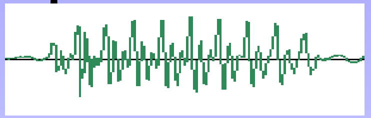
# Crash course in acoustics

# Outline

- What is sound and how does it travel?
- Physical characteristics of simple sound waves
  - Amplitude, period, frequency
- Complex periodic and aperiodic waves
- Source-filter theory

# Take home message

- Two types of visual representation of speech:
  - Waveforms (time by amplitude displays)
  - Spectrograms (time by frequency displays)



- Characteristics of waveforms
  - simple, complex, periodicity, frequency…
- Speech sounds
  - The source
    - vocal folds: produces a range of frequencies (harmonics)
    - Other sources, e.g. noise
  - Articulators filter this source

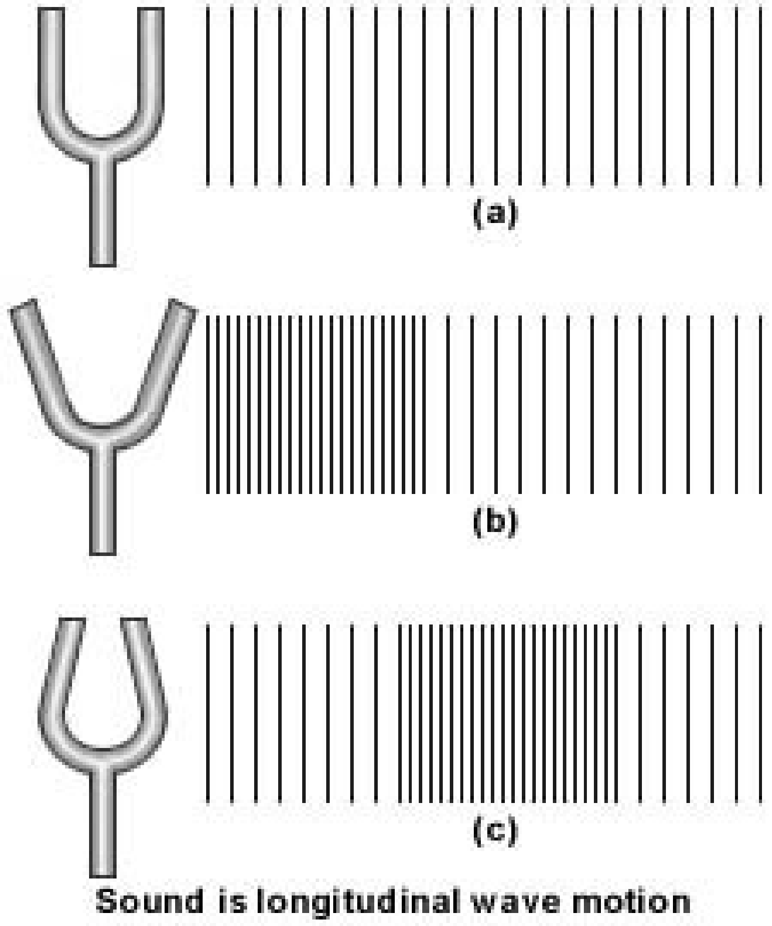
# What is speech acoustically?

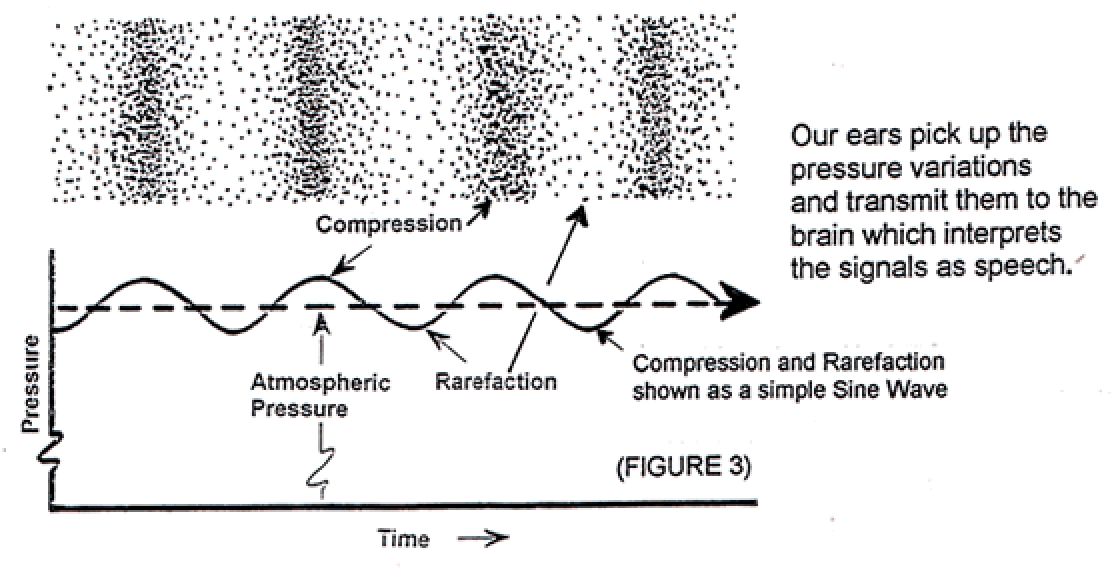
- Acoustically, speech sounds are similar to other sounds.
- What is sound?
  - Movement of particles which causes pressure fluctuation in the surrounding air.
- What happens to the air pressure?
  - Compression: Pressure increases so the air particles are close together
  - Rarefaction: Pressure decreases so the air particles are further apart
- Air pressure changes cause our eardrum and bones of the middle ear to vibrate.

# How does sound travel?

- In order to even reach our ears, sound must travel from its source.
- Pressure fluctuations caused by the source propagate through elastic media:
  - there is room for molecules to move back and forth.
- In other words, sound moves in waves
  - A sound wave is a traveling pressure fluctuation

