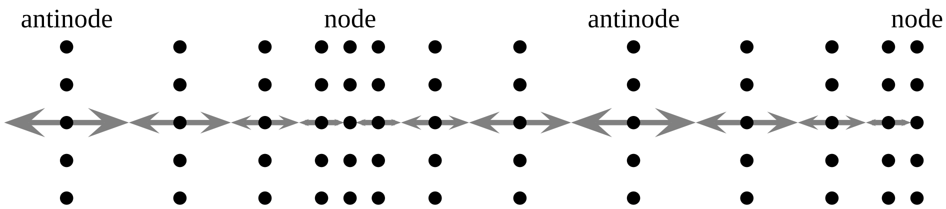
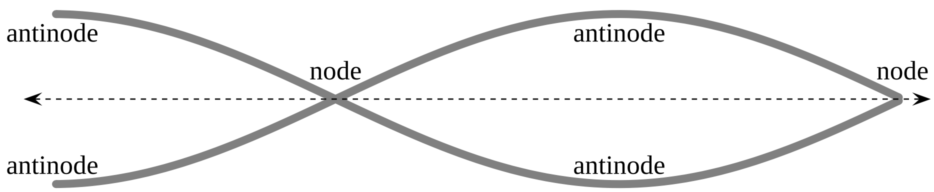
LING 210 **Handout: Perturbation Theory** **Spring 2009**

Monday, March 9th

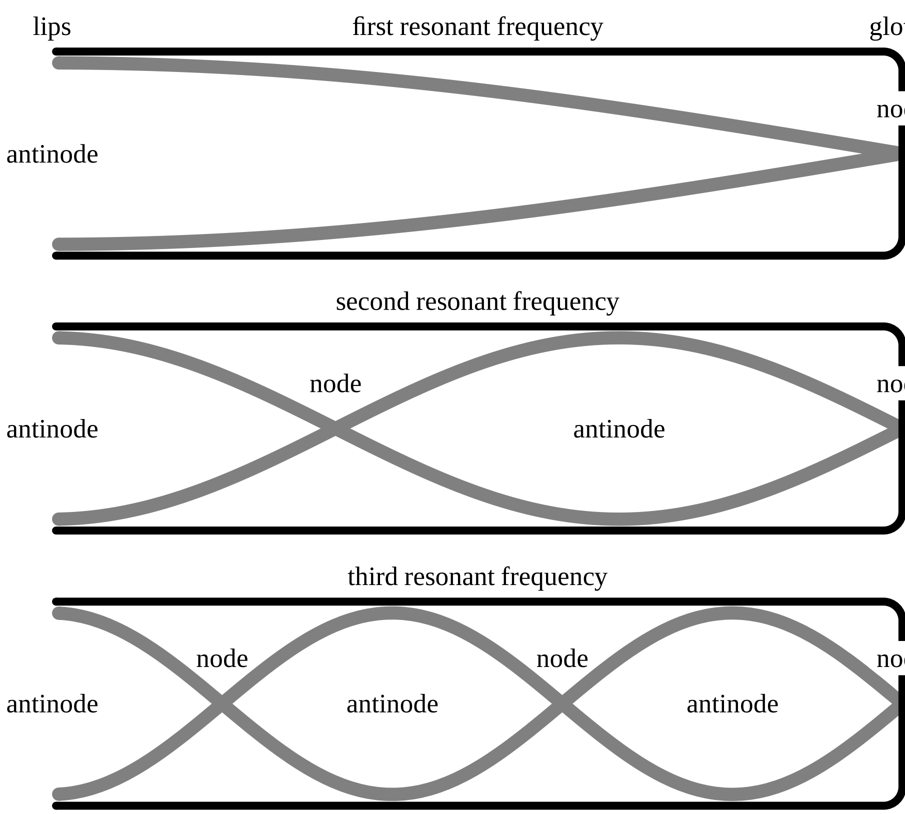
Sound waves are transmitted by air molecules banging against each other. In some parts of the wave, the molecules move very little, since they are close together. In other parts, the molecules move a lot, since they are farther apart. A displacement **node** is a part of a wave where the molecules essentially do not move at all because they are already packed together, while a displacement **antinode** is where density is at a minimum and the molecules move the most:



