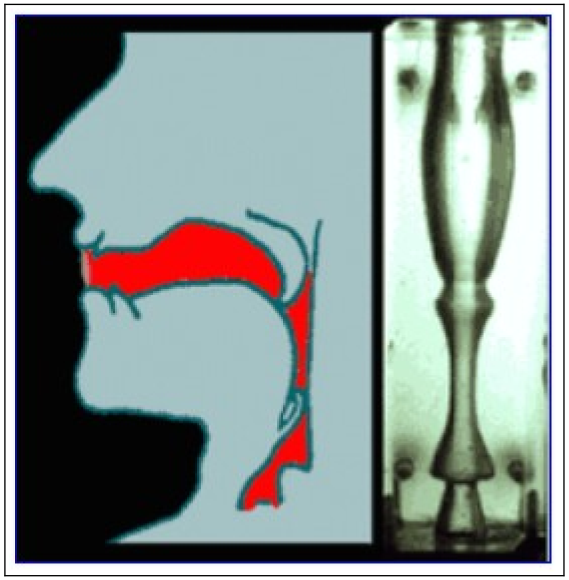
3. The mulitiple-tube model: Formants of [ɑ]

- Note: Central [a] is more typical in a 5-vowel system, while languages with more vowels (such as English) may use back [ɑ]
- The acoustics (and articulations) of these vowels are similar; we will follow *AAP* and discuss [ɑ] in the multiple-tube model

3. The mulitiple-tube model: Formants of [ɑ]

- [ɑ] has a narrowing in the **pharynx**

- We can model this as a series of ***two*** tubes
  - A **wide** tube in front (top of pharynx to lips)
  - A **narrow** tube in back (glottis to top of pharynx)



from [Exploratorium]

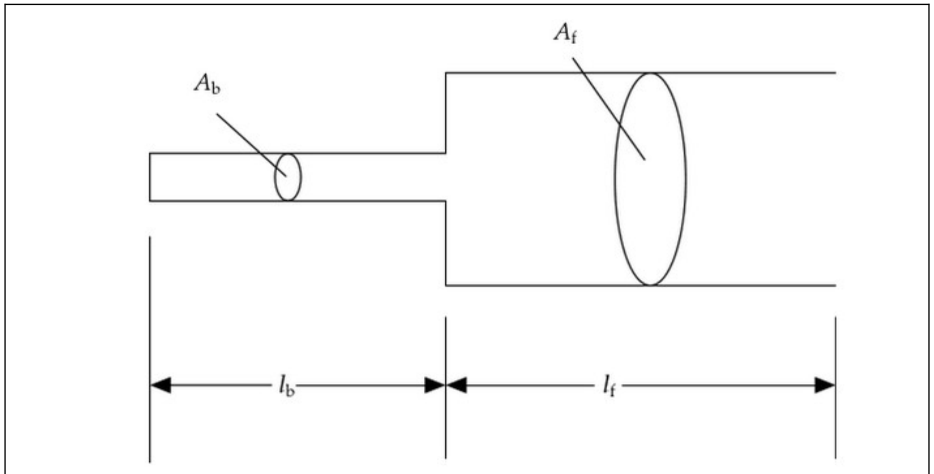


Figure 6.1 Two-tube model of the vocal tract that approximates the shape of the vocal tract for [ɑ].

AAP, Fig 6.1 (note lips on *right* side)

3. The mulitiple-tube model: Formants of [ɑ]

- We can model [ɑ] as a series of two tubes
  - A **wide** tube in front (top of pharynx to lips)
  - A **narrow** tube in back (glottis to top of pharynx)
- What are the **boundary conditions** for these tubes?

3. The mulitiple-tube model: Formants of [ɑ]

- What are the **boundary conditions** for these tubes?

  Front-cavity tube:

  - *open* at the lips
  - *closed* at the other end (as far as wave reflection is concerned — back-cavity tube has a smaller diameter)

  Back-cavity tube:

  - *closed* at the glottis
  - *open* at the other end (as far as wave reflection is concerned — front-cavity tube has *larger* diameter)

- **Nodes** or **antinodes**?

3. The mulitiple-tube model: Formants of [ɑ]

- What are the **boundary conditions** for these tubes?

  Front-cavity tube:

  - *open* at the lips
  - *closed* at the other end (as far as wave reflection is concerned — back-cavity tube has a smaller diameter)

  Back-cavity tube:

  - *closed* at the glottis
  - *open* at the other end (as far as wave reflection is concerned — front-cavity tube has *larger* diameter)

- Each tube → **node/antinode** (quarter-wave) system!

3. The mulitiple-tube model: Formants of [ɑ]

- **Front**-cavity tube is a node/antinode system
  Back-cavity tube is a node/antinode system

→ What do we need to know in order to calculate the **resonance frequencies** of these tubes?

