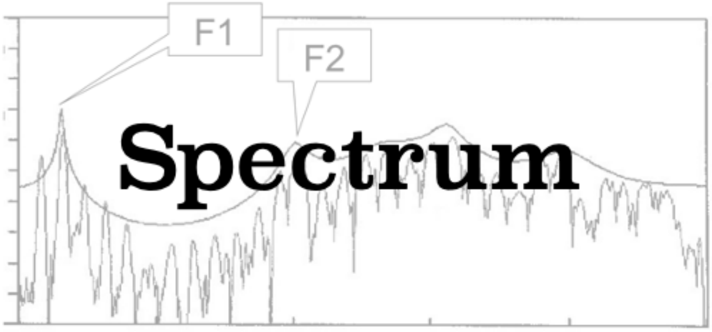
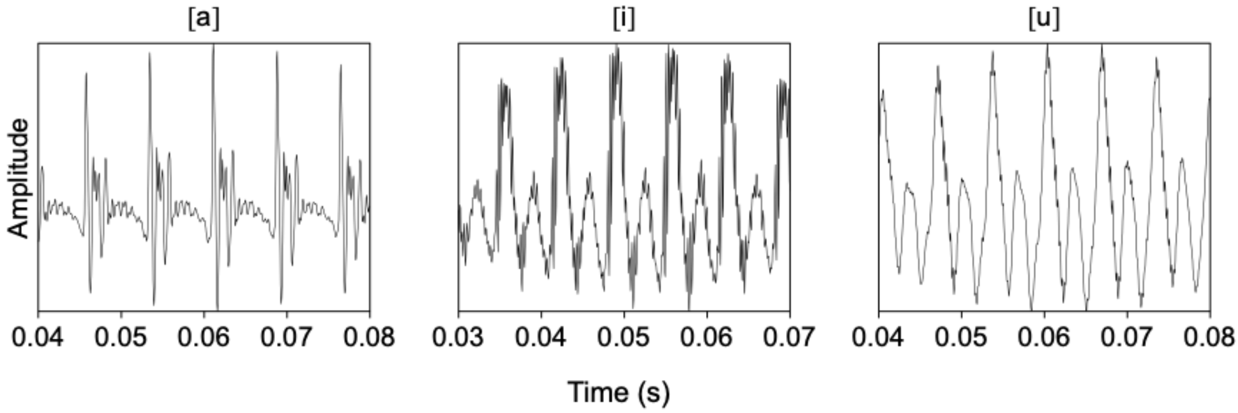
Spectrum



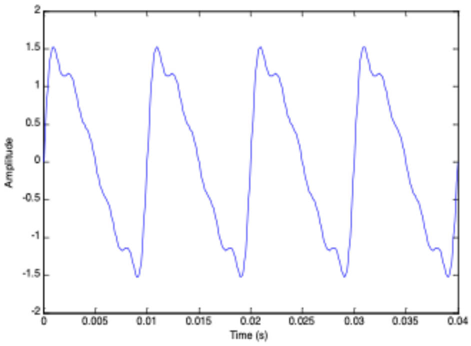
Fourier spectrum of [i], with an LPC spectrum overlaid.

From Kent & Read (1992) The Acoustic Analysis of Speech.

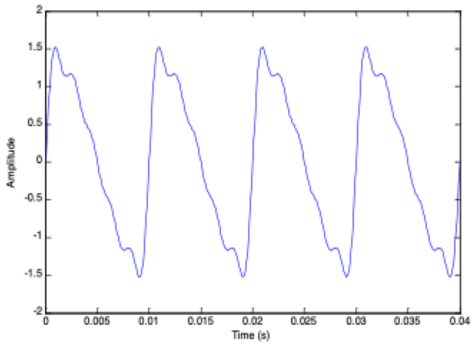


The complexity of a wave contributes to our experience of the quality of the sound. These differences in complexity are visible in the waveform, but somewhat difficut to characterize.