# Propagation of a sound wave

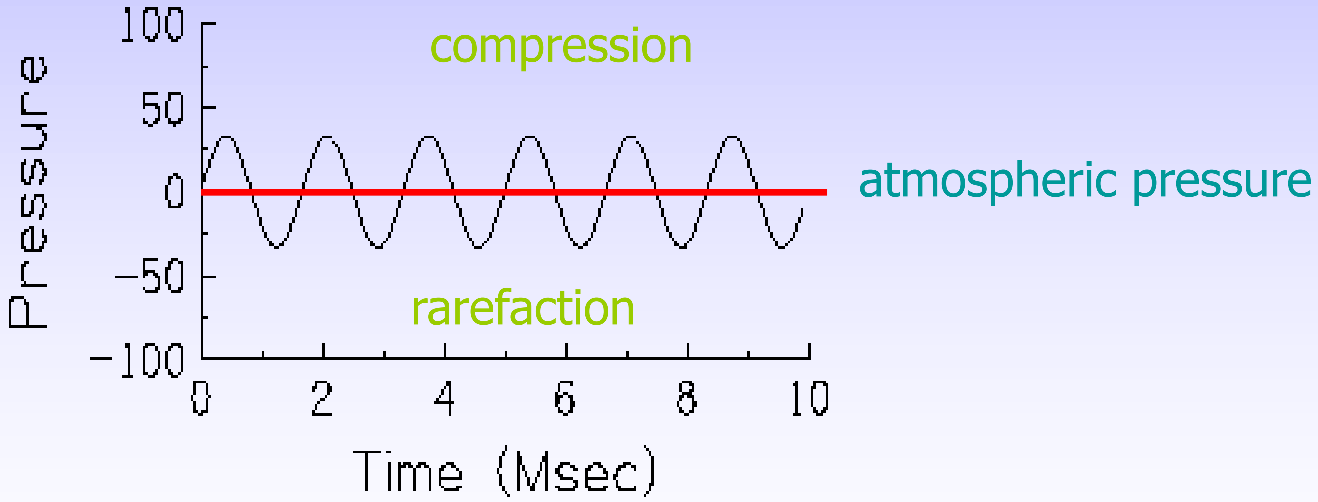



From http://home.att.net/~cat4a/wave-II.htm
See also http://entertainment.howstuffworks.com/sci-fi10.htm

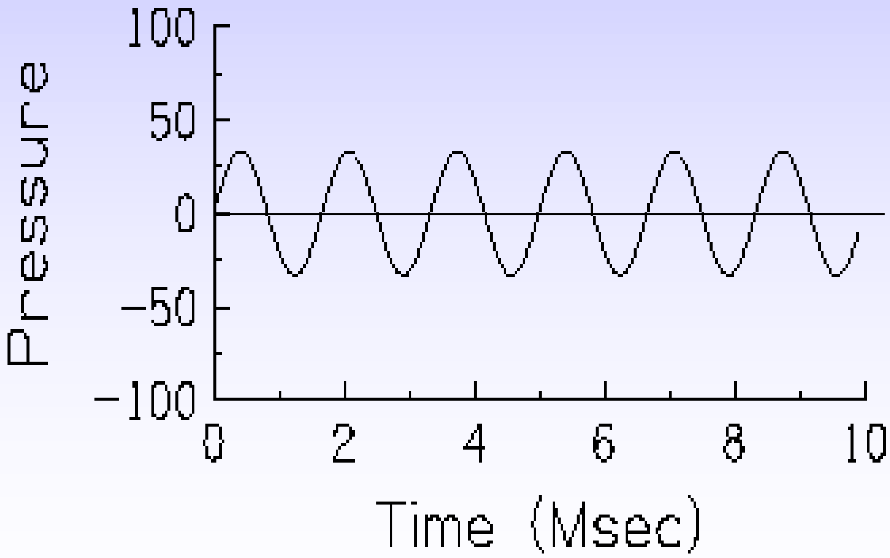
# The waveform as pressure changes

- Graphic representation showing changes in air pressure:



# The simplest sound

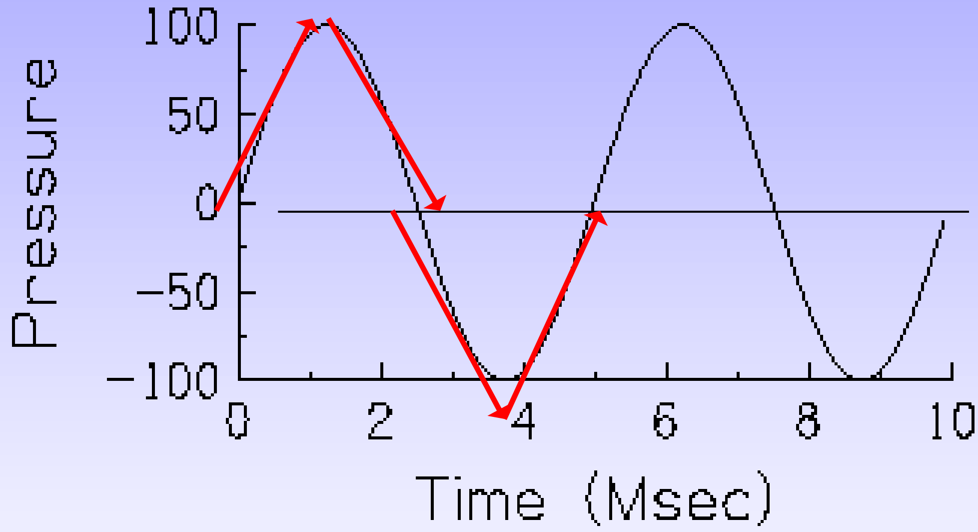
- The simplest type of sound is a pure tone (e.g. the sound of a tuning fork)
  - Generated by simple harmonic motion (think pendulum)

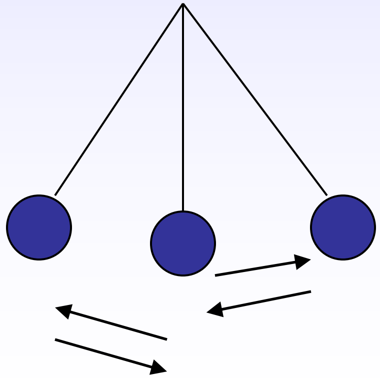


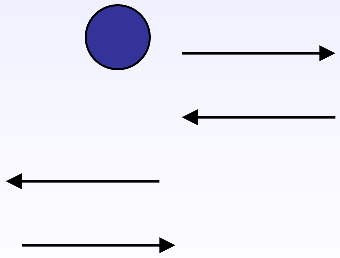

Waveform of a pure tone = a sine wave

# What does this graph represent?

- We can interpret this graph as showing the movement of one air particle from its rest position

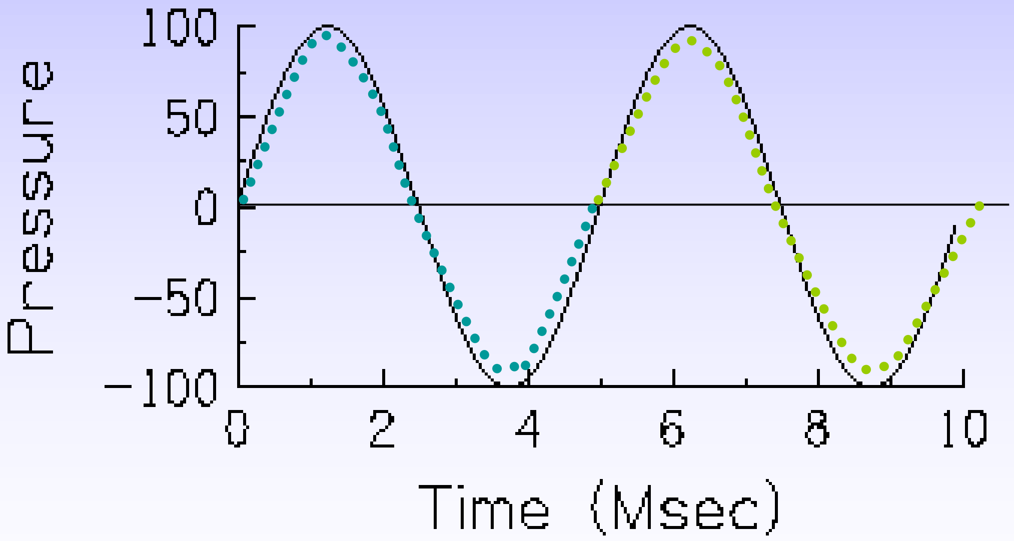




# What is a periodic wave?

- This type of wave is called periodic because the motion of the particle is repeated at regular intervals