3. The mulitiple-tube model: Formants of [ɑ]

- **Front**-cavity tube is a node/antinode system
  Back-cavity tube is a node/antinode system
  → What do we need to know in order to calculate the **resonance frequencies** of these tubes?
  - All we need is **tube length** and $f_n = (2n–1) \cdot c/4L$

- **Try it:** Assume a speaker with vocal tract = 16cm, back cavity = 7cm (and therefore front cavity = 9cm)

3. The mulitiple-tube model: Formants of [ɑ]

- **Front**-cavity tube is a node/antinode system
  Back-cavity tube is a node/antinode system
  - All we need is **tube length** and $f_n = (2n-1) \cdot c/4L$

- **Try it:** Assume a speaker with vocal tract = 16cm, back cavity = 7cm (and therefore front cavity = 9cm)
  - **Front**-cavity resonances (at c=350m/s)
    f_{f1} = **972Hz**, f_{f2} = 3*972Hz = **2916Hz**, ...
  - **Back**-cavity resonances (at c=350m/s)
    f_{b1} = **1250Hz**, f_{b2} = 3*1250Hz = **3750Hz**, ...

3. The mulitiple-tube model: Formants of [ɑ]

- What are the **formant frequencies of this speaker's [ɑ]** as predicted by this model?

3. The mulitiple-tube model: Formants of [ɑ]

- What are the **formant frequencies of this speaker's [ɑ]** as predicted by this model?
  - Remember: Formants are the **resonances of the vocal tract** when a vowel is produced
- So...what are the resonance frequencies produced by this two-tube system **as a whole**?
  - **Front**-cavity resonances (at *c*=350m/s)
    f_{f1} = **972Hz**, f_{f2} = 3*972Hz = **2916Hz**, ...
  - **Back**-cavity resonances (at *c*=350m/s)
    f_{b1} = **1250Hz**, f_{b2} = 3*1250Hz = **3750Hz**, ...

3. The mulitiple-tube model: Formants of [ɑ]

- What are the **formant frequencies of this speaker's [ɑ]** as predicted by this model?
  - Remember: Formants are the **resonances of the vocal tract** when a vowel is produced
- So…what are the resonance frequencies produced by this two-tube system **as a whole**?

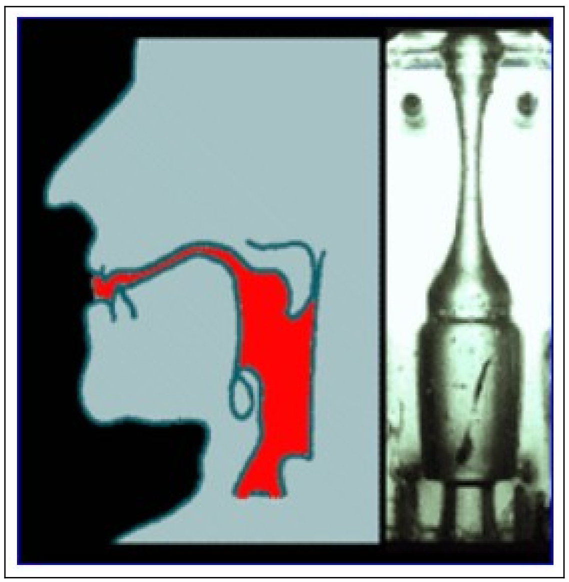
  → The front and back tube resonances all contribute! (List them in increasing order of *f*)

  F1 = **972Hz** (f_{f1}) F3 = **2916Hz** (f_{f2})

  F2 = **1250Hz** (f_{b1}) F4 = **3750Hz** (f_{b2})

4. The mulitiple-tube model: Formants of [i]

- [i] has a narrowing at the **palate**
- We can model this as a series of ***three*** tubes
  - A **wide** tube in front (palate to lips)
  - A **small narrow** tube at the palate
  - A **wide** tube in back (glottis to palate)



from Exploratorium

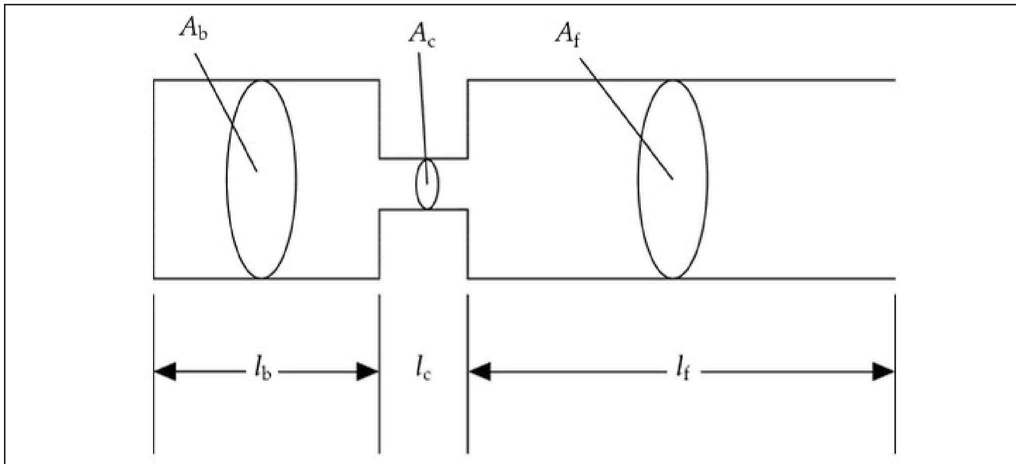


Figure 6.3 Tube model of vocal tract configurations that have a short constriction at some point in the vocal tract.

