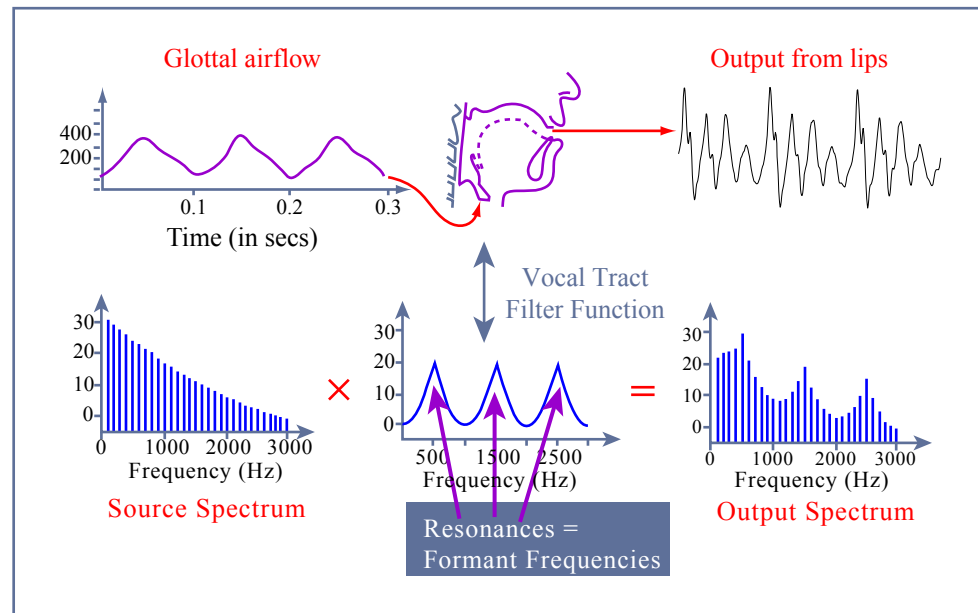


24.915/24.963

Linguistic Phonetics

The source-filter model of speech production



- Readings for next week: Ladefoged (1996) 'Elements of Acoustic Phonetics (2nd edn) ch. 10 from p.160, ch. 11 to p.183.
 - Ladefoged FFT, Ladefoged LPC
- optional: read the rest of chapter 11 (on LPC).
- Assignment: Write up a discussion of aspiration in Mandarin affricates, answering question 2 (due 10/6).

Source-Filter Model of Speech Production

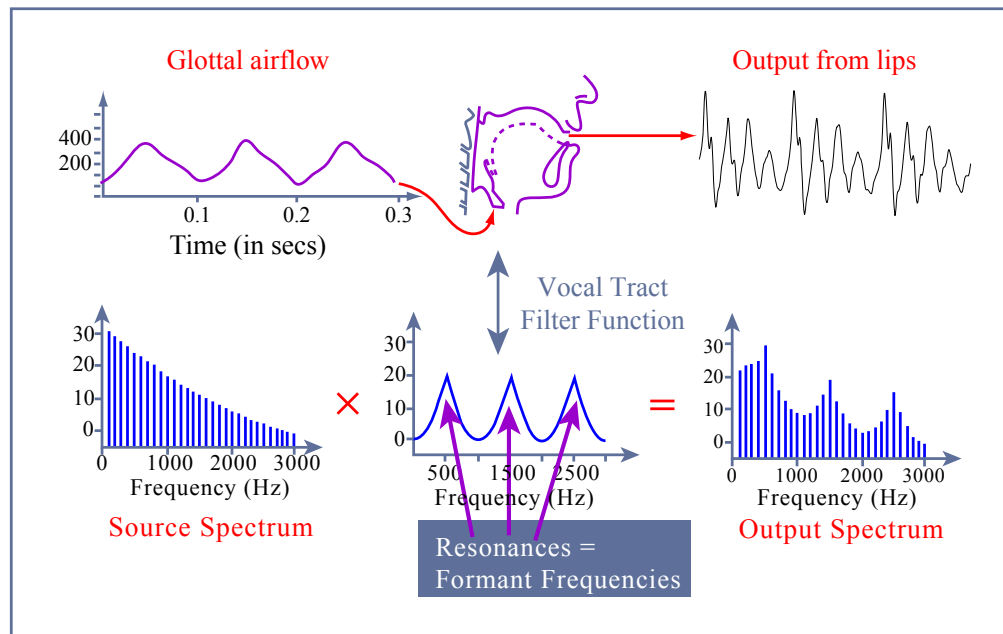


Image by MIT OCW.

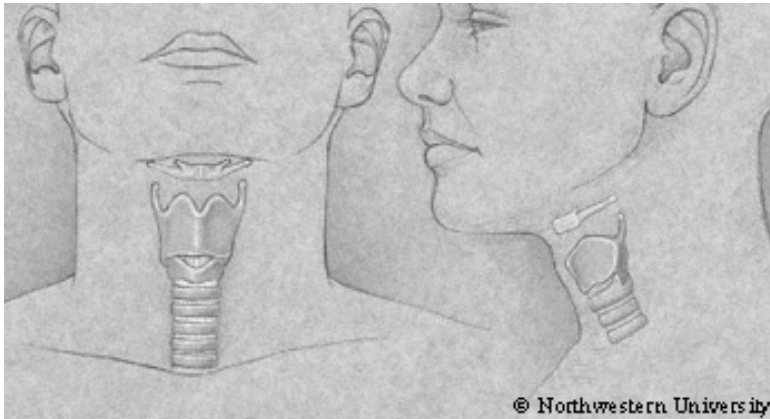
Sources

- A source is an input of acoustic energy into the speech production system.
- There are two basic types:
 - Voicing
 - Noise

Voice source

- Voicing is a periodic source produced by modulation of the airflow from the lungs by the vocal folds.
- The vocal folds are muscular folds located in the larynx.
- The space between the vocal folds is the glottis.

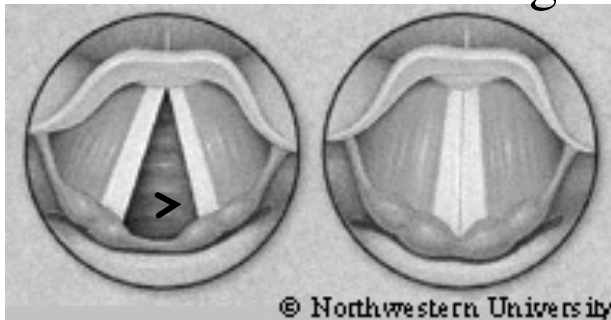
location of the larynx:



© Northwestern University. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>.

view of the vocal folds looking down into the larynx

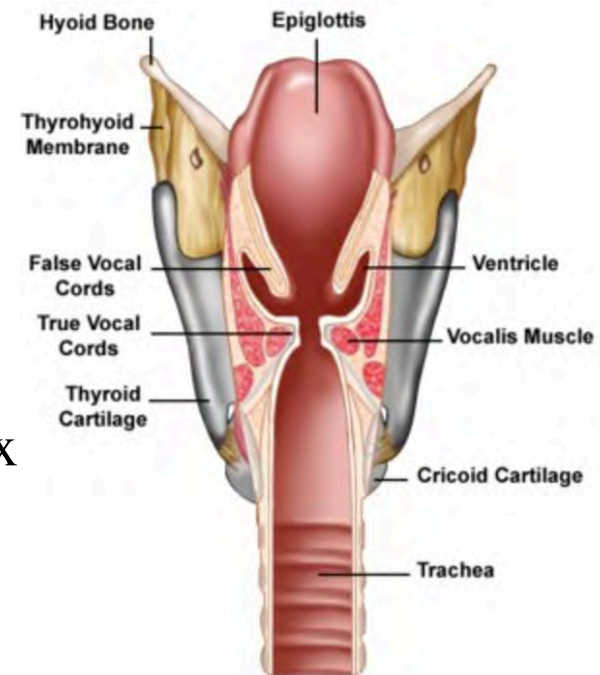
folds
open
glottis



folds
closed

© Northwestern University. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>.

mid-coronal section through the larynx:



© Greater Baltimore Medical Center. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <http://ocw.mit.edu/help/faq-fair-use/>

Voice source

- If the vocal folds are close together, then air pressure from the lungs can cause them to vibrate periodically, generating voicing.