# Frequency of a periodic wave

- Cycle: The completed movement of a wave
- Period (T): The time it takes for each cycle in a wave to be completed
- Frequency: The number of cycles completed in one second
- Hertz (or cycles per second): Frequency is measured in Hz and is perceived as *pitch*
- Humans can hear sounds whose frequency ranges from 20-20,000 Hz, though speech sounds essentially utilize the lowest half of this range

# Frequency in more detail

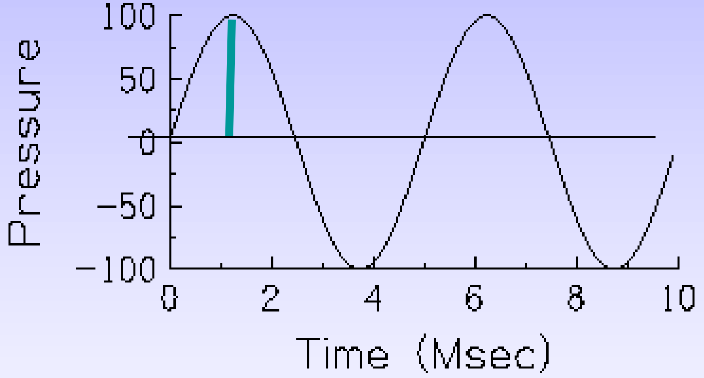
- Frequency: the number of times a cycle is repeated within a time interval (in speech, the standard unit of time is 1 second)

$$\text{Frequency} = 1/T \text{ (period)}$$

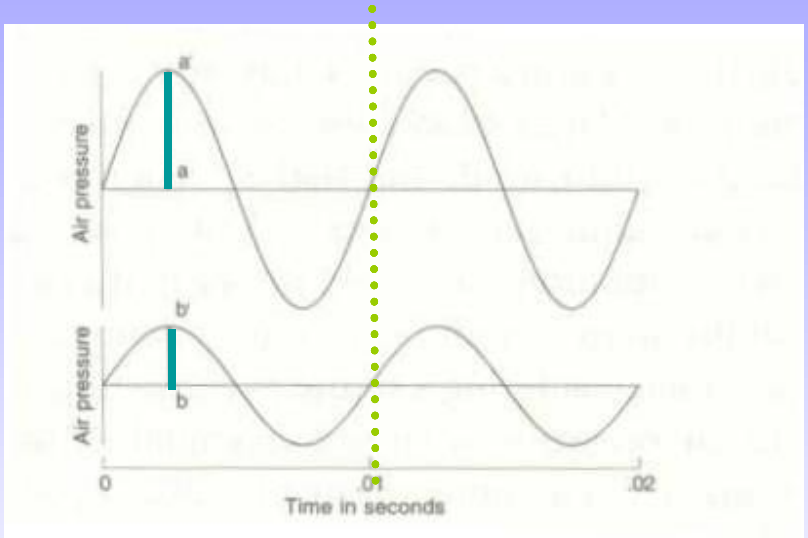
- If one cycle of vocal fold vibration is completed every 1/100 of a second, then T= 0.01 and frequency is 1/0.01 = 100 cycles per second = 100 Hz (Hertz)
- Conversely, if the frequency of a sound is 200 Hz, then there are 200 cycles of the sound wave per second, and the period is 1/200 = 0.005 s = 5 ms (milliseconds)

# Amplitude of a periodic wave

- The amplitude of the wave is related to the extent of particle displacement
- Roughly speaking, the greater the displacement (amplitude), the louder the sound
- NB: Amplitude and frequency are independent of one another

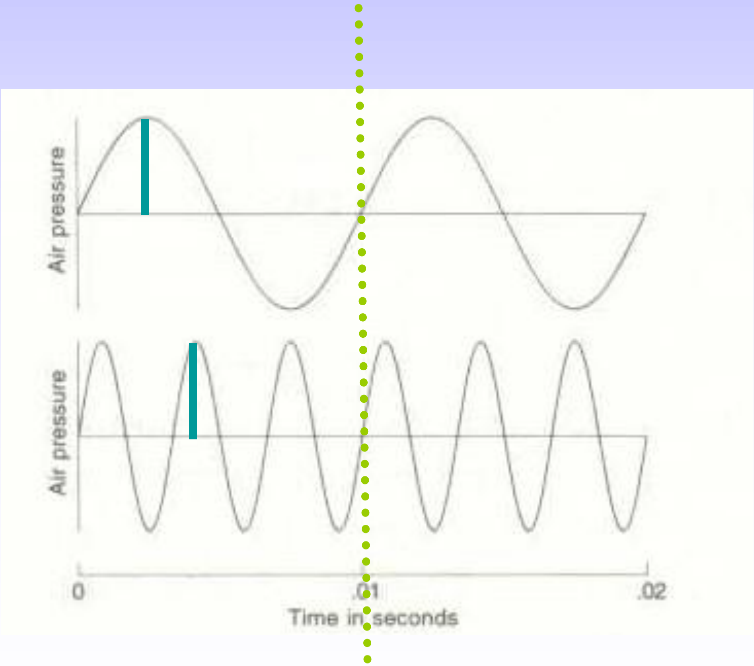


# The independence of frequency and amplitude




Two sine waves with the same frequency but different amplitude

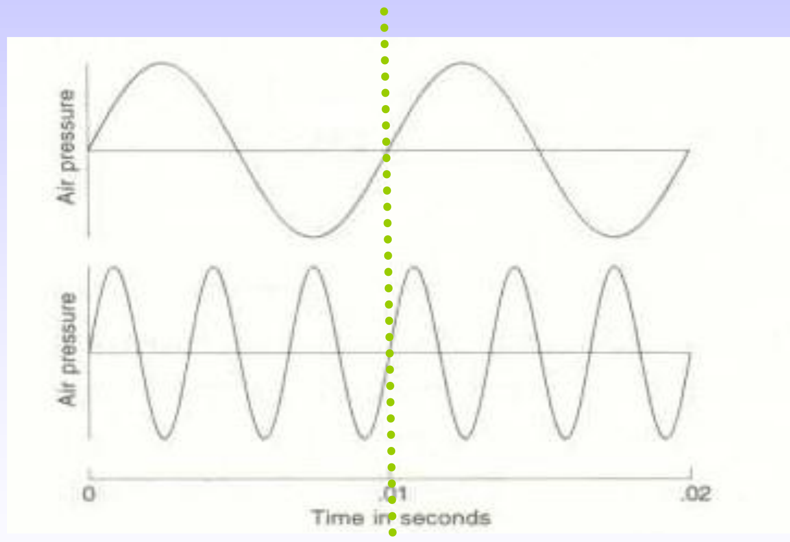
Two sine waves with the same amplitude but different frequency




From Ladefoged, 1996

# Question

- What are the frequencies of these two sine waves? (Hint: Calculate how many cycles must be completed in one second.)




