



## Speaker notes

- A common approach involves first finding and marking the beginning and end of realizations of particular phonemes or of major allophones of particular phonemes, for example, all tokens of /i/ or all tokens of /aɪ/ not adjacent to a nasal, lateral, rhotic, or labiovelar.

# FORMANT FREQUENCIES



## Speaker notes

Since questioned-speaker recordings are often telephone recordings and traditional telephone systems have bandpasses of around 300 Hz – 3.4 kHz, first formant (F1) frequencies close to 300 Hz (e.g., in [i] and [u]) are often distorted, and high frequency spectral information in bursts and fricatives (e.g., in [th] and [s]) is often missing.

Once they have made their measurements, some practitioners make tables or plots of the values, and use their training and experience to subjectively assess strength of evidence via examination of those tables or plots. For example, first versus second formant (F1-F2) values could be plotted for realizations of a particular vowel phoneme in both the known- and questioned-speaker recording, and the visual degree of overlap considered. Measurements made on recordings of foil speakers may also be plotted.

# TIMING CATEGORIES IN SPEECH ACOUSTIC

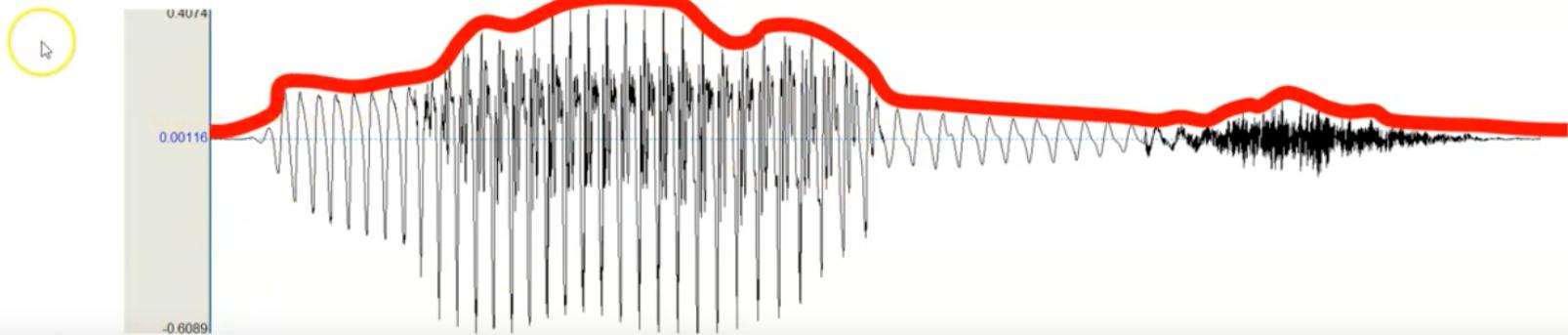
- Fastest changes: temporal fine structure
- Medium changes: periodicity
- Slowest changes: envelop