AAP, Fig 6.3 (note lips on *right* side)

4. The mulitiple-tube model: Formants of [i]

- We can model [i] as a series of three tubes
  - A **wide** tube in front (palate to lips)
  - *A **small narrow** tube at the palate*
  - A **wide** tube in back (glottis to palate)
- What are the **boundary conditions** for the **wide** tubes?
- *The small narrow tube is modeled as a Helmholtz resonator; we will return to this shortly*

4. The mulitiple-tube model: Formants of [i]

- What are the **boundary conditions** for these tubes?

  Front-cavity tube:
  - *open* at the lips
  - *closed* at the other end (as far as wave reflection is concerned — narrow center tube has a *smaller* diameter)

  Back-cavity tube:
  - *closed* at the glottis
  - *closed* at the other end (as far as wave reflection is concerned — narrow center tube has a *smaller* diameter)

- **Nodes** or **antinodes**?

4. The mulitiple-tube model: Formants of [i]

- What are the **boundary conditions** for these tubes?

Front-cavity tube:

- *open* at the lips
- *closed* at the other end (as far as wave reflection is concerned — narrow center tube has a *smaller* diameter)

→ **Node/antinode** (***quarter***-wave) system

Back-cavity tube:

- *closed* at the glottis
- *closed* at the other end (as far as wave reflection is concerned — narrow center tube has a *smaller* diameter)

→ **Antinode/antinode** (***half***-wave) system

4. The mulitiple-tube model: Formants of [i]

- **Front**-cavity tube is a node/antinode system
  Back-cavity tube is an antinode/antinode system

→ What do we need to know in order to calculate the **resonance frequencies** of these tubes?

4. The mulitiple-tube model: Formants of [i]

- **Front**-cavity tube is a node/antinode system
  Back-cavity tube is an antinode/antinode system
  - → What do we need to know in order to calculate the **resonance frequencies** of these tubes?
  - \- We need **tube length** and the two formulas
    $f_n = (2n–1) \cdot c/4L$ and $f_n = n \cdot c/2L$
- **Try it:** Assume a speaker with vocal tract = 16cm, back cavity = 10cm, narrow central constriction = 3cm (and therefore front cavity = 3cm)

4. The mulitiple-tube model: Formants of [i]

- **Front**-cavity tube is a node/antinode system
  Back-cavity tube is an antinode/antinode system
  - We need **tube length** and the two formulas
    $f_n = (2n-1) \cdot c/4L$ and $f_n = n \cdot c/2L$
- **Try it:** Assume a speaker with vocal tract = 16cm, back cavity = 10cm, constriction = 3cm, front = 3cm
  - **Front**-cavity resonances (at *c*=350m/s)
    f_{f1} = **2917Hz** , f_{f2} = 3*2917Hz = **8750Hz**, ...
  - **Back**-cavity resonances (at *c*=350m/s)
    f_{b1} = **1750Hz** , f_{b2} = **2***1750Hz = **3500Hz**, ...

4. The mulitiple-tube model: Formants of [i]

- Now we can look at the narrow central constriction (at the palate for [i])
  - This constriction also has its own resonance frequencies, as a (very short!) tube, but they are too high to matter for speech analysis
- This constriction plus the back cavity form a **Helmholtz resonator**, which also has a (single!), low-frequency resonance
  - A Helmholtz resonator involves a small volume of air that oscillates into and out of a larger one
  - For an animation, see "Helmholtz Resonance" (UNSW); scroll down to the *bottle* animation

4. The mulitiple-tube model: Formants of [i]

- There is a formula for calculating the Helmholtz resonance in this three-tube vowel system, based on the relative volumes of the central constriction and the back tube (see *AAP*, p 135, equation 6.2)
  - You do not need to know this formula
  - We will just **estimate** the Helmholtz resonance in a high vowel as being around **200Hz**

4. The mulitiple-tube model: Formants of [i]

- Putting it all together:
  - **Front**-cavity resonances (at *c*=350m/s)
    f_{f1} = **2917Hz**, f_{f2} = 3*2917Hz = **8750Hz**, ...
  - **Back**-cavity resonances (at *c*=350m/s)
    f_{b1} = **1750Hz** , f_{b2} = **2***1750Hz = **3500Hz**, ...
  - Helmholtz resonance = around **200Hz**

4. The mulitiple-tube model: Formants of [i]

- What are the **formant frequencies of this speaker's [i]** as predicted by this model?

4. The mulitiple-tube model: Formants of [i]

- What are the **formant frequencies of this speaker's [i]** as predicted by this model?