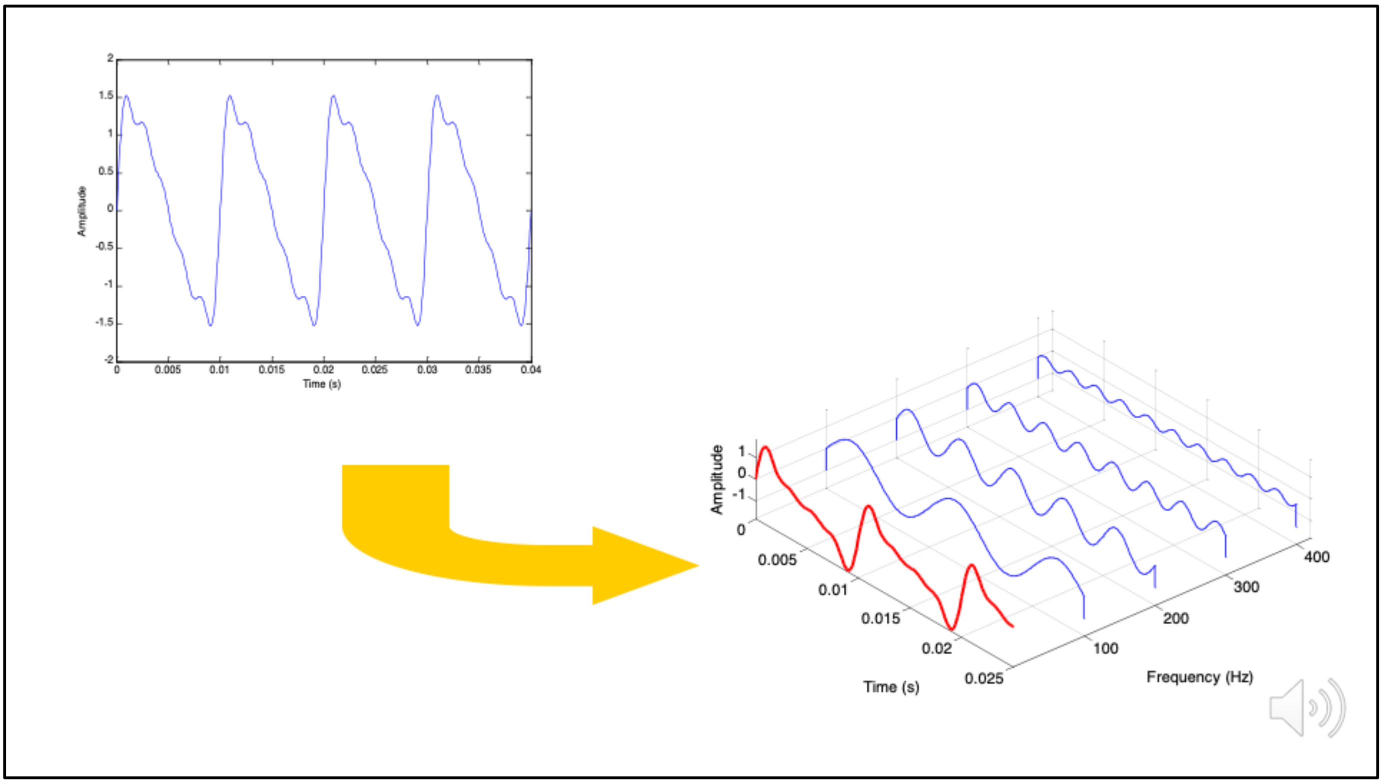
For example, we see here waveforms of three different vowels. Each of these vowels was produced with roughly the same F0, but differs from the others in terms of complexity. We can get a better of sense of the structure of this complexity by removing the time dimension from our visualization, and instead look at the frequency and amplitude of the component waves.



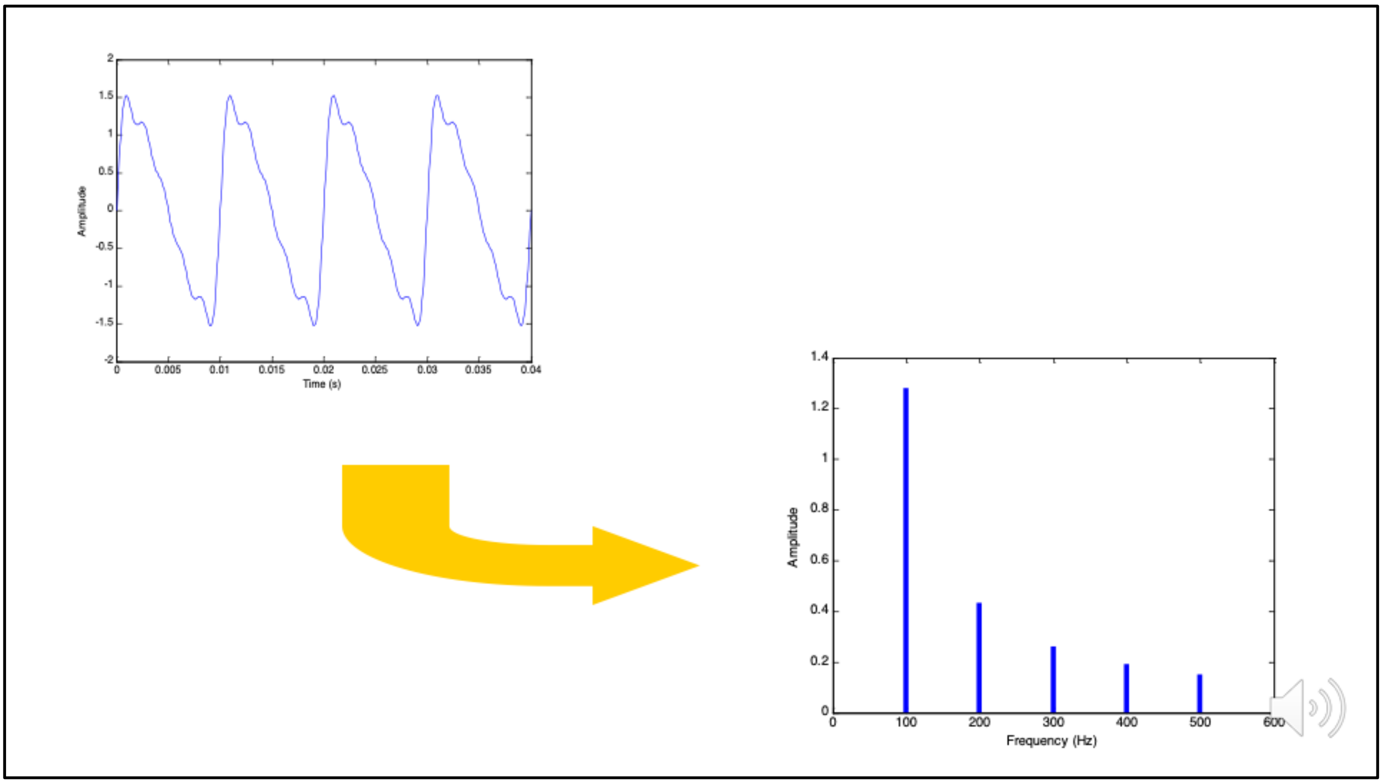
We know that complex waves are made up of at least two component sine waves. However, we cannot easily determine from the waveform, what the frequency and amplitudes of the component waves are.