0 .01 .02

Time in seconds

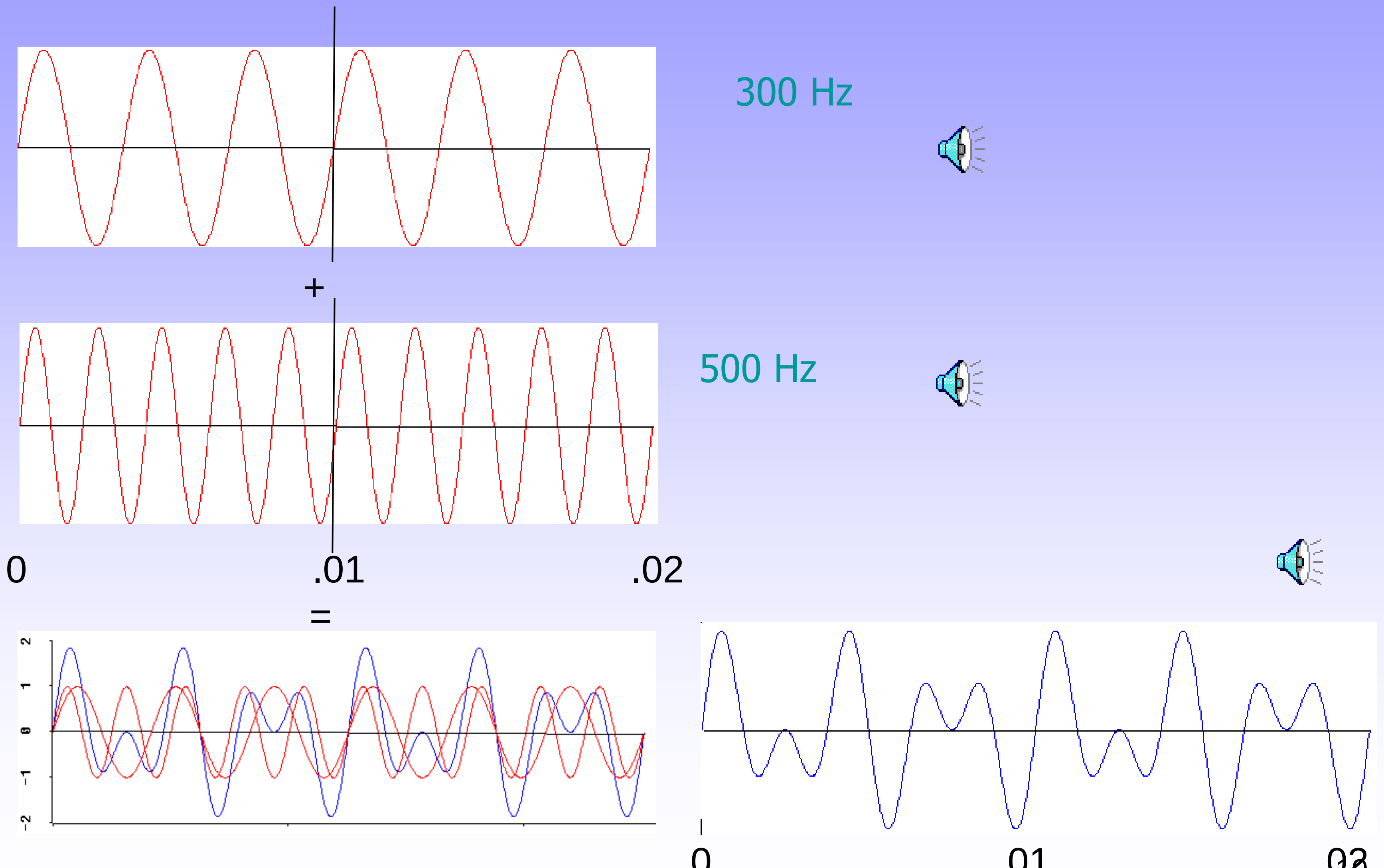
# Beyond tuning forks

- If only speech were that simple. But in fact:
- Speech sounds are not pure tones
- Rather, they are made up of tones of several frequencies, and can be distinguished into two categories
  - (complex) periodic sounds
  - aperiodic sounds

# Complex periodic sounds

- Complex periodic sounds are made up of two or more sine waves.
- Complex waves still have a repeating pattern.
- The component sine wave with the lowest frequency is called the fundamental frequency (F0).