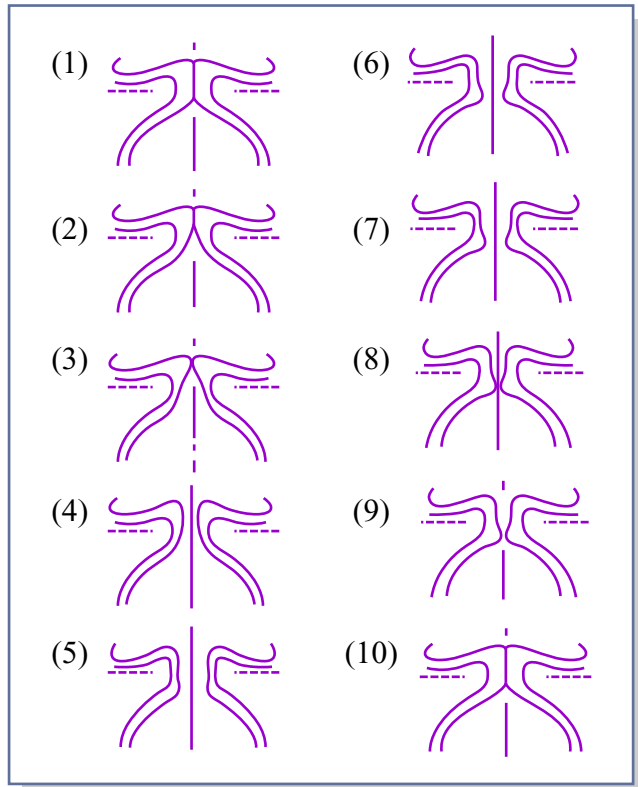


Image of human vocal folds removed due to copyright restrictions.

Image by MIT OCW.

Voicing source

- Vocal fold vibration produces a complex periodic wave whose spectrum contains energy at the fundamental frequency of laryngeal vibration and multiples of the fundamental frequency (harmonics).

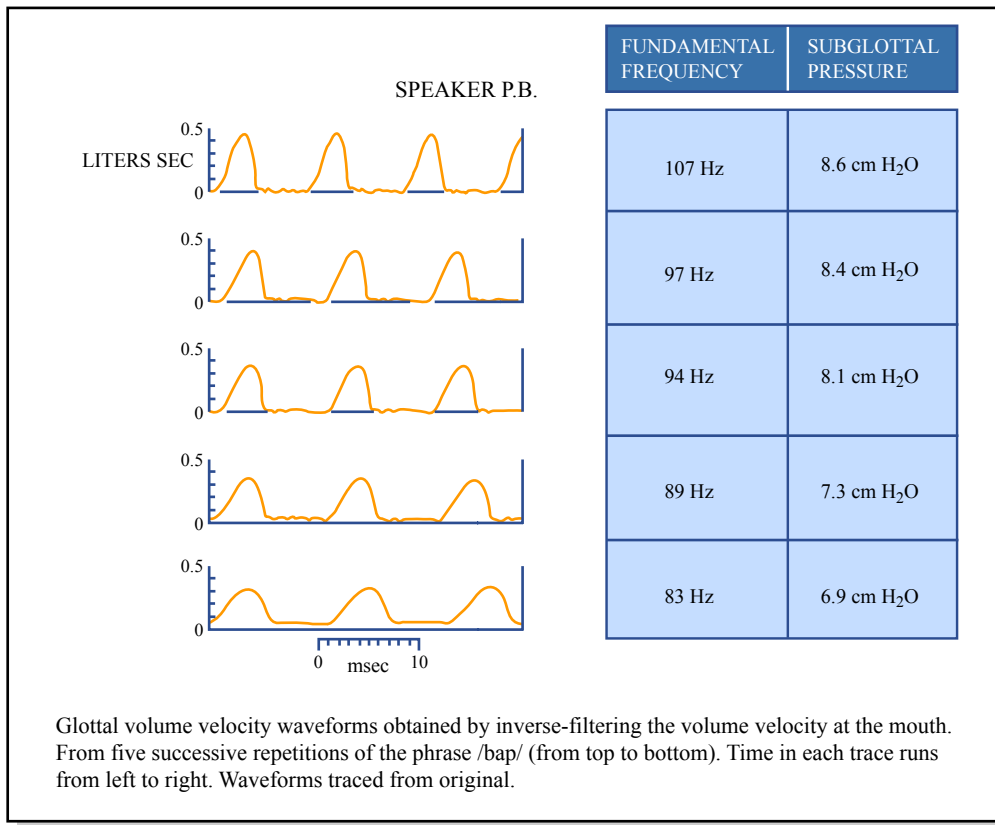
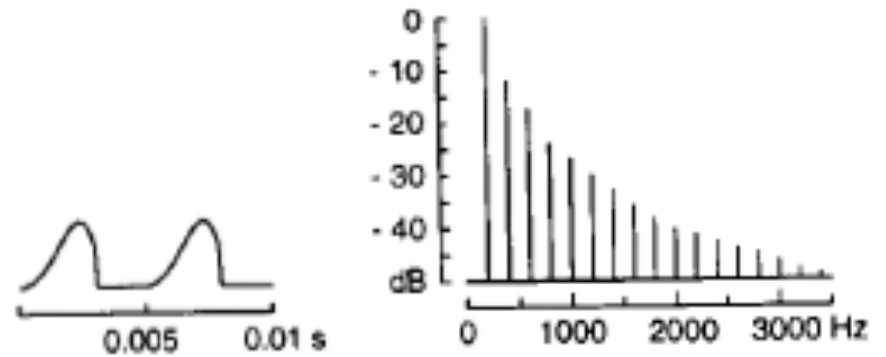


Image by MIT OCW.

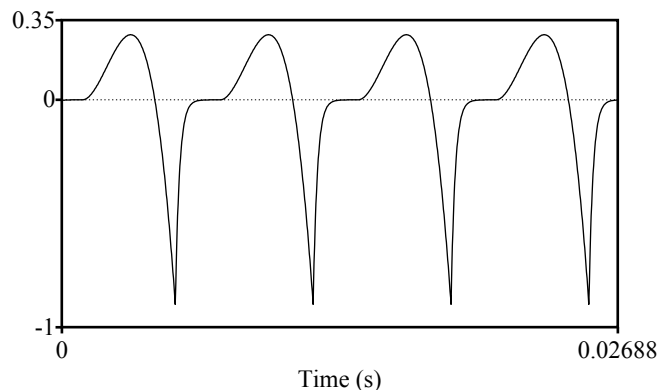
Adapted from Rothenberg, M. "The Glottal Volume Velocity Waveform During Loose and Tight Glottal Adjustments." *Proceedings of the VII International Congress of Phonetic Sciences* (1971): 380-388.