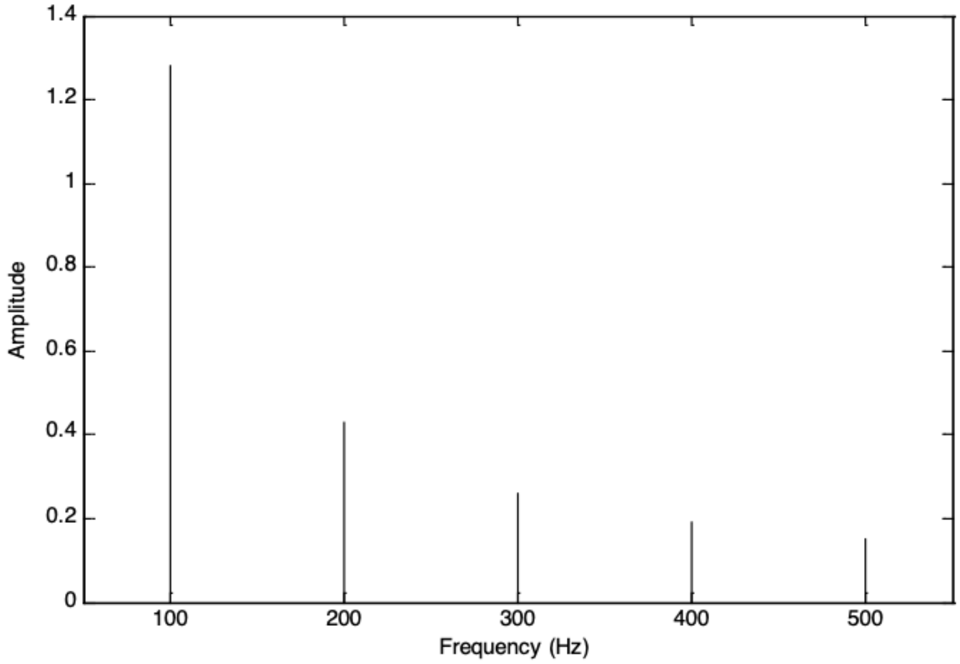
To that we use a mathematical process called the Fourier transform. This process is somewhat analogous to a prism breaking white light into its various wavelengths.