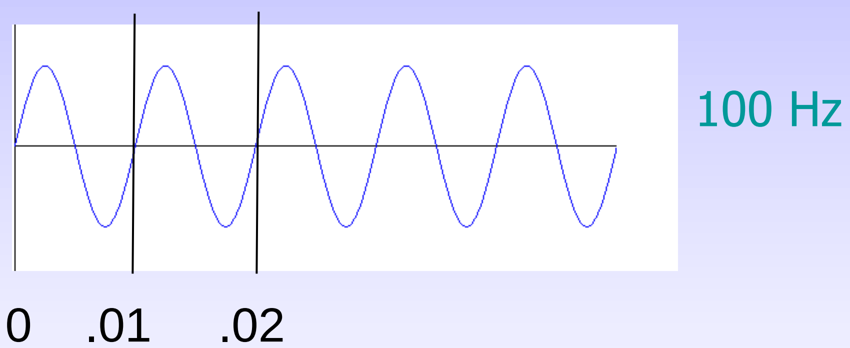
# A complex periodic wave

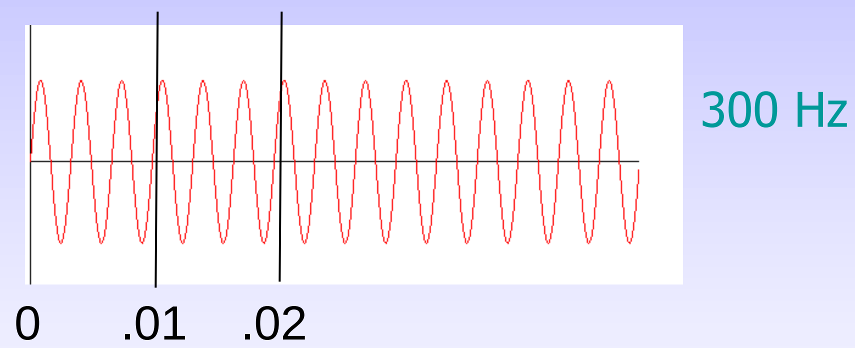


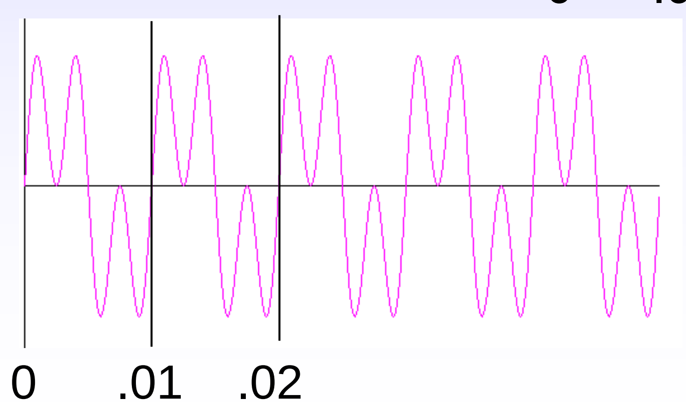

Source: http://www.umanitoba.ca/faculties/arts/linguistics/russell/138/sec4/acoust1.htm

# Fundamental Frequency

- The waveform of a complex periodic sound has a complex shape that is repeated with the same periodicity as its lowest frequency component.

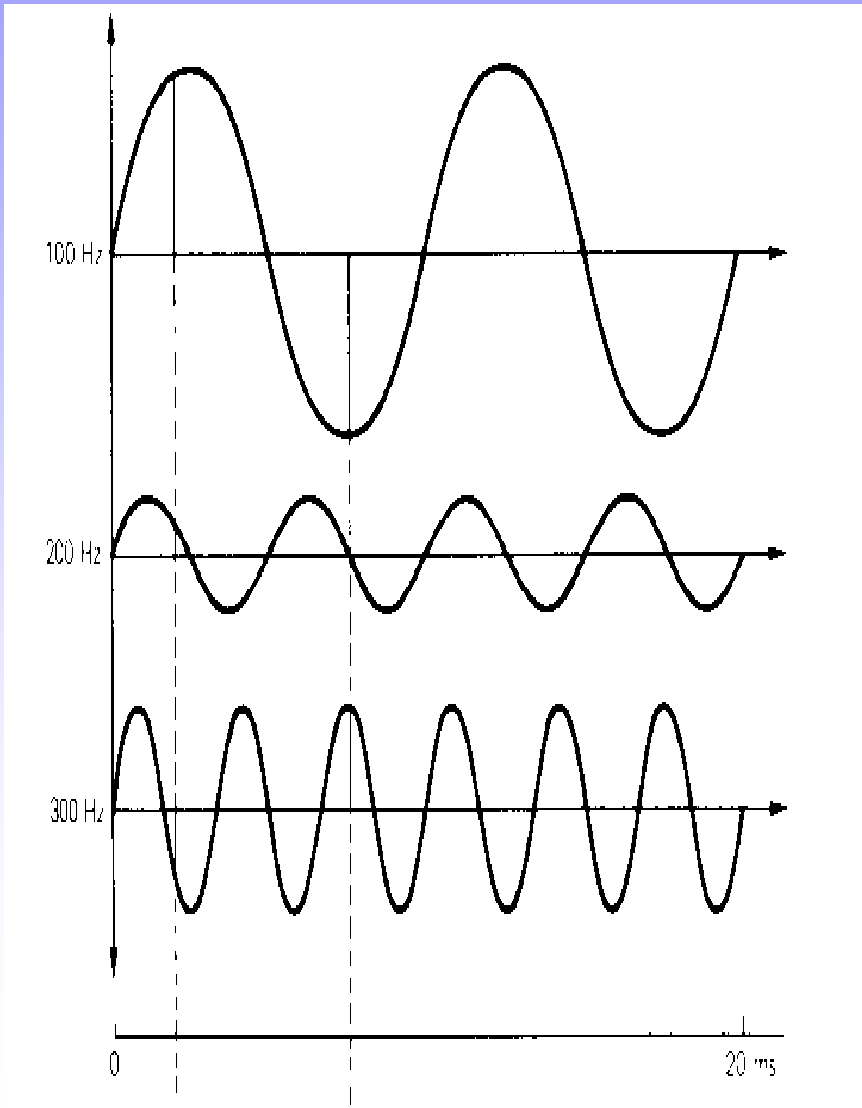



What is the fundamental frequency?

# Waves of different frequencies (and amplitudes)



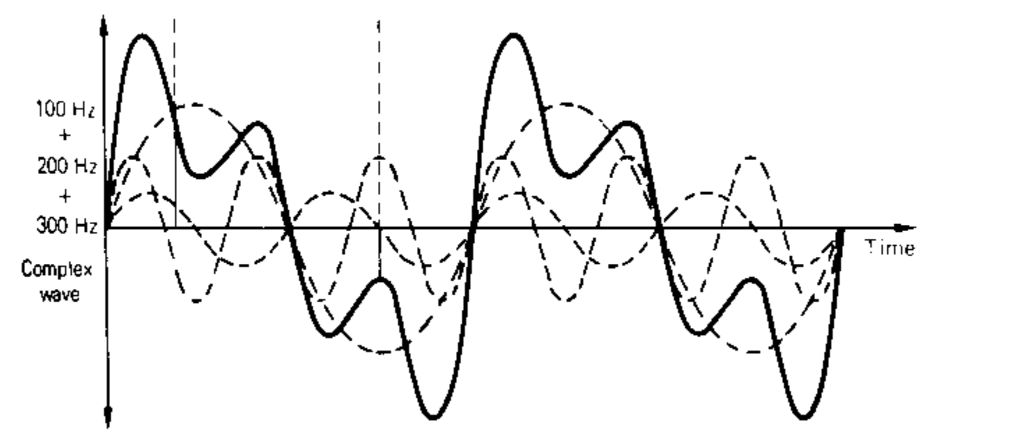




**FIGURE** 7.4.1 Complex wave with three sinusoidal components (100 Hz, 200 Hz, 300 Hz)
***Adapted from:*** Ladefoged 1962, p. 35.