University of Chicago Press. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/help/faq-fair-use/>.

glottal volume velocity spectrum

acoustic wave of unfiltered source



Noise source

- Air turbulence generated at an obstruction involves random (aperiodic) pressure fluctuations over a wide range of frequencies.
- Noise generated at the glottis is called aspiration, while noise generated elsewhere is called frication.

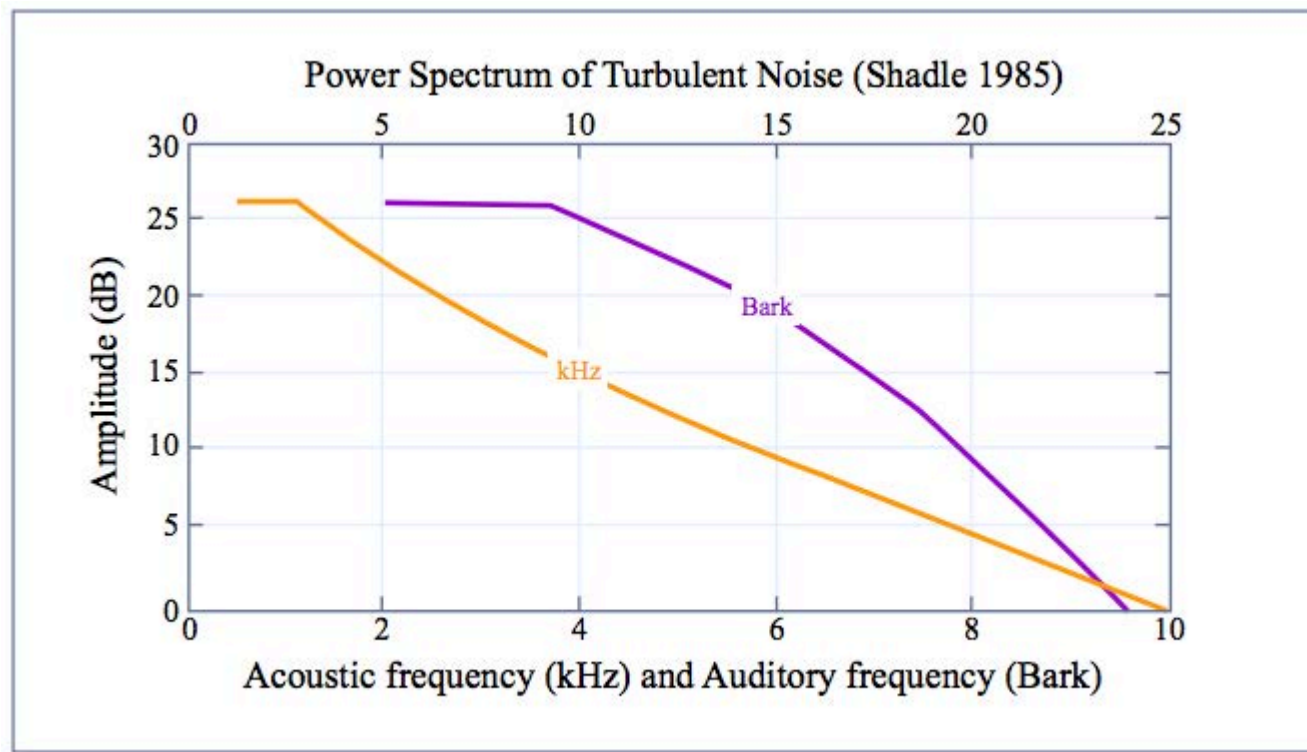


Image by MIT OCW. Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997.

Source

A sound may involve more than one source.

- E.g. a voiced fricative combines voicing and frication noise
- breathy voice can combine voicing and aspiration noise.
- voiceless fricatives can have noise generated at the glottis and at the supralaryngeal constriction.

Filters

- The vocal tract acts as a filter, modifying the source waveform.
- An acoustic filter is a device which passes certain frequencies and attenuates others.
- An important characteristic of a filter is its **transfer function** - the ratio of the output to the input depending on frequency.

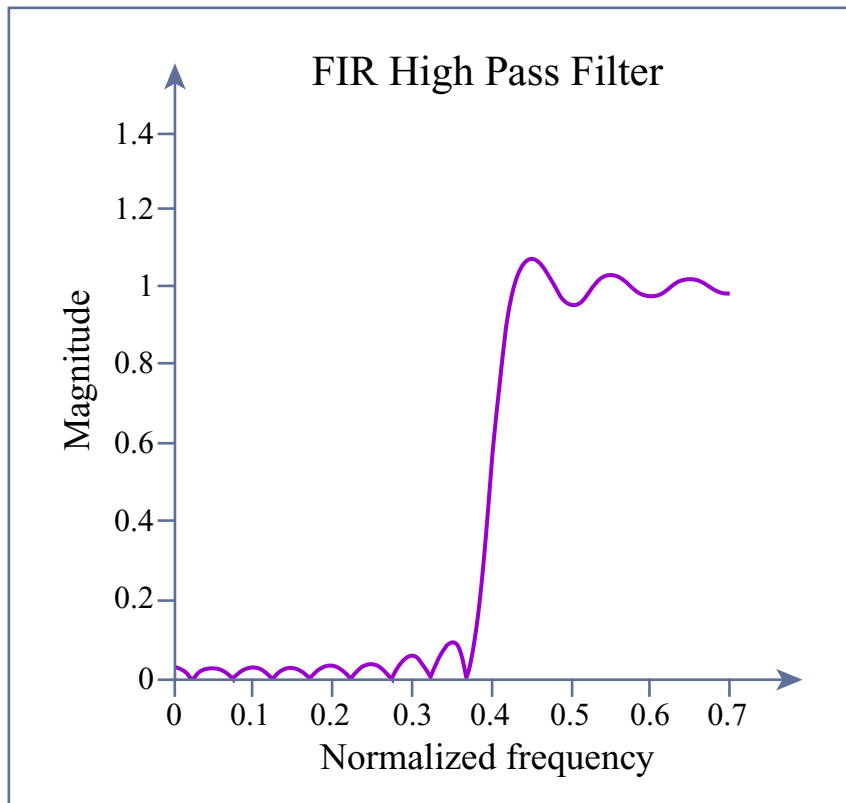


Image by MIT OCW.

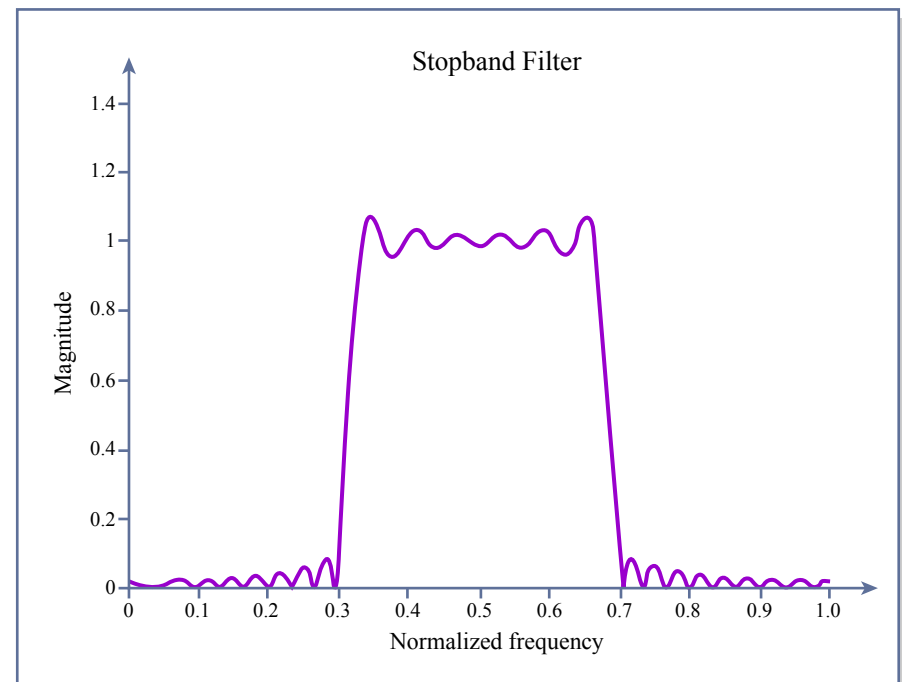


Image by MIT OCW.