- List them all in order

  F1 = **200Hz** (Helmholtz resonance)

  F2 = **1750Hz** (f_{b1}) | note—a little low for a typical [i]

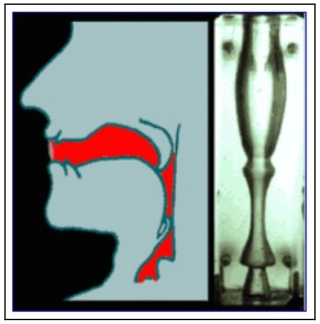
  F3 = **2917Hz** (f_{f1})

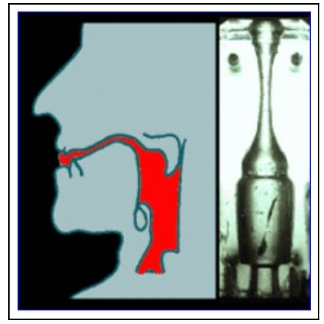
  …

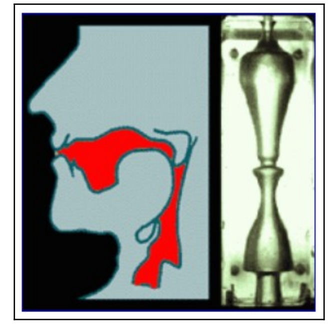
5. The mulitiple-tube model: Some final points

- Any vowel can in principle be modeled with this multiple-tubes model
  - It might need more than three tubes…
- For this class, we will focus on understanding the [ɑ]-type vowels and [i]-type vowels discussed above
- A **nomogram** is a graph that displays resonances as calculated for tube systems of particular sizes
  - See AAP, sec 6.1, reproduced on this handout
  - Check out the nomograms and try to relate them to today's discussion

6. Perturbation theory: Overview

[ɑ] 

[i] 

[u] 

low back unrd   high front unrd   high back round

- Using the **multiple-tube model**, we can model vowel **vocal-tract shapes** as a ***series*** of **tubes**
- With **perturbation theory**, we can model vowel **vocal-tract shapes** as a **perturbation** (modification) of a ***uniform*** **tube**
- Both models are simplifications, but are useful ways of understanding and predicting speech acoustics

6. Perturbation theory: Overview

- We can use our understanding of vowel articulations as **narrowings** in the vocal tract...
  - to model expected **deviations** in the **resonance frequencies** from those of a **uniform tube** ([ə])
  - and thereby **predict formants** of non-[ə] vowels

- Later in the course, we will also use perturbation theory to model **place-of-articulation** effects in consonant acoustics

6. Perturbation theory: Overview

- First step: Model the formant frequencies of [ə] (uniform vocal-tract tube)
- Then: Predict how formant frequencies will differ in other speech sounds
  - Find **where** there is a narrow region in the vocal tract—where the uniform tube is “perturbed”
  - Determine how a perturbation at that vocal-tract location should **change** the resonance frequencies of the tube
    - Note: Consider each resonance separately

7. Review: Vowel articulations for [i ɑ u]

- **Fill in the chart** on the next slide
  (download and save the PDF if you want to keep your work)
  - Complete the articulatory descriptions
  - State whether or not each vowel has a narrowing at the indicated point in the vocal tract

- Reminders
  - Narrowing at the *lips* when a vowel is *round*
  - Narrowing at the *palate* when a vowel is *high* and *front*
  - Narrowing at the *velum* when a vowel is *high* and *back*
  - Narrowing at the *pharynx* when a vowel is *low* (especially if also *back*)

7. Review: Vowel articulations for [i ɑ u]

- Fill in the chart: Where is there a constriction?

vowel	**description**	**lips?**	**palate?**	**velum?**	**pharynx?**
[i]		☐	☐	☐	☐
[ɑ]		☐	☐	☐	☐
[u]		☐	☐	☐	☐

"Description" = high/mid/low, front/central/back, round/unrounded

8. Review: Nodes & antinodes in v. t. tube

- Model the vocal tract as a uniform tube that is **open** at the lips and **closed** at the glottis
- Consider the **pressure** wave:
  - What are the **boundary conditions**? Node? Antinode?
  - For the **first three resonances**, where are *all* of the nodes and antinodes in the tube?
- Review: "Standing Sound Waves" animation of standing waves in air in a tube (Dan Russell, Penn State)
  - Graphs compare the pressure and displacement waves, and relate both to the air molecules in the tube