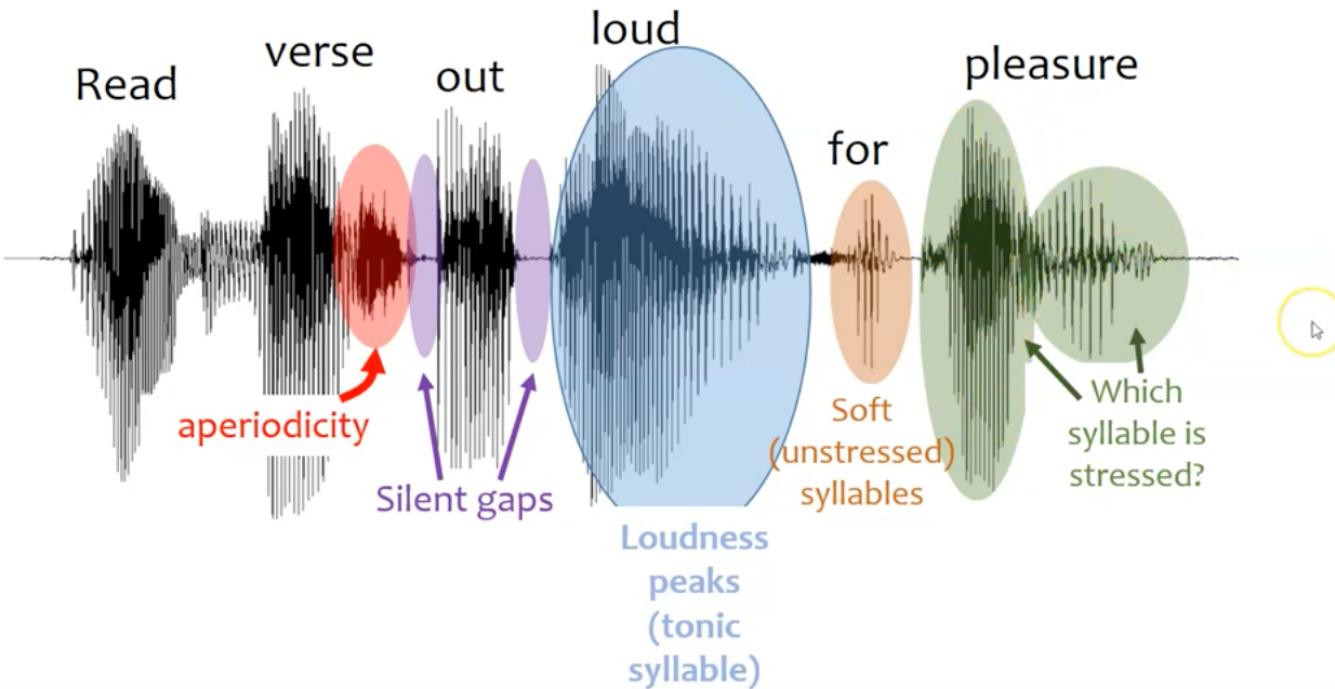
# ENVELOPE - SHOWS STRESSED SYLLABLES

- Cues for manner of articulation
- Cues for consonant voicing
- Durational cues
- Prosodic cues
- Emphasized syllables (stress)