The vocal tract as a filter

- The sound wave at some distance from the speaker is the result of filtering the source with the vocal tract filter, plus the radiation characteristics of the lips/nose.
- How does the vocal tract operate as a filter?
- What determines the characteristics of the filter?

Outline

- Resonators are filters.
- The column of air in the vocal tract is a resonator, hence a filter.
- The characteristics of the filter depend on the shape of the vocal tract - we will explore the relationship by considering simple cases.

Resonators

- A body, like a mass of air in a tube, naturally vibrates at one or more frequencies.
- If a pulse of energy is imparted to the air, it will vibrate at these natural frequencies.
- If a source (driving force) which is vibrating at a natural frequency of the body is applied to it (e.g. a tuning fork is held over the tube of air), the body will resonate with the source, i.e. vibrate strongly at the same frequency.
- If a source of a different frequency is applied to the body it will vibrate with less amplitude.

Resonators

- We can plot the response of a resonator to a range of input frequencies. Peaks mark the resonant frequencies.
- Resonators vary in the range of frequencies they will respond strongly to
 - a sharply tuned resonator only responds strongly to frequencies very close to the resonant frequency (it has a narrow bandwidth).
 - A resonator with a wide bandwidth will vibrate strongly in response to a wider range of input frequencies.

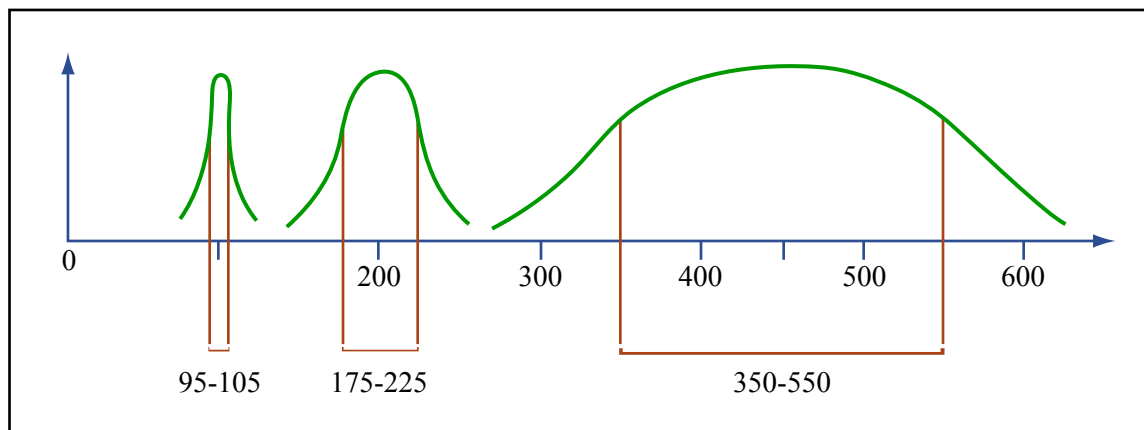


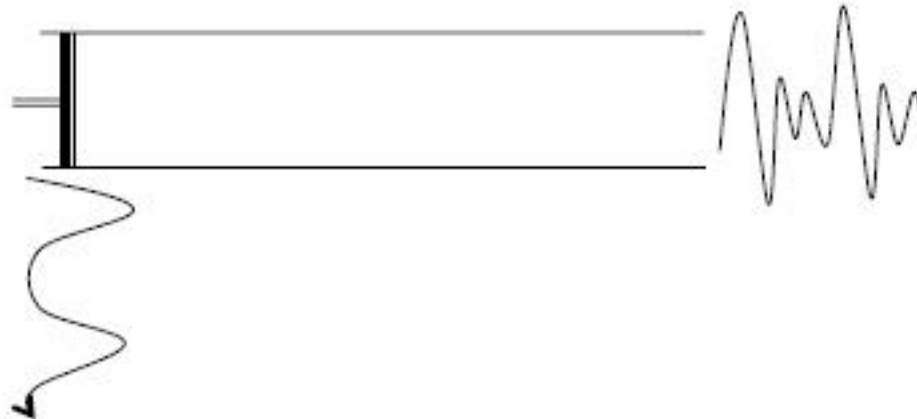
Image by MIT OCW.

Adapted from Ladefoged, P. *Elements of Acoustic Phonetics*, 2nd ed. University of Chicago Press, 1996.

Resonators are filters

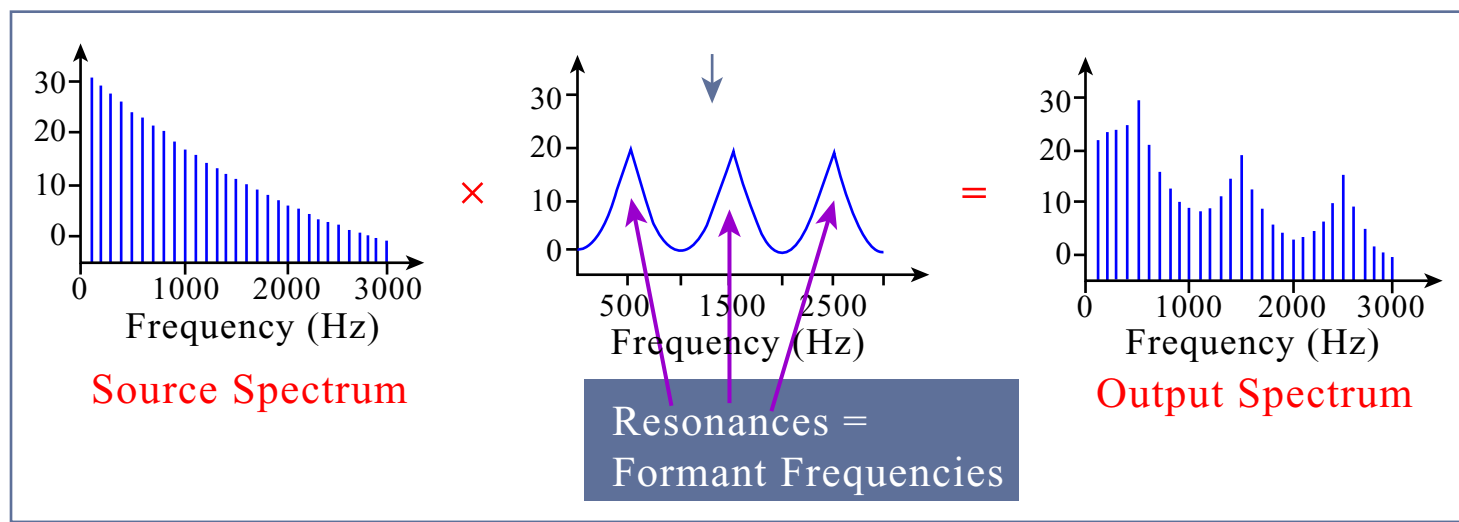
A resonator acts as an acoustic filter:

- Consider the vibration of a column of air driven by a vibrating source.
- The source is the input to the filter. This input produces vibration in the column of air.
- This vibration produces a sound wave in the external air - the output of the filter.
- The strength with which an input frequency is output from the filter will depend on the resonance characteristics of the air column - if it is near a resonant frequency, it will be passed through to the output at full strength, if it is not near a resonant frequency, it will have low amplitude in the output.