8. Review: Nodes & antinodes in v. t. tube

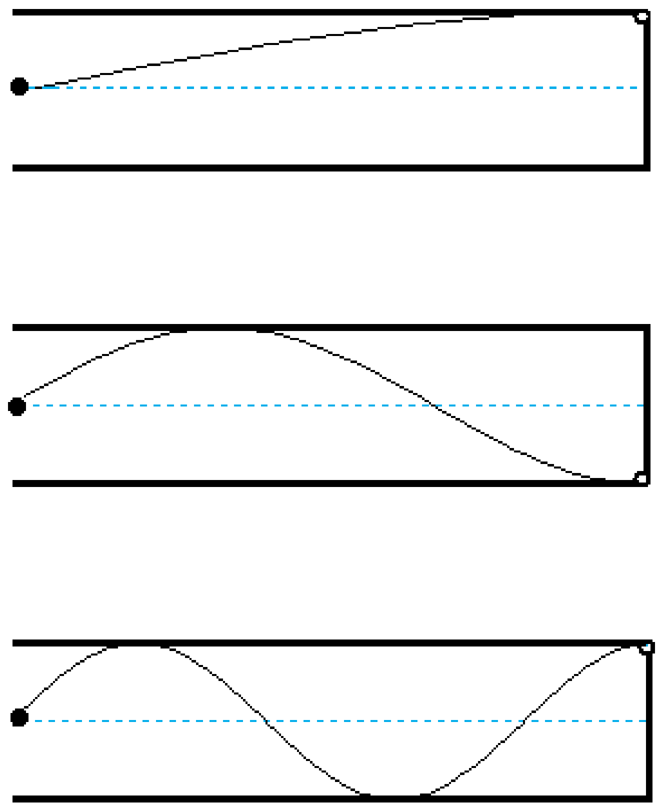
- What we saw before: The first three resonances
  Where are *all* the **nodes** and **antinodes** for each?

(lips) *(glottis)*

1st resonance

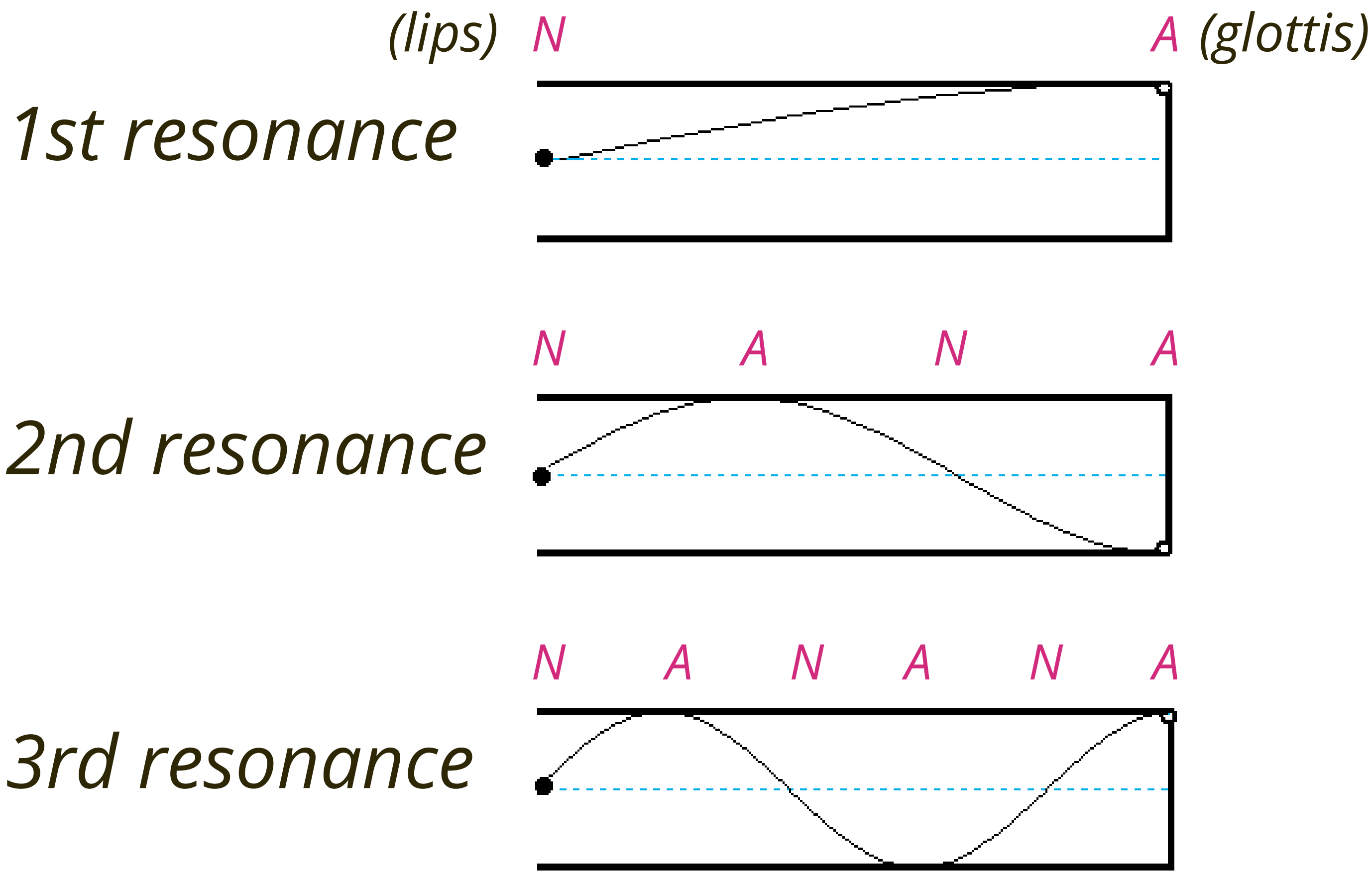
2nd resonance

3rd resonance



8. Review: Nodes & antinodes in v. t. tube

- What we saw before: The first three resonances
  Where are *all* the **nodes** and **antinodes** for each?