**Envelope** (slow overall changes)  
Between 2 Hz and 50 Hz

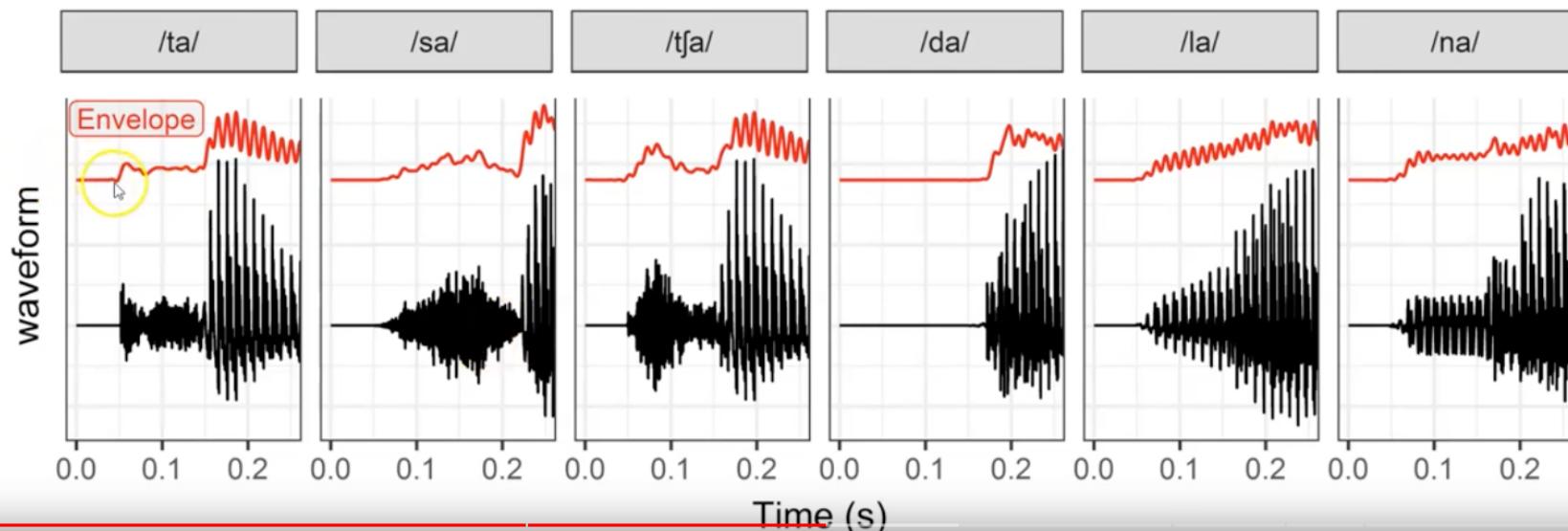


# Sentences contain lots of different temporal modulations



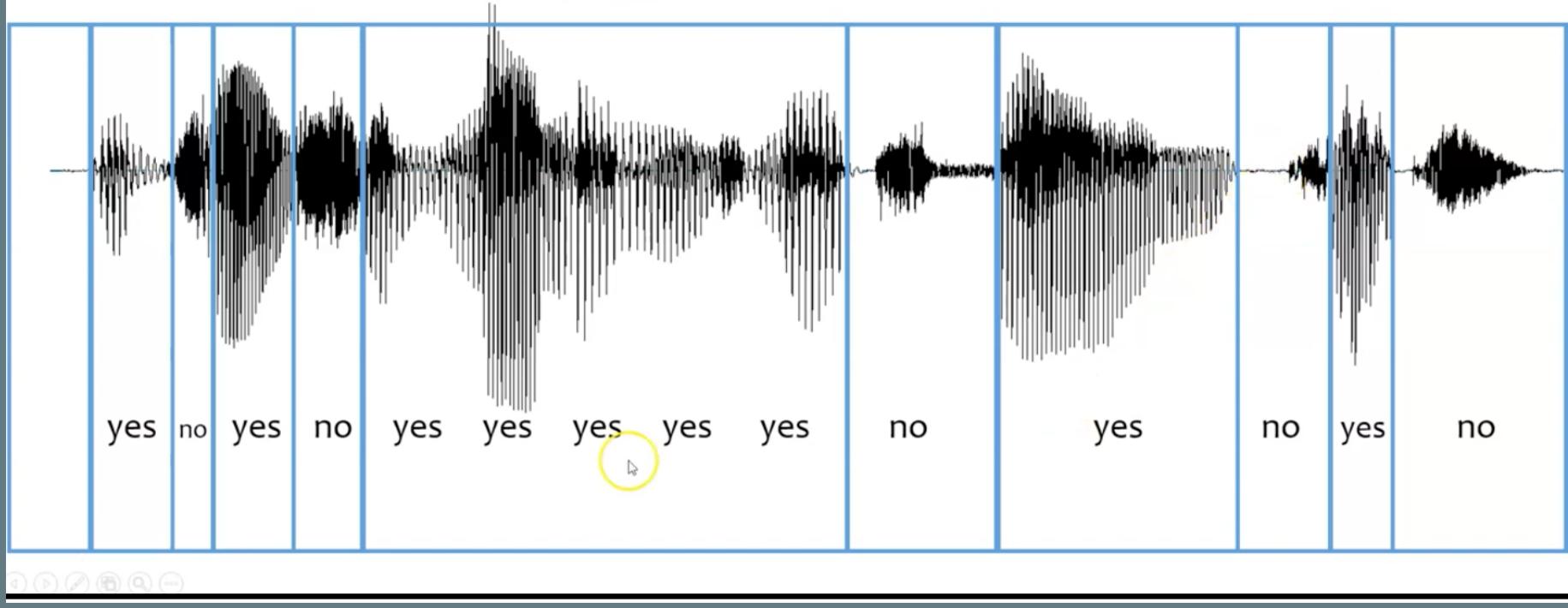
Sounds with different manner of articulation have different amounts of airflow constriction

Amount and type of airflow constriction will determine the envelope of the sound.