The vocal tract is a filter

- In speech the input has energy at a range of frequencies (voicing or noise, or both) and the resonator has multiple resonant frequencies of varying bandwidths.
- The amplitude in the output wave of each frequency component of the source depends on the extent to which the air in the vocal tract resonates at that frequency. Thus the vocal tract filters the source.
- The resonances of the vocal tract are called formants.



Resonant frequencies of air in a tube

What determines the resonant frequencies of a mass of air in a tube?

- The resonant frequencies (formants) of air in the vocal tract depend on how the cross-sectional area of the vocal tract varies over its length.
- The relationship between vocal tract shape and transfer function is complex - we will start by considering the simple case of a uniform tube.

How air in a uniform tube vibrates

- The vocal tract in a vowel can be approximated by a tube which is closed at one end (the glottis) and open at the other (the lips).
- A compression or rarefaction travelling down a tube is reflected off the ends of the tube.
- There is reflection even at the open end of a tube, but with a change of phase - a compression is reflected as a rarefaction, and vice versa.

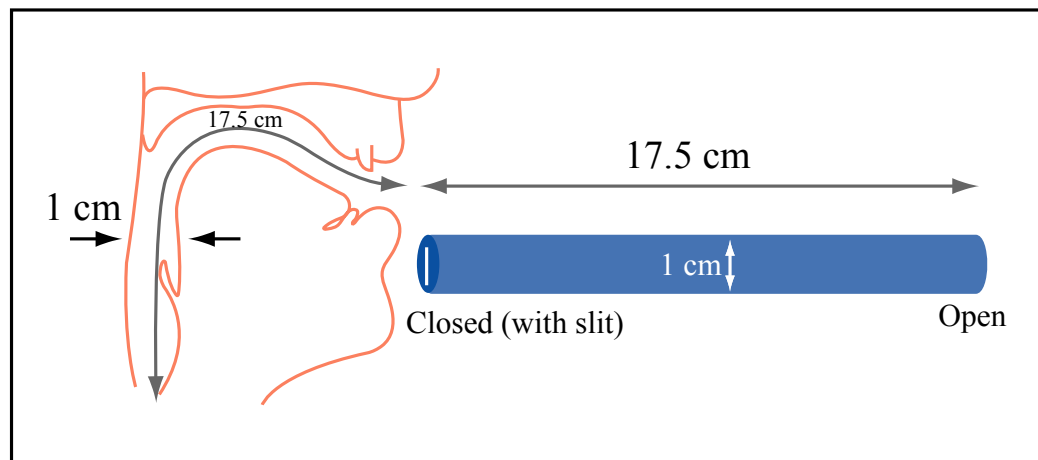


Image by MIT OCW.

Adapted from Ladefoged, P. *Elements of Acoustic Phonetics*. 2nd ed. Chicago, IL: University of Chicago Press, 1996.