8. Review: Nodes & antinodes in v. t. tube

- Compare the diagrams in *AAP* Fig. 6.7 (p 139): These show the first four resonances of an open/closed tube, superimposed on the vocal tract

  Some points to consider here:

  - These diagrams look "backward" because they show the velocity wave (or displacement wave), which has *nodes* where the pressure waves we draw have *antinodes*
  - The points labeled V, V', V", etc. show antinodes on the **v**elocity waves, and so *nodes* on the pressure waves

  - Again, see "Standing Sound Waves" for pressure vs. displacement(=velocity) standing wave diagrams

9. Perturbing the tube

- Perturbation theory uses two “rules of thumb” to predict what will happen to a resonance frequency when a tube has a narrowing at a node or antinode
- Use the discussion in *AAP* sec 6.2 to fill in the blanks on the following slide with ***up* or *down***
  - Remember that the *AAP* “node”/”antinode” descriptions are given in terms of the **velocity** or displacement wave, not the **pressure** wave!

9. Perturbing the tube

- **Perturbation rules — MEMORIZE THIS**

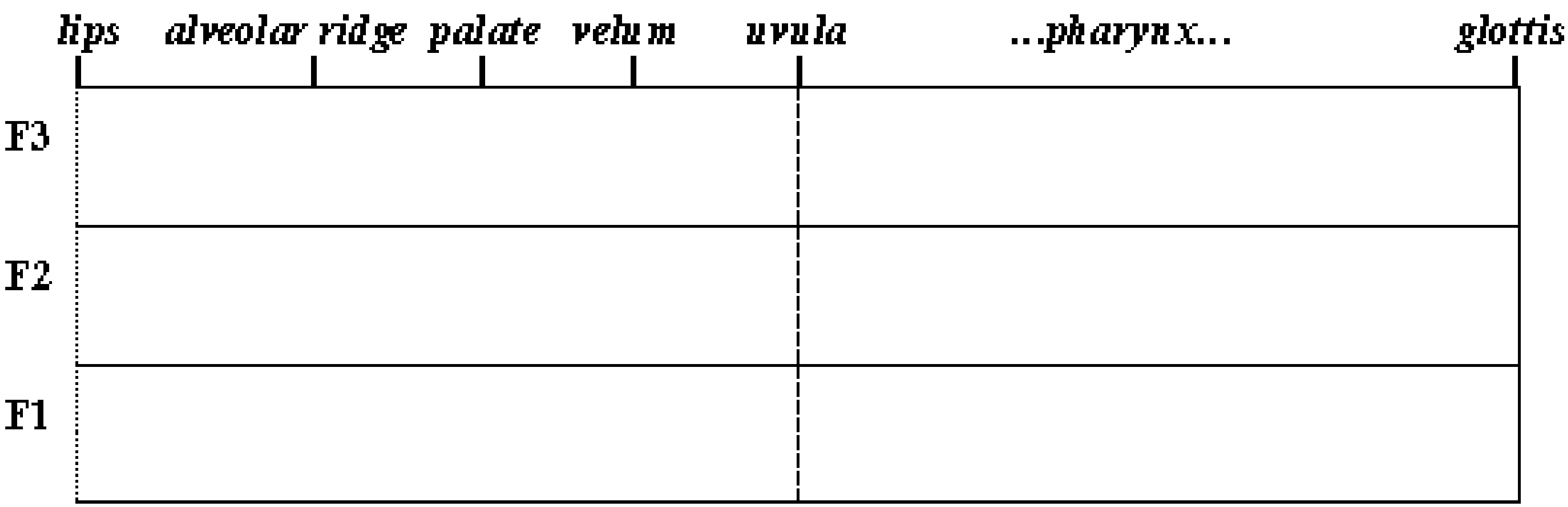
1. If there is a narrowing in the vocal tract near a velocity/displacement antinode = **pressure node**, the **formant frequency goes** ______.
2. If there is a narrowing in the vocal tract near a **pressure antinode** (velocity/displacement node), **formant frequency goes** ______.

9. Perturbing the tube

- Forming a constriction or narrowing in the vocal tract **affects each formant separately**
  - The effect on each formant depends on whether the narrowing is closer to a (pressure) node or closer to a (pressure) antinode **for that formant**
  - The **same point** in the vocal tract could be near a node for one formant and an antinode for another!