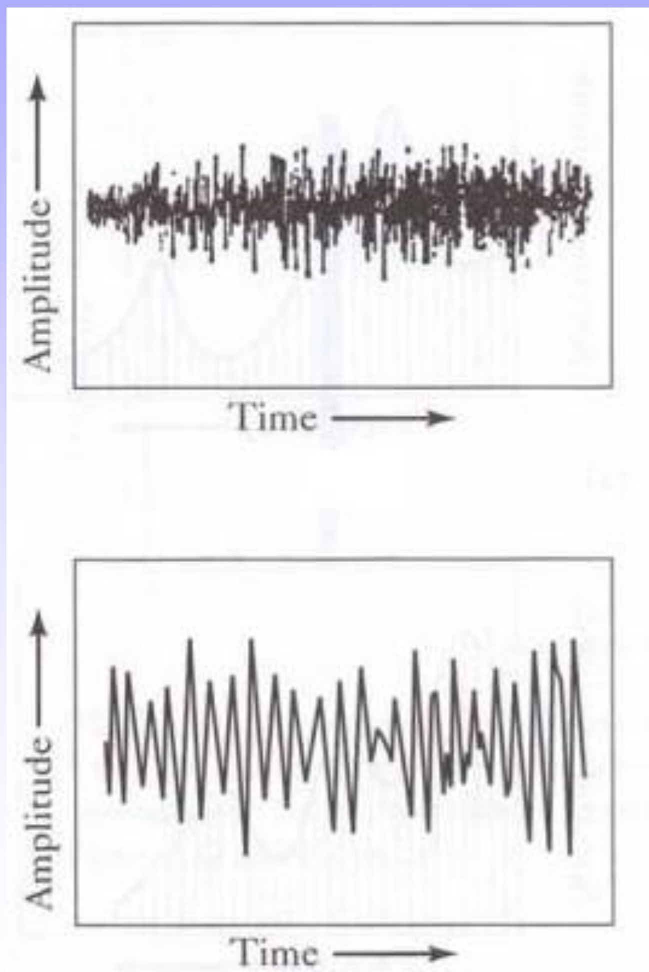
# Further aspects of complex waves

- The waveform of a complex periodic wave is derived by adding up the amplitude of the component waves at every point in time.
- When the complex wave is known but the components aren't, the wave can be broken down using a method called Fourier analysis.
- Vowels (and to an extent other vowel-like speech sounds, i.e. approximants) are complex periodic waves

# Aperiodic waves

- Aperiodic waves have components at several frequencies, not only at frequencies that are whole number multiples of the fundamental.
- The result is an irregular non-repetitive waveform
- An example is radio static or white noise, which include components of equal amplitude at all audible frequencies
- In speech, fricatives are aperiodic.

# Examples of aperiodic waves



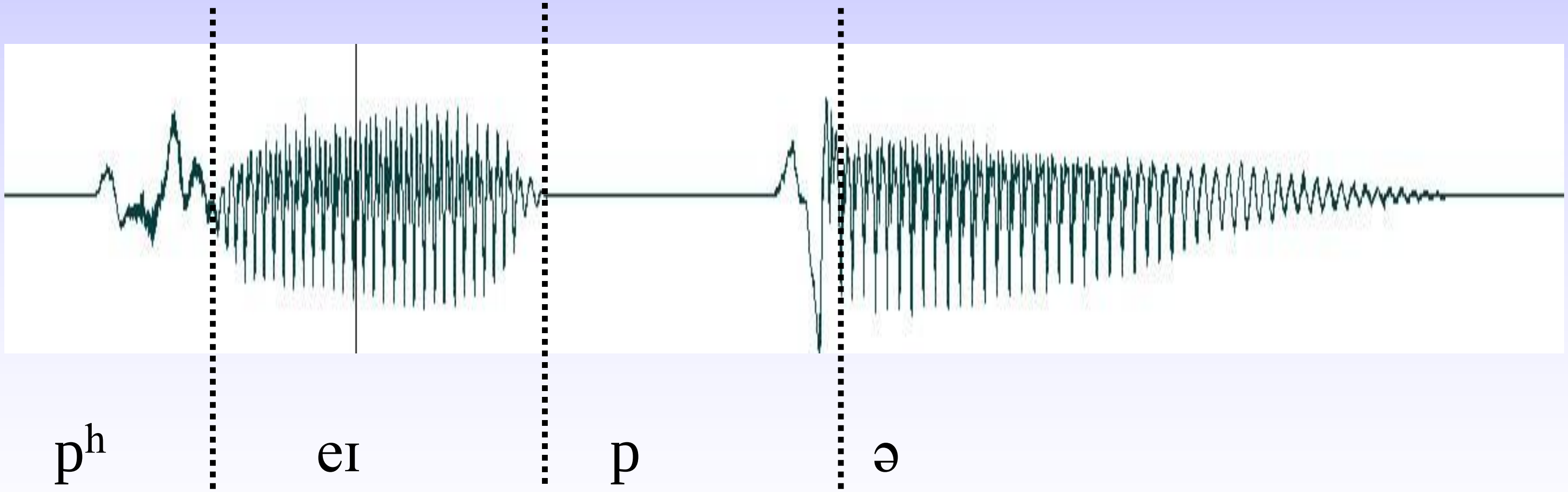

White noise

∫

From Denes & Pinson, 1973

# Visual waveforms of speech

- One way speech can be examined is in the time domain.
- This is plotted as a time by amplitude waveform.
- No frequency information is contained in a waveform.



# Limitations of waveforms

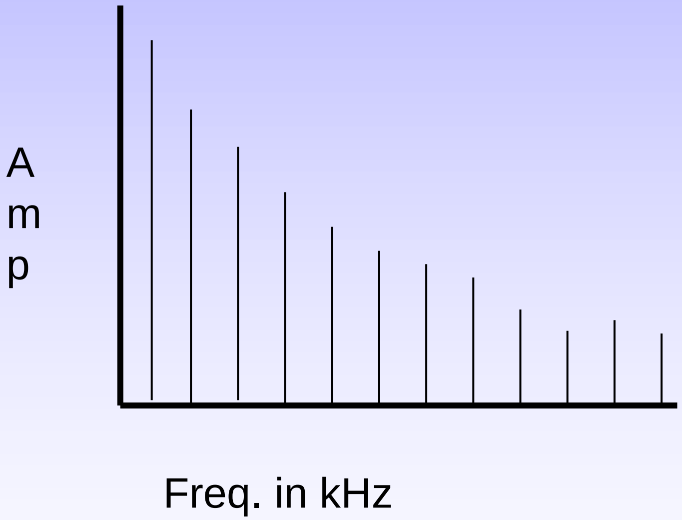
- By examining waveforms we may not be able to tell exactly which sound is shown: e.g. we can tell the sound is a stop, but we may not be able to tell whether it is  a [p] or a [t]
- The difficulties are similarly great for vowels
- Frequency information is needed, in addition to time and amplitude
- **Spectrograms** are a representation that provides us with time, frequency and amplitude information

# Harmonics (overtones)

- What is the difference between the tuning fork and the vocal folds in how they vibrate?
- The fork corresponds to (more or less) a simple sine wave whereas the vocal folds correspond to a complex wave.
- Components of complex waves are called harmonics.
- Harmonics are not arbitrary sine waves:
  - Rather, waves with frequencies that are whole number multiples of the fundamental frequency (F0, the lowest frequency component)
- The amplitude of each component wave is not necessarily the same.