How air in a tube vibrates

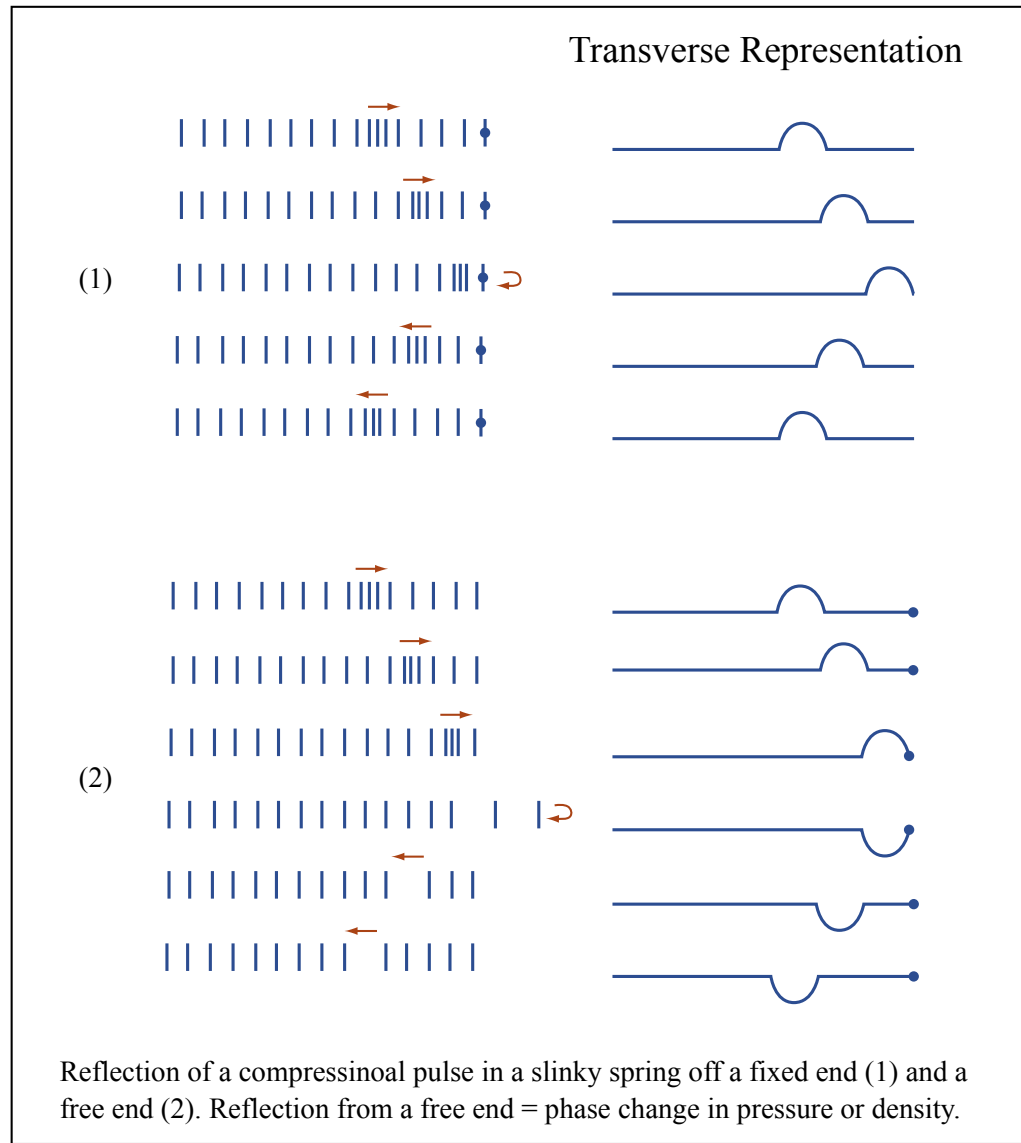


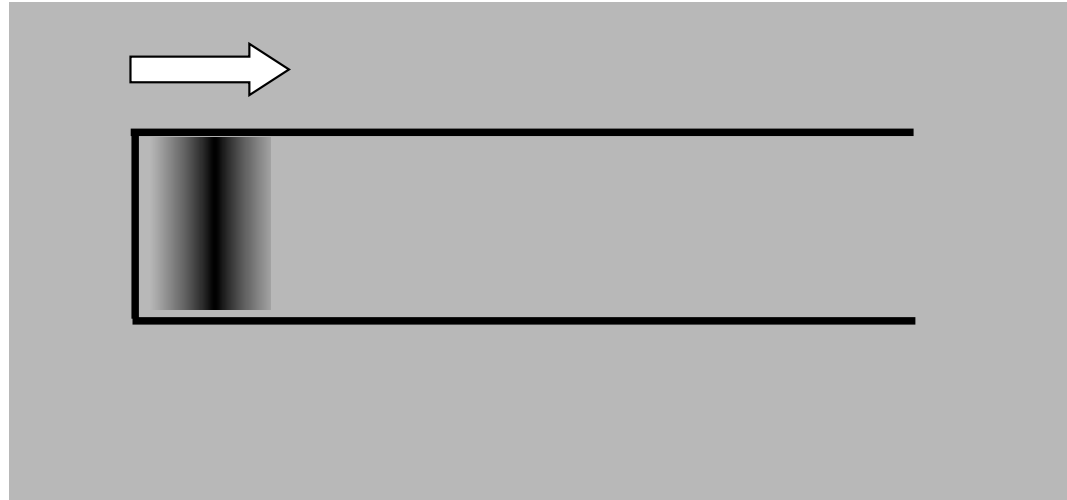
Image by MIT OCW.

Adapted from Berg, Richard E., and David G. Stork. *The Physics of Sound*. New York, NY: Prentice-Hall, 1982.

Resonance in a uniform tube, open at one end

- For resonance to occur, compressions must reach the closed end at the same time as the excitation source produces another compression peak, and rarefactions must reach the closed end at the same time as the source produces a rarefaction.
- In a tube which is open at one end:
 - i. a compression is reflected from the open end as a rarefaction.
 - ii. This rarefaction is reflected back off the closed end.
 - iii. The rarefaction is then reflected from the open end as a compression.
 - rarefactions do the opposite.
- So source and reflected wave will interfere constructively if, e.g., the time it takes for the wave to travel 2 times up and down the tube is equal to the half the period of the wave.

Resonance in a tube open at one end



Resonance in a tube open at one end

- Source and reflected wave will interfere constructively if, e.g., the time it takes for the wave to travel 2 times up and down the tube is equal to the half the period of the wave.
- For a tube of length L , for a wave with period T , travelling with speed c ,

$$\frac{T}{2} = \frac{2L}{c}$$

- $f = 1/T$ so this resonant frequency $F = c/4L$
- 17.5cm is a reasonable figure for length of a (large) vocal tract, speed of sound in air is approximately 35000cm/s.

Resonance in a tube open at one end

- This is only one resonant frequency (the lowest, or first, resonance)
- The general requirement is that compressions emitted from the source must coincide with reflected compressions (likewise for rarefactions) - they don't have to coincide with every reflected compression.
- So the time it takes the wave to travel $2L$ must be equal to $1/2, 3/2, 5/2 \dots$ i.e. $(2n-1)/2$ periods of the source wave.

$$\frac{(2n - 1)T}{2} = \frac{2L}{c}$$

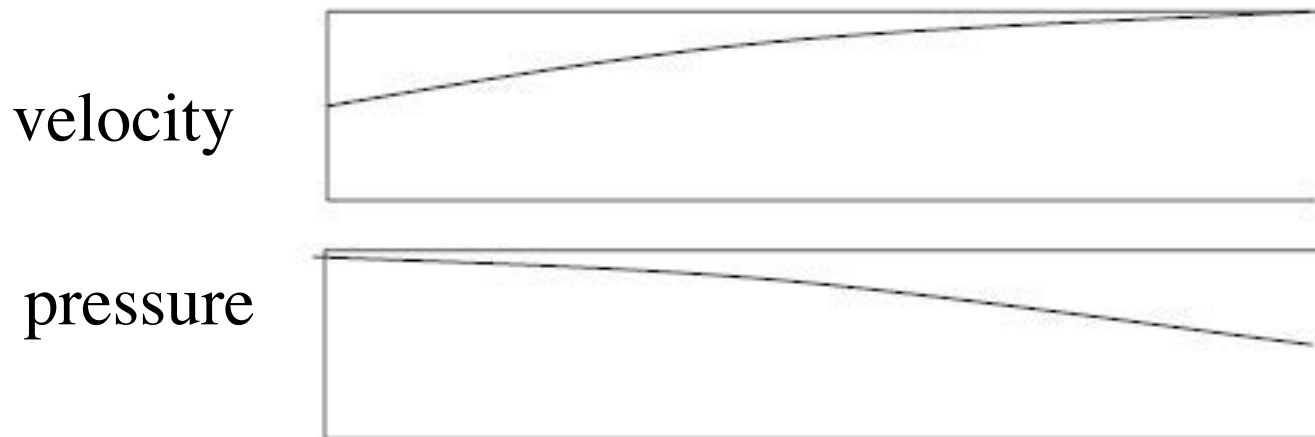
$$T = \frac{4L}{(2n - 1)c} \quad \text{so:} \quad F_n = \frac{(2n - 1)c}{4L}$$

Standing waves

- Under the conditions for resonance, the air in a tube vibrates in a standing wave.
- Standing waves arise when identical waves are travelling in opposite directions - e.g. a wave and its reflection from a surface. standing wave demo
- A standing wave contains fixed points at which air particles do not move (velocity nodes). Movement is maximal between the nodes (at the velocity anti-nodes).
- Velocity anti-nodes correspond to pressure nodes - pressure does not change - and velocity nodes correspond to pressure anti-nodes.

Standing waves and resonance in tubes

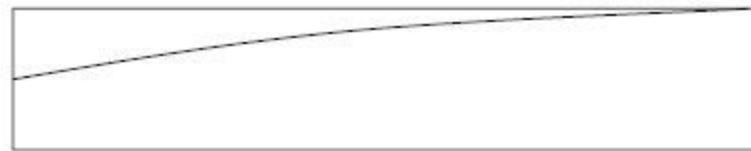
- The air in a tube which is open at one end naturally vibrates in standing waves.
- There must be a velocity node at the closed end since the fixed end prevents particle movement.
- There must be a pressure node at the open end - pressure is always equal to external pressure.
- The simplest wave that meets these conditions is a standing wave that goes from maximum pressure at the glottis to zero (atmospheric pressure) at the lips.



Standing waves and resonance in tubes

What is the frequency of this wave?

velocity



pressure



Relating frequency to wavelength

- The distribution of pressure fluctuations in space due to a sound wave depends on the frequency of the wave and the speed at which the wavefront travels.
- Consider a wave with frequency f Hz and speed c cm/s.
- In one second it travels c cm.
- In one second it goes through f cycles.
- So there are f cycles in c cm.
- So the length of one cycle (the wavelength) is c/f .