# Periodicity (between 50 and 500 Hz)

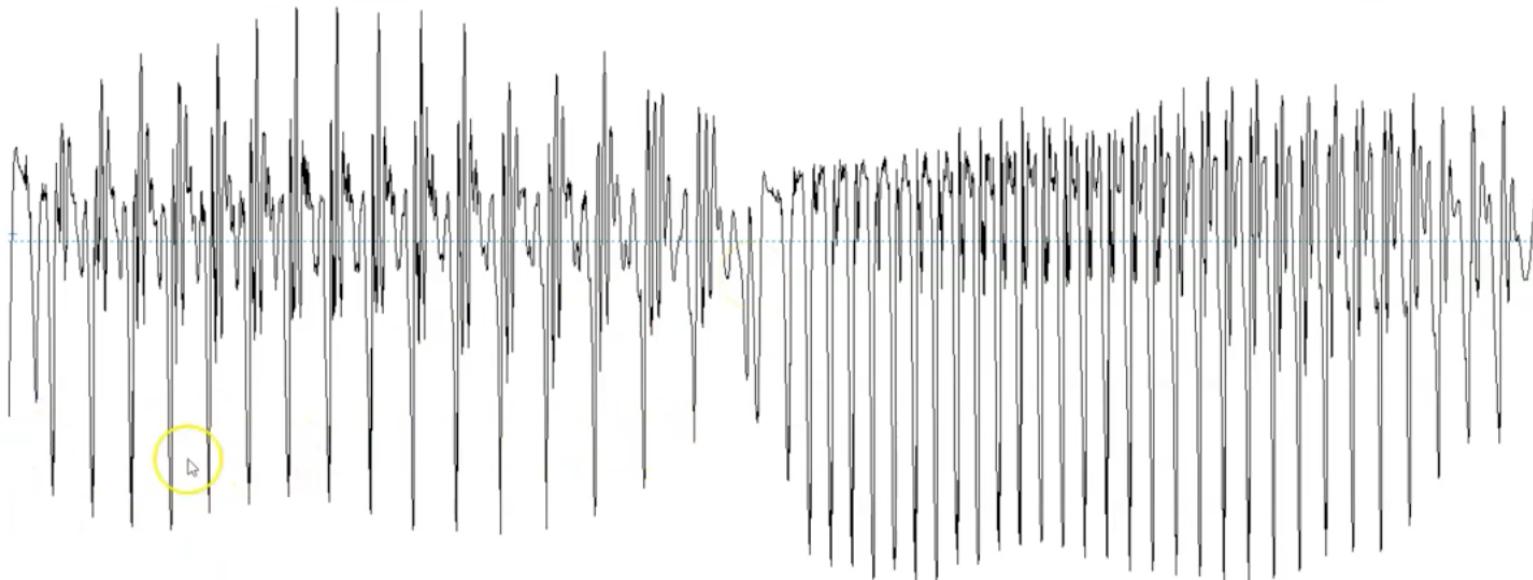
- Presence of voicing (all vowels and voiced consonants)
- Cues for intonation / pitch



## Speaker notes

- Periodicity for the yes words, also louder
  - also known as voiced sounds
- non periodicity for the no words

Periodicity tracking can be used  
to compare fundamental frequency (pitch)