If we graph the wave as a function of molecular displacement, displacement nodes occur where the wave crosses the center line, and displacement antinodes occur at the peaks and valleys:

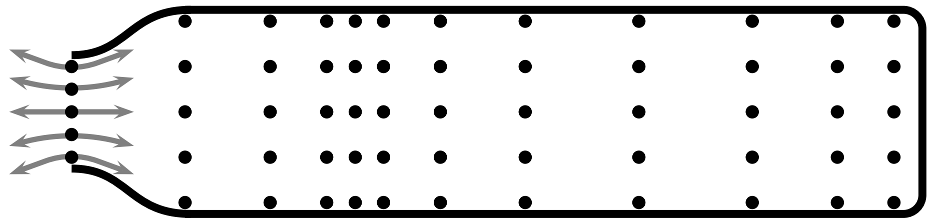


As we already know, a resonance wave fits inside a tube such that the molecules do not move very much at closed ends (because there is nowhere for them to move to), and they move maximally at open ends. So for a half-open tube, like our estimate for the human vocal tract, the first three resonant frequencies fit as shown in the following diagrams:



Note that for every resonant frequency of the mouth, there is an antinode at the lips. When a constriction is made at the lips (i.e., the lips are rounded), we find that every vowel formant is lower than for an unconstricted tube of the same length. That is, the resonant frequencies of the tube all decrease when a constriction is made at the open end.

Consider what is going on with the air molecules at the open end of the tube: as an antinode, a particular molecule has a large distance to travel before bumping into another molecule, which is required in order to transmit the sound wave. When a constriction is made, the molecule's path is obstructed, and it must travel even further than usual:



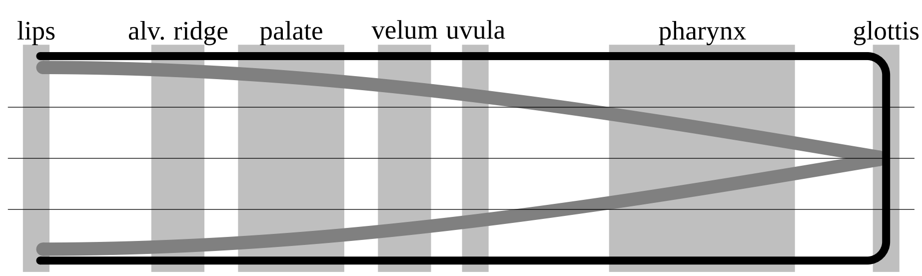
Because the molecules have farther to travel, it takes them longer than usual to vibrate a full cycle. This means the period of the wave will increase. When the period increases, the frequency decreases. Thus, **a constriction at a displacement antinode of a resonant frequency decreases the effective resonant frequency**.

At a node, travel time is not an issue, because the molecules are very close together. A constriction at a node causes the molecules to be packed even tighter. This causes the vibrations to be more responsive and happen quicker (this is the same reason sound waves travel faster through solids than gases). Since the vibrations are happening faster, the period of the wave will decrease. When the period decreases, the frequency increases. Thus, **a constriction at a displacement node of a resonant frequency increases the effective resonant frequency**.

In addition, **the raising/lowering effect on a resonant frequency is increased if the constriction is tighter**: for an antinode, the distance traveled becomes longer; for a node, the molecules become even more compact, and thus, more responsive.

These principles are the basis of **Perturbation Theory**, and using them, we can predict what sort of resonant frequencies a given vocal tract configuration will have and how they compare to those of other configurations. Despite its lack of quantitative precision, Perturbation Theory has an advantage over tube models: in Perturbation Theory, it is relatively easy to qualitatively describe the resonant frequencies for rather complex articulations, including the third, fourth, and higher resonant frequencies.