# Spectrum

- Spectrum is a plot of frequency against amplitude at a particular point in time



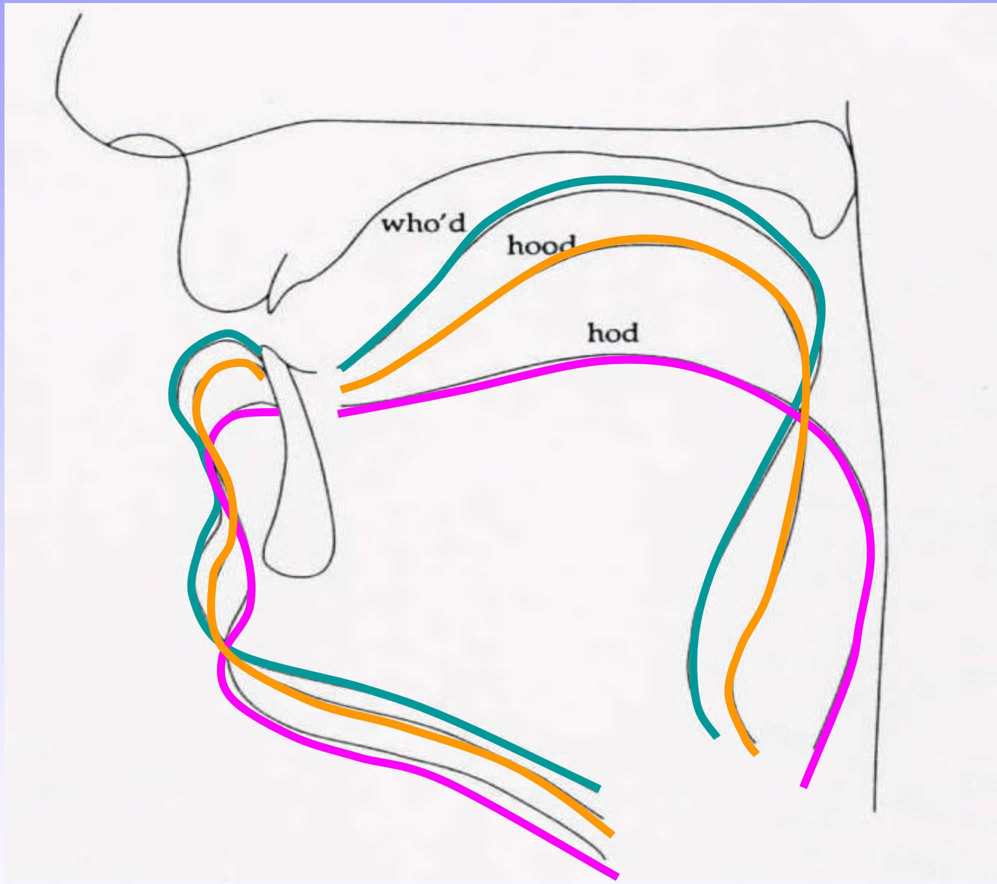

Source input (vocal folds)

# Source-Filter Theory

- In the speech wave, not all harmonics have equal amplitude.
- This is because the human vocal tract acts as a filter:
  - Frequency-selective
  - Amplifies some frequencies (called formants), damps others;
- Phonation (vocal fold vibration) is the sound source that is shaped by the vocal tract

# The vocal tract filter

- Different sounds have different vocal tract configurations.
- For example, consider the difference between [u], [ʊɑ], and [ɑ].

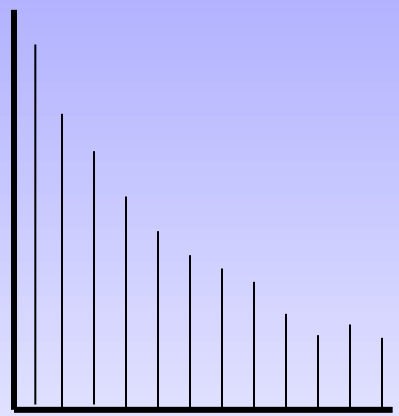



The sound source from the vocal folds always contains the same frequencies, but different vocal tract shapes will amplify some frequencies and damp others.

# Spectrum

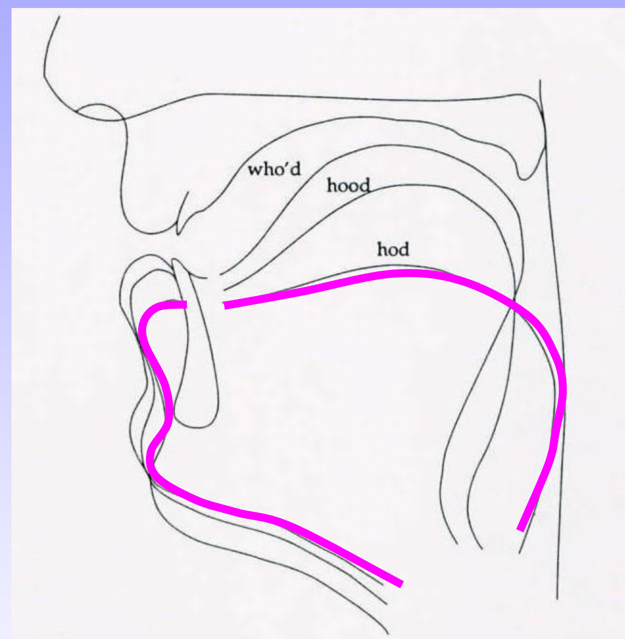
Frequency response for [ɑ]: amplified frequencies (formants) at 800Hz, 1300Hz, 2500Hz (approx)