- Usual formulation: $c = f\lambda$ Where λ is wavelength

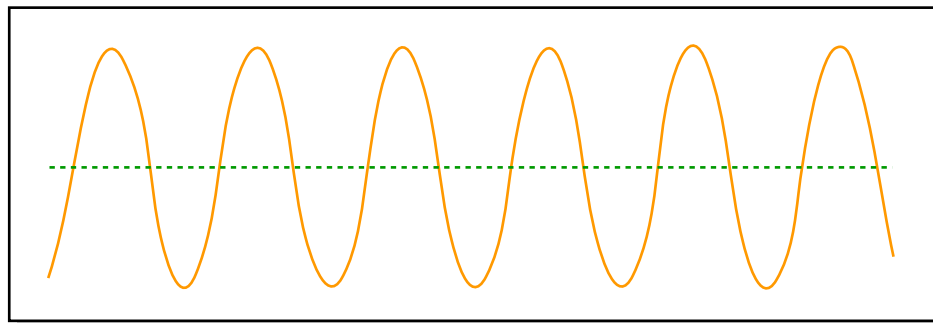


Image by MIT OCW.

Standing waves and resonance in tubes

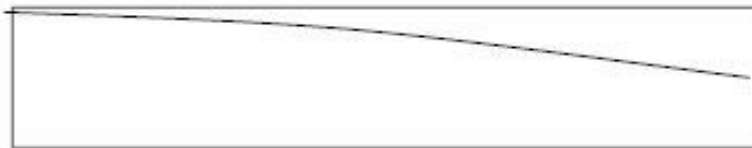
What is the frequency of this wave?

- The transition from maximum to zero pressure is only one quarter of a cycle.
- For a tube of length L , $L = \lambda/4$
- $\lambda = 4L$
- $f = c/\lambda = c/4L$

velocity



pressure



Standing waves and resonance in tubes

- Other standing waves that meet the boundary conditions:

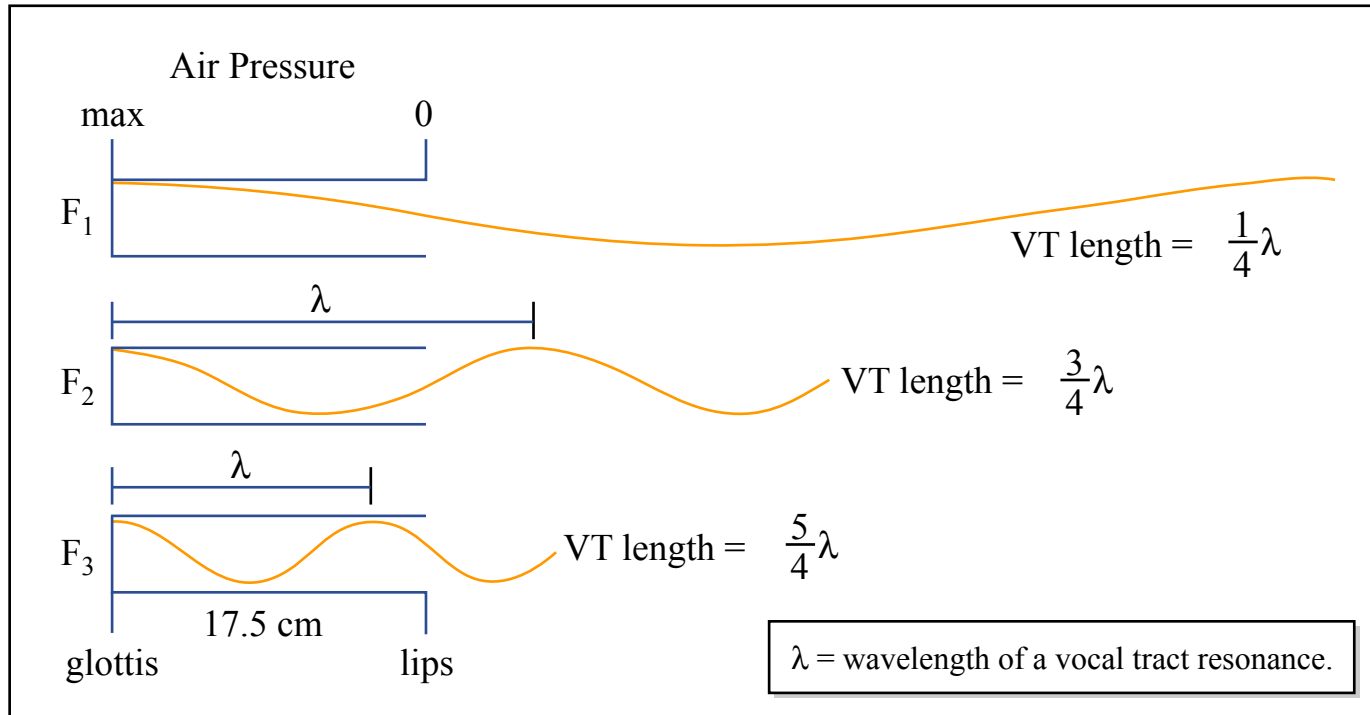


Image by MIT OCW.

Adapted from Johnson, Keith. *Acoustic and Auditory Phonetics*. Malden, MA: Blackwell Publishers, 1997.

- In general, air in a tube of length L , closed at one end, will have resonant frequencies (formants) at:
$$F_n = \frac{(2n-1)c}{4L}$$

Tubes models of the vocal tract

- The vocal tract in vowels can be approximated by a tube that is closed at one end and open at the other.

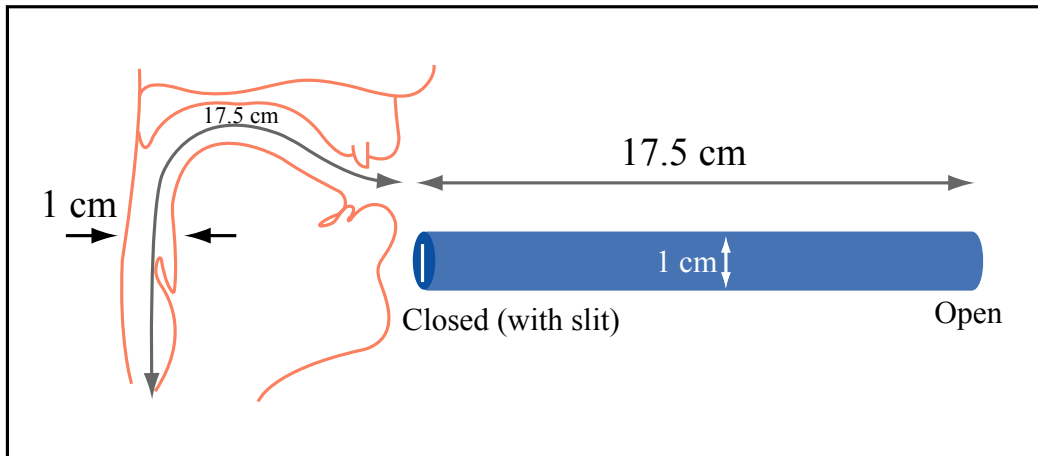


Image by MIT OCW.

Adapted from Ladefoged, P. *Elements of Acoustic Phonetics*. 2nd ed. Chicago, IL: University of Chicago Press, 1996.

$$F_n = \frac{(2n - 1)c}{4L}$$

- The cross-sectional area of the vocal tract is closest to being uniform in a mid central vowel [ə] as in RP English *bird* [bɜːd], or at the end of *sofa* [sɒfə].
- Caution: Most things that are transcribed as schwa do not have a uniform vocal tract.

Tubes models of the vocal tract

- Another simple tube model that approximates a vocal tract shape is a tube (almost) closed at both ends.
- This is similar to a sound made with a narrow constriction at the lips, or the cavity behind an interior constriction.
- Air in a tube closed at both ends naturally vibrates in standing waves with nodes at both ends.