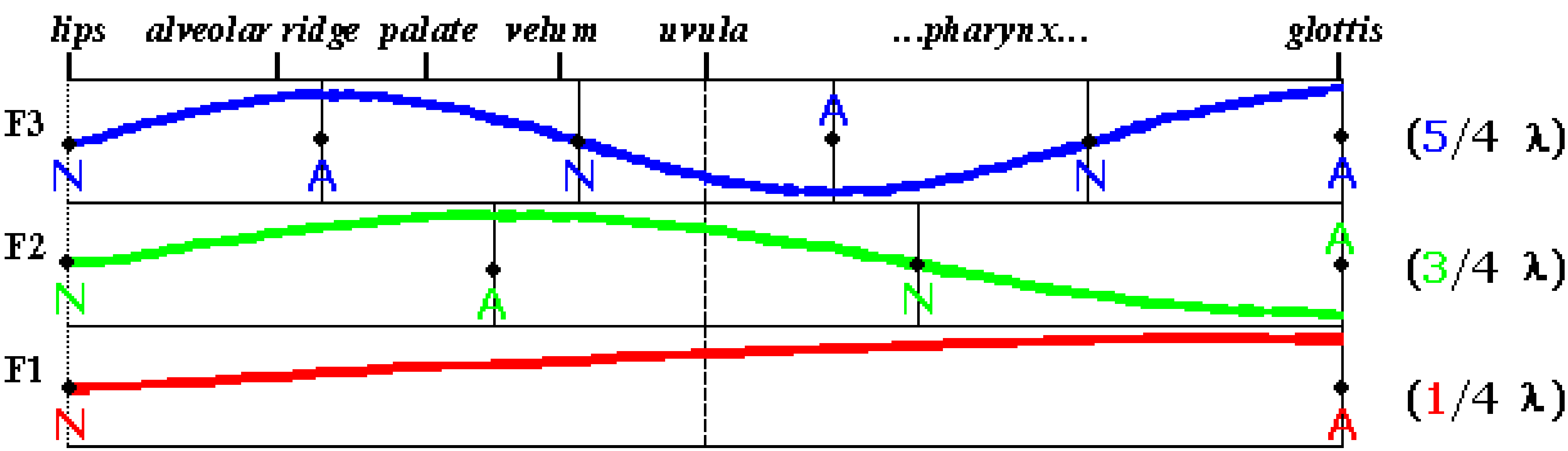
9. Perturbing the tube

- Articulatory landmarks in the vocal tract
  - The following diagram shows approximately where different vocal-tract structures lie along the length of the vocal tract



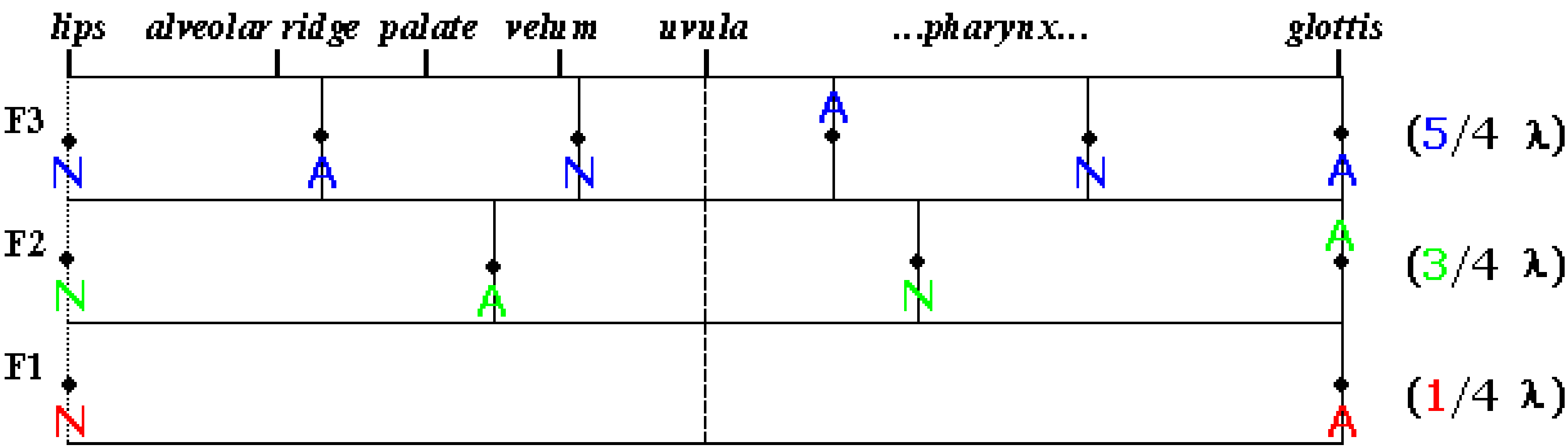
9. Perturbing the tube

- Articulatory landmarks in the vocal tract
  - Draw the first three standing (pressure) waves
  - Label their **nodes** and **antinodes**



9. Perturbing the tube

- Articulatory landmarks in the vocal tract
  - What we need to pay attention to is not the standing-wave diagram itself, but specifically where the **nodes** and **antinodes** are