A
m
p

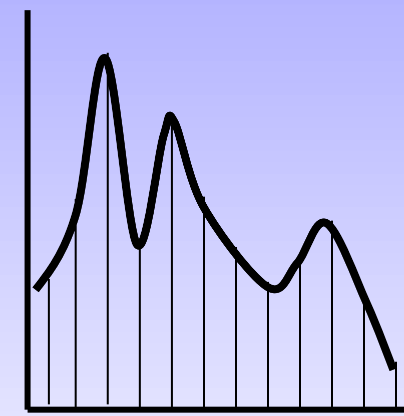


Freq. in kHz

→




→

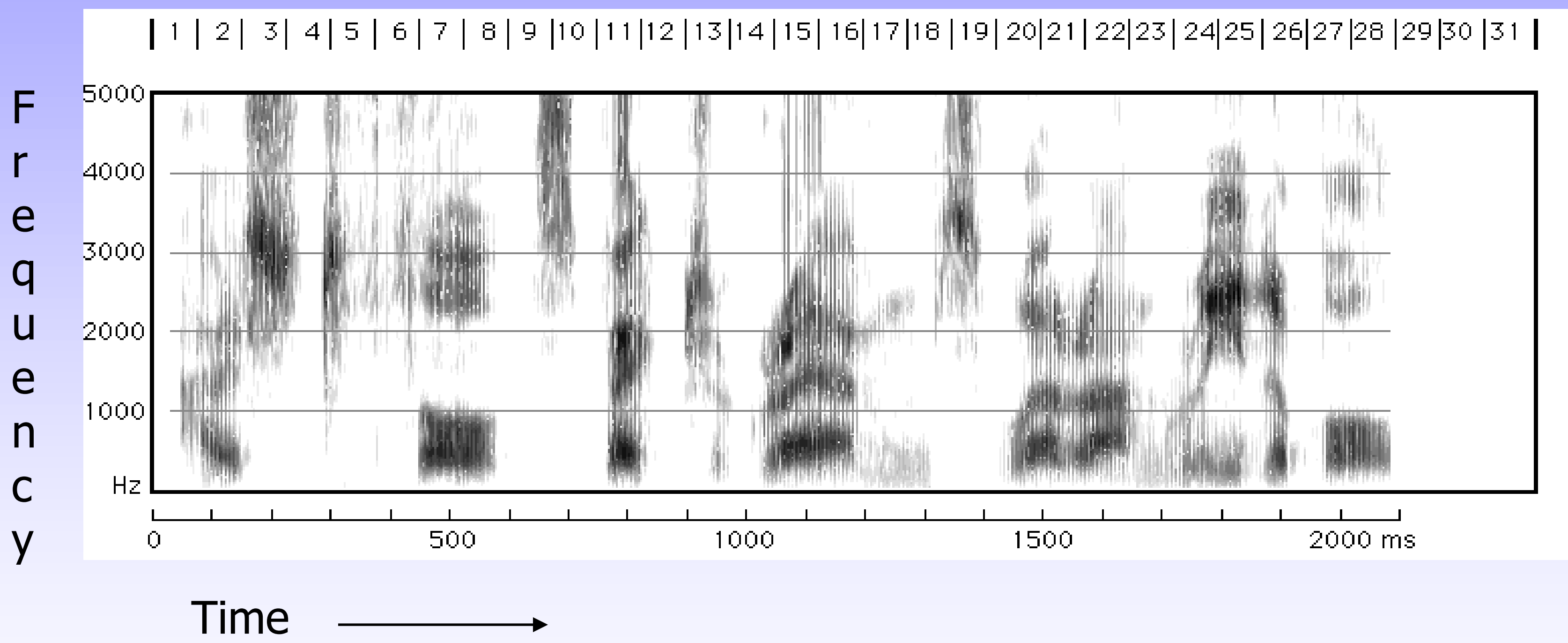


Source input
(vocal folds)

Resonator filter
(vocal tract)

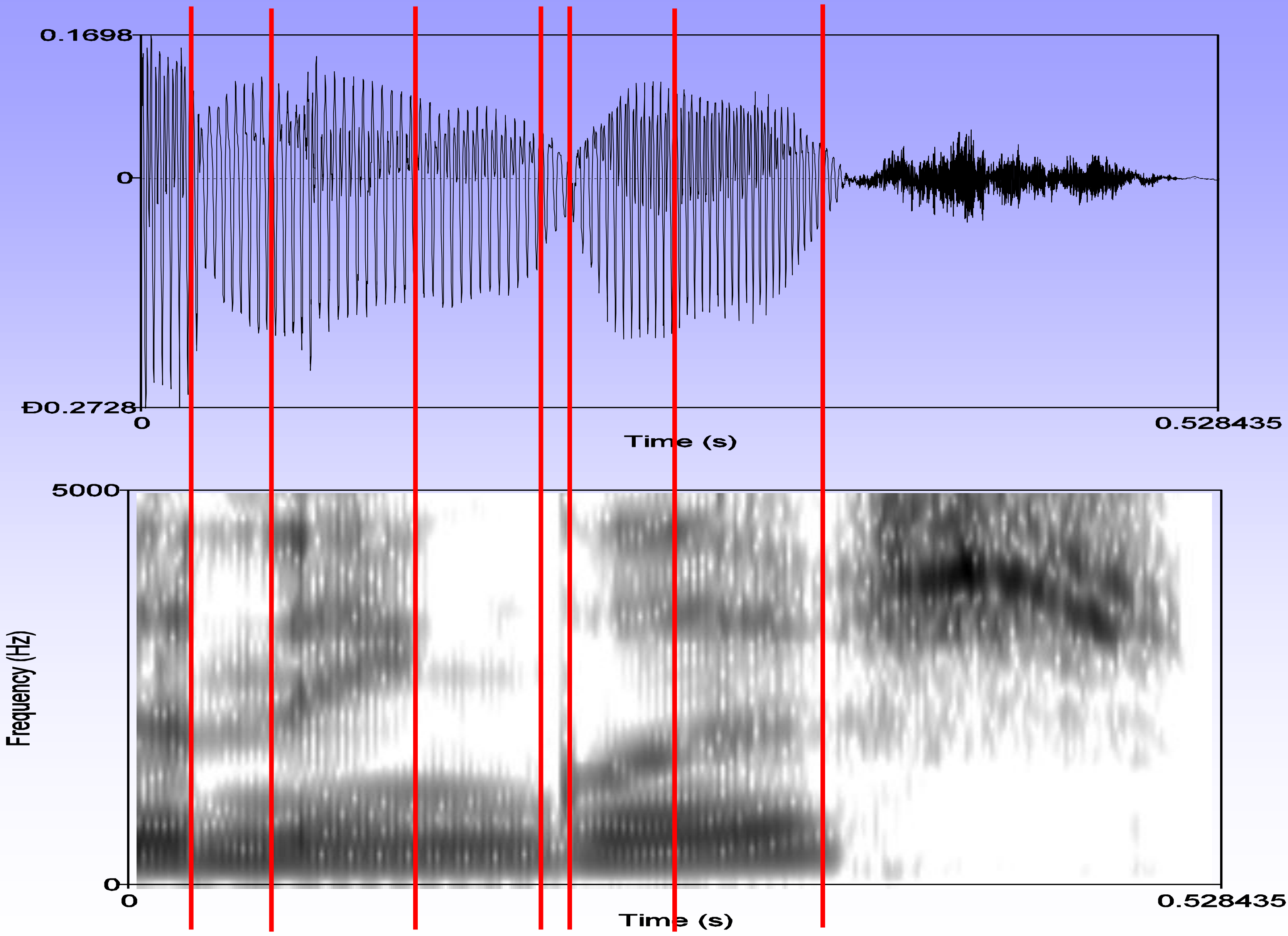
Output
(vowel)

# Spectrogram



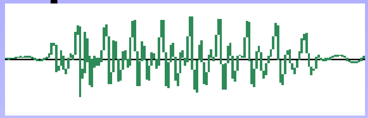

X-axis is time; Y-axis is frequency; darkness is intensity

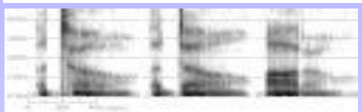
Try to find the individual phonemes of “In English” in the waveform:



# Summary

- Two types of visual representation of speech:
  - Waveforms (time by amplitude displays)
  - Spectrograms (time by frequency displays)





- Characteristics of waveforms
  - simple, complex, periodicity, frequency…
- Speech sounds
  - Articulators filter the source
  - Spectrum, harmonics