In order to effectively use Perturbation Theory, we need to know how the geography of the vocal tract lines up with nodes and antinodes of the resonant frequencies. The following diagram shows the first resonant frequency of the vocal tract, with vertical shading marking some of the important landmarks the vocal tract and horizontal lines dividing the wave into regions based on their proximity to nodes and antinodes.



A uniform tube has a first resonant frequency around 500 Hz, so this is the baseline value for F1.

The low vowels [æ] and [ɑ] are made by lowering the tongue, constricting the pharynx, which is located closer to a node, so this constriction causes F1 to raise above 500 Hz. Because [ɑ] is farther back than [æ], the pharyngeal constriction is slightly more extreme for [ɑ], so [ɑ] is predicted to have a somewhat higher F1 than [æ]. In addition, [æ] involves a slight narrowing near the hard palate, which is near an antinode, causing more lowering of F1 for [æ].

The high front vowel [i] is made by raising the front of the tongue towards the front of the hard palate, which is near an antinode. There is also slight pharyngeal expansion as the tongue moves upward. Both of these articulations (antinode constriction, node expansion) result in lowering F1.

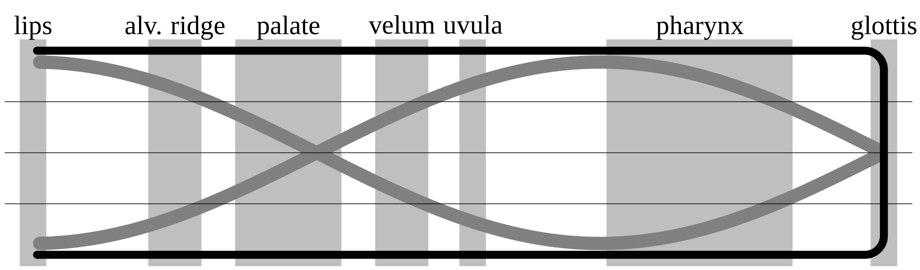
The front mid vowel [e] is intermediate between [i] and [ə], so we predict that it will have F1 less than 500 Hz, but not as low as for [i]. Similarly, [ɛ] is intermediate between [æ] and [ə], so it is predicted to have F1 higher than 500 Hz, but not as high as for [æ].

The high back vowel [u] is made by raising the back of the tongue towards the velum. There is basically no pharyngeal expansion, since the tongue doesn't move as for as it does for [i]. Since the velum is farther back than the palate, a velar constriction is farther away from the antinode at the lips, so the lowering effect isn't as great for [u] as it is for [i]. However, [u] is also rounded, which involves a constriction at the lips (an antinode), and therefore F1 encounters some further lowering. So we can say that [u] has roughly the same F1 as [i], though the exact comparison would depend on the degree of lip rounding.

The mid back vowel [o] is intermediate in height between [u] and [ə], so we predict that it will have F1 less than 500 Hz, but not as low as for [u]. Similarly, [ɔ] is intermediate between [ɑ] and [ə], so it is predicted to have F1 higher than 500 Hz, but not as high as for [ɑ].

Summary of F1: [i u] < [e o] < [ə] < [ɛ ɔ] < [æ] < [ɑ]
500 Hz

The next diagram is for the second resonant frequency, which for a uniform tube is about 1500 Hz:



The pharyngeal constriction of the low vowels occurs close to an antinode for F2, so [ɑ] and [æ] are predicted to have F2 lower than 1500 Hz. As with F1, the degree of constriction is less for [æ], so its F2 will not be as low as the F2 of [ɑ]. In addition, the palatal constriction for [æ] occurs directly on a node, which will cause F2 to raise significantly. The combined effect for [æ] is that F2 is slightly higher than for [ə].