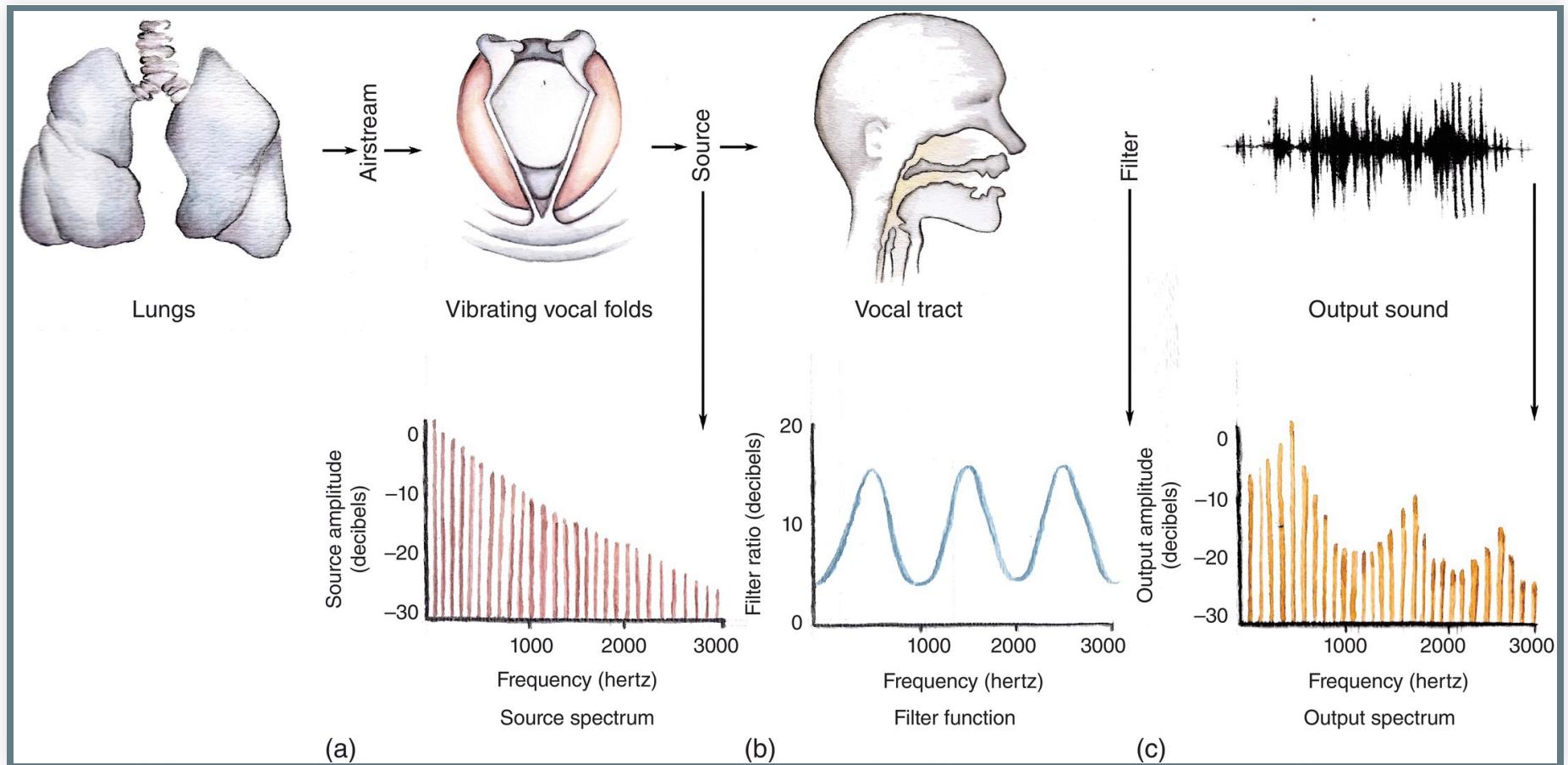
Slower rate of repetition  
Lower Fo

Faster rate of repetition  
Higher Fo

## Speaker notes

We can at least see that the signal on the right has a higher frequency than the signal on the left

# SOURCE-FILTER MODEL



## Speaker notes

Here's a basic model for how our vocal tract works.

# Functions of the source and the filter

- Source:

- Periodicity and pitch  
(fundamental frequency)
- Intensity
- Duration
- Phonation quality (breathy, modal, creaky)

harmonic frequencies

- Filter:

- vowel height, vowel advancement, rounding

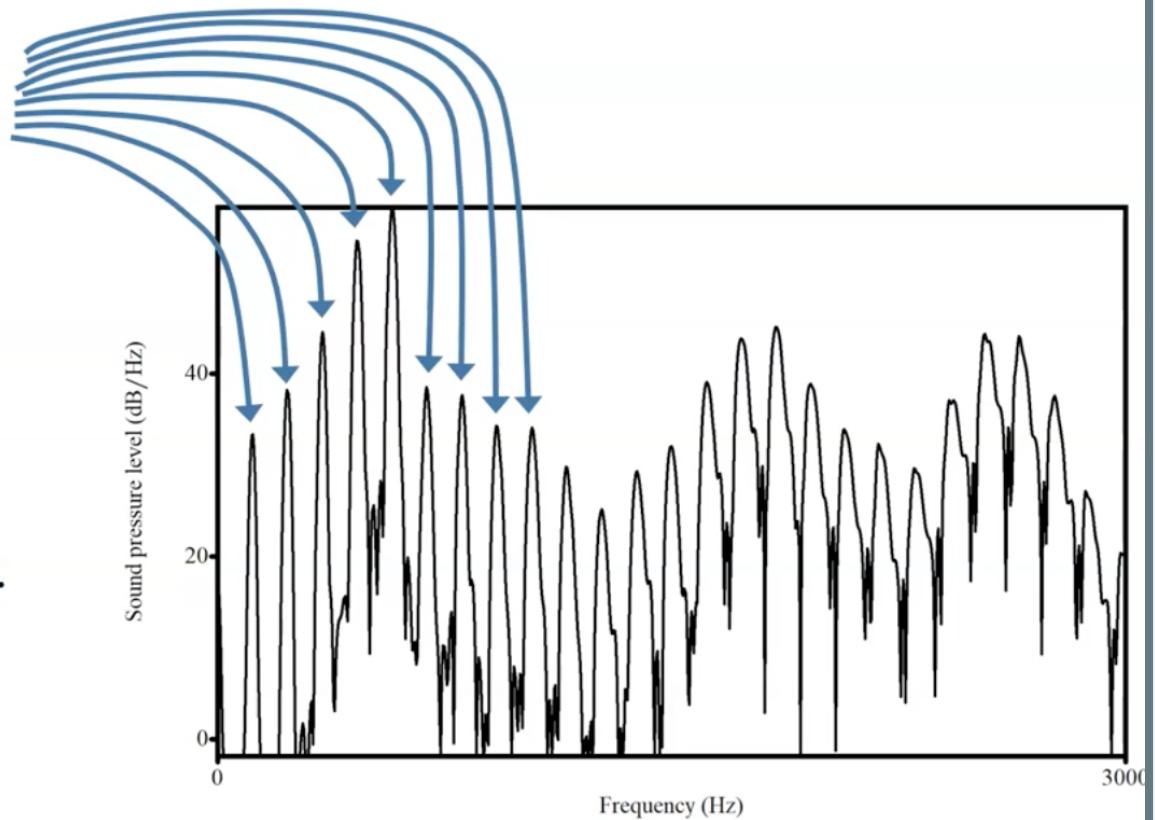
formant frequencies

## Speaker notes

The source is responsible for harmonic frequencies, the filter is responsible for formant frequencies.

# Harmonics are visible on the spectrum

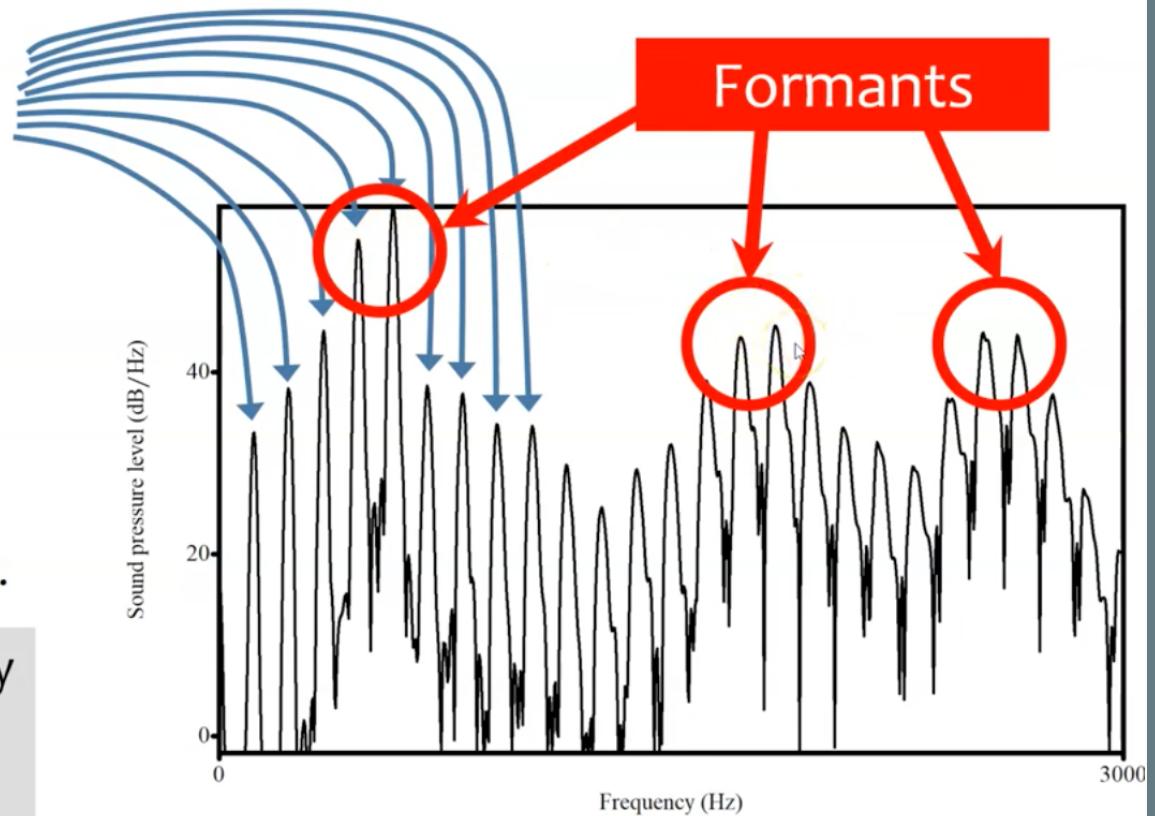
- Each of these vertical lines represents a harmonic
- Harmonics are equally spaced frequency components that are integer multiples of a fundamental frequency
  - E.g. 200, 400, 600, 800, etc.



# Harmonics are visible on the spectrum

- Each of these vertical lines represents a harmonic
- Harmonics are equally spaced frequency components that are integer multiples of a fundamental frequency
  - E.g. 200, 400, 600, 800, etc.

The fundamental frequency is repetition rate of the glottal SOURCE



## Speaker notes

Here we can see the first three formant frequencies.

# INTERACTIVE VOCAL TRACT

Here