9. Perturbing the tube

- This tube diagram of the vocal tract is based on *AAP* Fig 6.12 (p 151), for the "typical male speaker". Landmarks:
  - uvula at midpoint of vocal tract
  - alveolar ridge at 1/3 of the distance from lips to uvula
  - palate at 1/3 and velum at 2/3 of the distance from the alveolar ridge to the uvula

10. Predicting formants in non-[ə] vowels

- Now you have what you need to use perturbation theory to make predictions about vowel formants!
- Is a particular narrowing in the vocal tract *closer* to a (pressure) node, or *closer* to a (pressure) antinode?
  → This predicts what the formants will do

10. Predicting formants in non-[ə] vowels

Try it out — we'll check in next time

- Use what you know about:
  - vowel articulations (slide 34)
  - the perturbation rules (slide 40)
  - locations of the (pressure) nodes/antinodes in the vocal tract (slide 44)

to fill in the charts on the next two slides (except F2 for [u])

10. Predicting formants in non-[ə] vowels

- Should the formant be higher (↑) or lower (↓) than the equivalent formant in [ə] when there is a narrowing as indicated?

	lips	palate	velum	pharynx
F3	○ ↑ ○ ↓	○ ↑ ○ ↓	○ ↑ ○ ↓	○ ↑ ○ ↓
F2	○ ↑ ○ ↓	○ ↑ ○ ↓	○ ↑ ○ ↓	○ ↑ ○ ↓
F1	○ ↑ ○ ↓	○ ↑ ○ ↓	○ ↑ ○ ↓	○ ↑ ○ ↓

10. Predicting formants in non-[ə] vowels

- Should the formants **in these vowels** be higher (↑) or lower (↓) than the equivalent formant in [ə]?

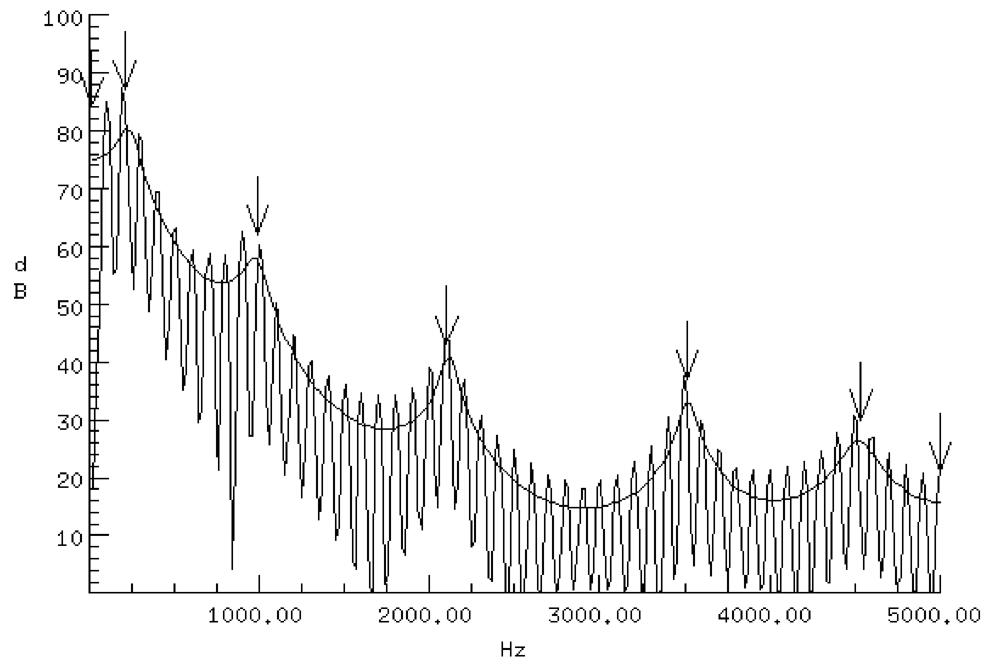
	[i]	[ɑ]	[u]
F3	○ ↑ ○ ↓	○ ↑ ○ ↓	○ ↑ ○ ↓
F2	○ ↑ ○ ↓	○ ↑ ○ ↓	*see below*
F1	○ ↑ ○ ↓	○ ↑ ○ ↓	○ ↑ ○ ↓

10. Predicting formants in non-[ə] vowels

- Why is it hard to make a prediction for F2 in [u] using perturbation theory?

10. Predicting formants in non-[ə] vowels

- Look at this [u] spectrum; formants are indicated with arrows (from [U Delaware Speech Research]())
  - What does F2 in [u] actually look like compared to [ə]? ([ə] produced by this synthesizer: F2=1550 Hz)



10. Predicting formants in non-[ə] vowels

Extension to **mid vowels**: [e], [o]

- [e] is less high and less front than [i]
  - Its formant frequencies are perturbed in the direction of [i], but not as far
- Likewise, [o] is less high and less back than [u]
  - Its formant frequencies are perturbed in the direction of [u], but not as far
  - Note: American English so-called “[u]” is more of a central vowel ([ʉ]) than a back one; [o] may be *further back* than [u]!