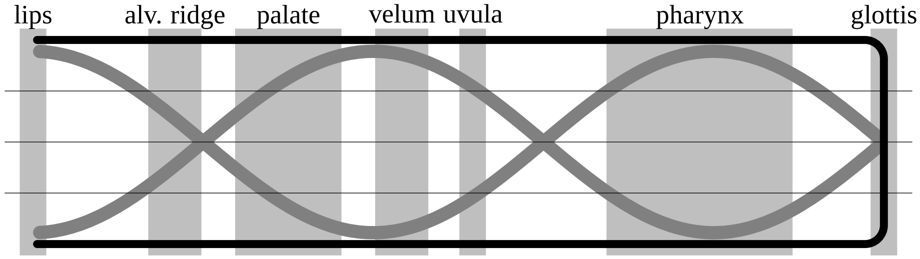
The palatal constriction and pharyngeal expansion for the high front vowel [i] cause dramatic raising of F2 well above 1500 Hz. Indeed, a very high F2 is one of the most salient acoustic characteristics of [i]. The front mid vowel [e] is intermediate in frontness between [i] and [ə], so we predict that it will have F2 greater than 1500 Hz, but not as high as for [i]. Similarly, [ɛ] is intermediate between [æ] and [e], so it is predicted to have F1 higher than for [æ], but lower than for [e].

The velar constriction for [u] is near the border between a node and antinode, though closer to an antinode, so it has a slight raising effect on F2. However, [u] involves significant lip rounding, which is directly located at an antinode, and this counteracts the mild raising effects, causing F2 to actually be lower than 1500 Hz. The effect of this lip rounding is greater than the effect of the pharyngeal constriction in [ɑ], so the F2 for [u] is lower than for [ɑ].

The mid back vowels [o] and [ɔ] have about the same backness and rounding as [u], so we predict that they will also have F2 less than 1500 Hz. Their rounding isn't quite as extreme as for [u], so their F2s won't be quite as low. However, the lower tongue position requires slight pharyngeal constrictions which counteract this effect, given the back round vowels roughly the same value of F2.

Summary of F2: [u o ɔ] < [ɑ] < [ə] < [æ] < [ɛ] < [e] < [i]
1500 Hz

Finally, we have a diagram for the third resonant frequency, which has a value around 2500 Hz for a uniform tube:



The pharyngeal constriction of the low vowels occurs directly on an antinode for F3, so [ɑ] and [æ] are predicted to have F3 lower than 2500 Hz. As with F1 and F2, the degree of constriction is less for [æ], so its F3 will not be as low as the F3 of [ɑ]. However, the palatal constriction for [æ] occurs near an antinode, which will cause F3 to lower. The combined effect for [æ] is that F3 is about the same as for [ɑ], which is lower than 2500 Hz.

The palatal constriction for the high front vowel has little effect on F3, but pharyngeal expansion for causes slight raising of F3 above 2500 Hz.

The front mid vowels [e] and [ɛ] have very little pharyngeal expansion, and their constrictions are at the palate, which is intermediate between a node and antinode. Thus, we expect that their F3's will be approximately the same as for [ə].

The velar constrictions for the back round vowels [u o ɔ] occur near an antinode, which has a lowering effect on F3. In addition, the lip rounding for these vowels causes F3 to be even lower than 2500 Hz. This effect is stronger for [u] and gets progressively weaker as the vowels get lower and less round.

Summary of F3: [u] < [o] < [ɔ] < [ɑ æ] < [ə] < [ɛ] < [e] < [i]
2500 Hz

We have yet to talk about the rhotacized vowels [ɝ ɚ]. They involve numerous articulations (simultaneous lip, mid-palatal, and pharyngeal constrictions), making a tube model very complex and difficult to compute. However, with Perturbation Theory, we can make rough predictions about the formants of these vowels:

F1 is predicted to be somewhat lower than 500 Hz (two antinodes and a node)
F2 is predicted to be somewhat lower than 1500 Hz (two antinodes and a node)
F3 is predicted to be significantly lower than 2500 Hz (three antinodes).

And indeed, this is precisely what we find for [ɝ ɚ], especially the very low F3.