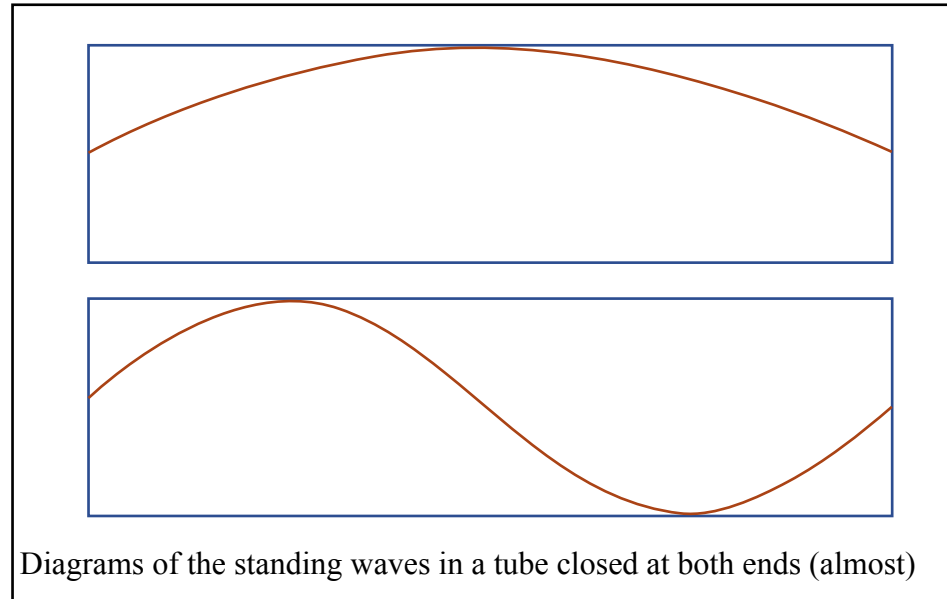


Image by MIT OCW.

Adapted from Ladefoged, Peter. L104/L204 Phonetic Theory course notes, University of California, Los Angeles.

Tubes models of the vocal tract

- $L = n\lambda/2$
- $\lambda = 2L/n$
- $F_n = c/\lambda$

$$F_n = \frac{nc}{2L}$$

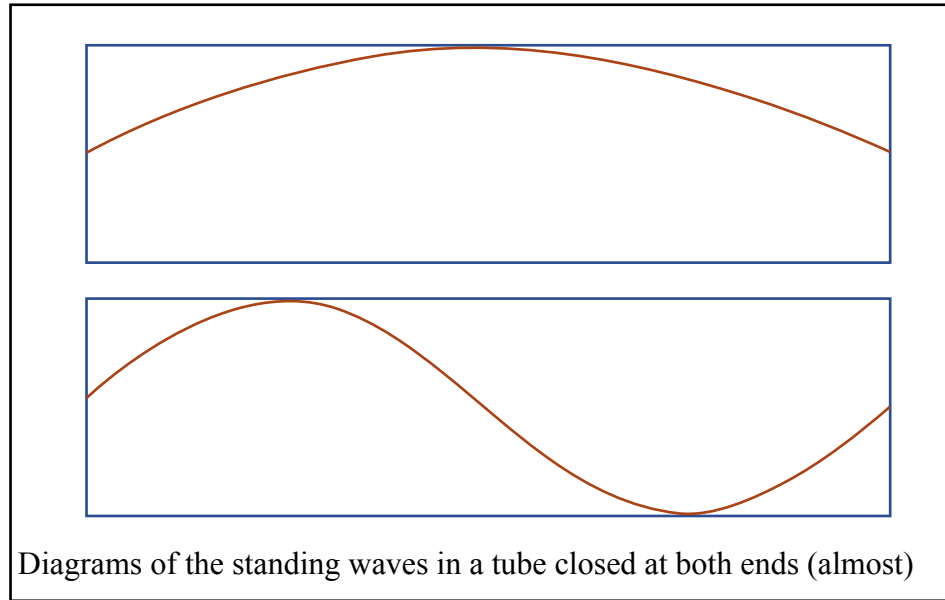


Image by MIT OCW.

Adapted from Ladefoged, Peter. L104/L204 Phonetic Theory course notes, University of California, Los Angeles.

Source-Filter Model of Speech Production

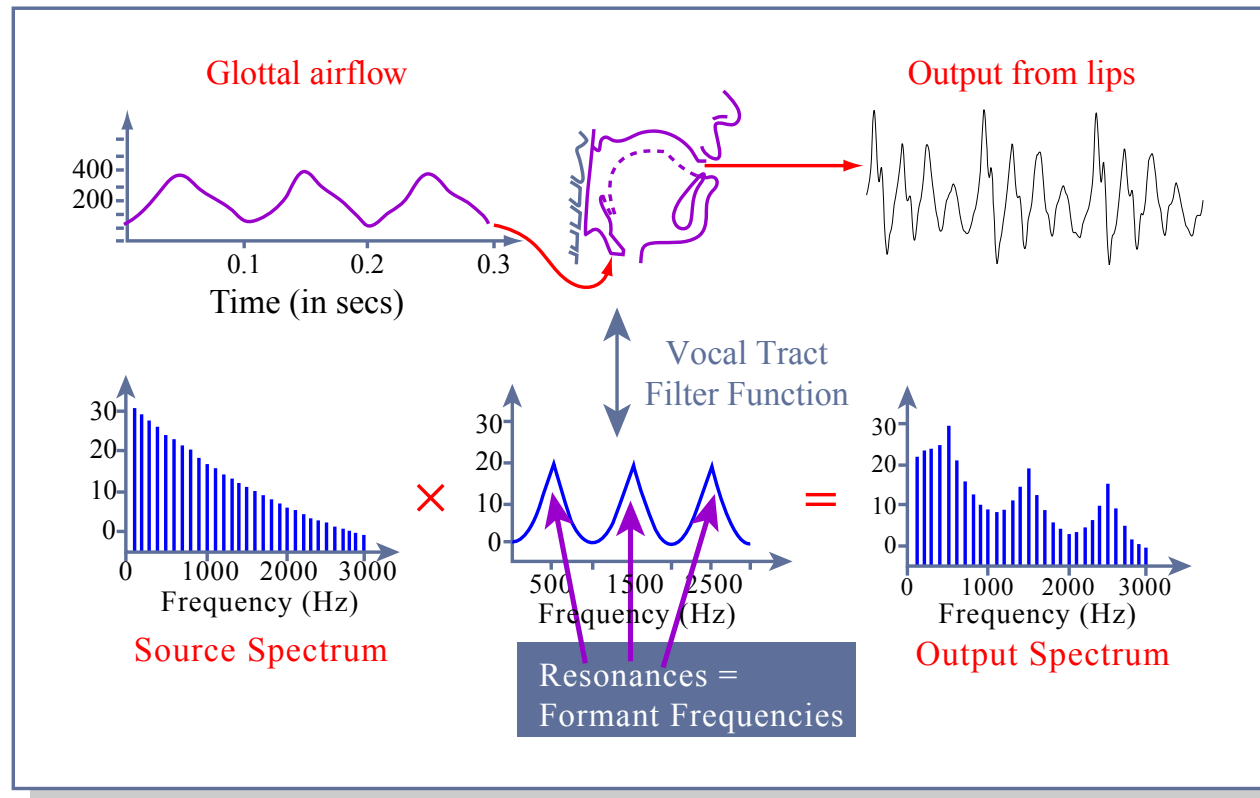


Image by MIT OCW.

MIT OpenCourseWare
<https://ocw.mit.edu>

24.915 / 24.963 Quantum Optical Communication
Fall 2015

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.