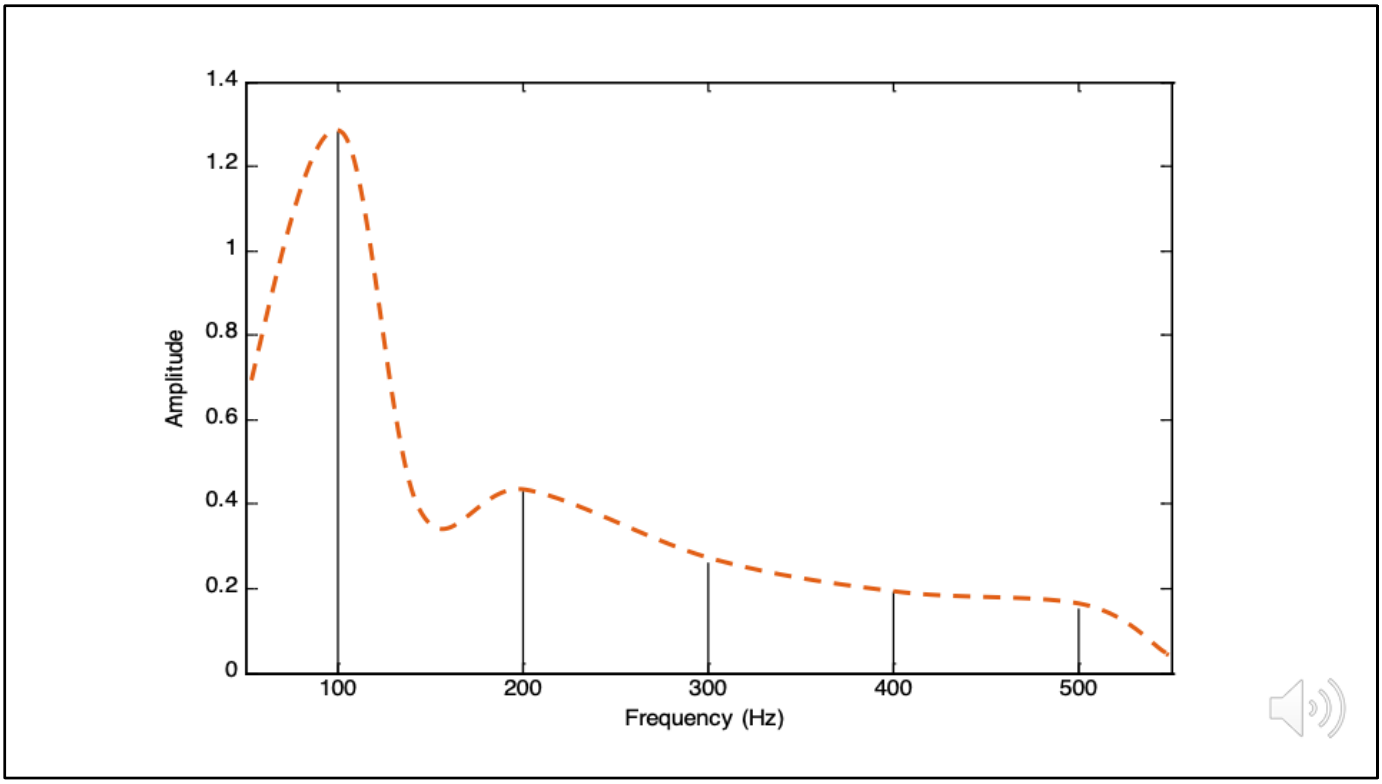
Here we have a visualization of a complex wave in red on the right, and each of its component simple waves in blue. Each simple wave appears as a vertical line along the frequency axis at the appropriate spot on the scale, with the amplitude is represented by the height of that line.



The spectrum represents each component simple wave. Since all simple waves are described by the same form (a sine wave), we can represent each component wave of a complex wave in terms of its amplitude and frequency and alone, eliminating the time dimension. This is what we are representing when we look at the spectrum of a complex wave.

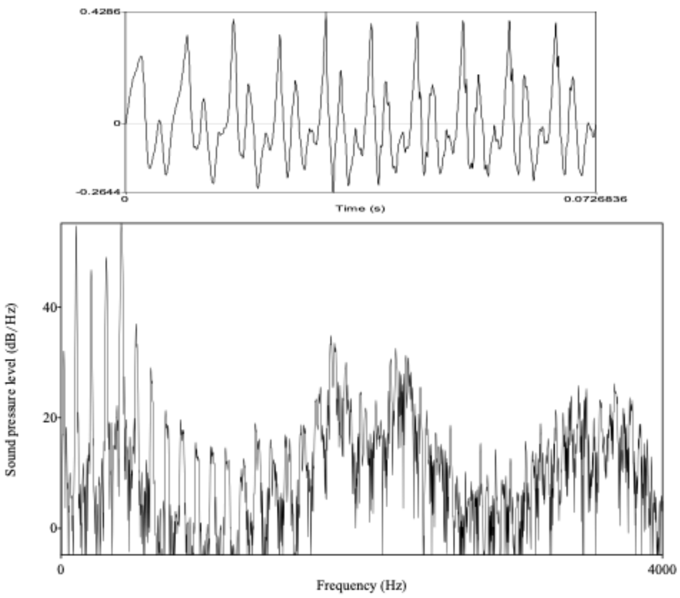


Here is an example of a line spectrum of a complex wave made up of 5 frequency components.



If we trace the shape of the harmonic peaks in the spectrum, we create what is called the spectral envelope. The precise shape of this envelope depends on the amplitudes of the component waves, as well as the method used to derive it.

(Question: what would the spectrum of a simple wave look like?)



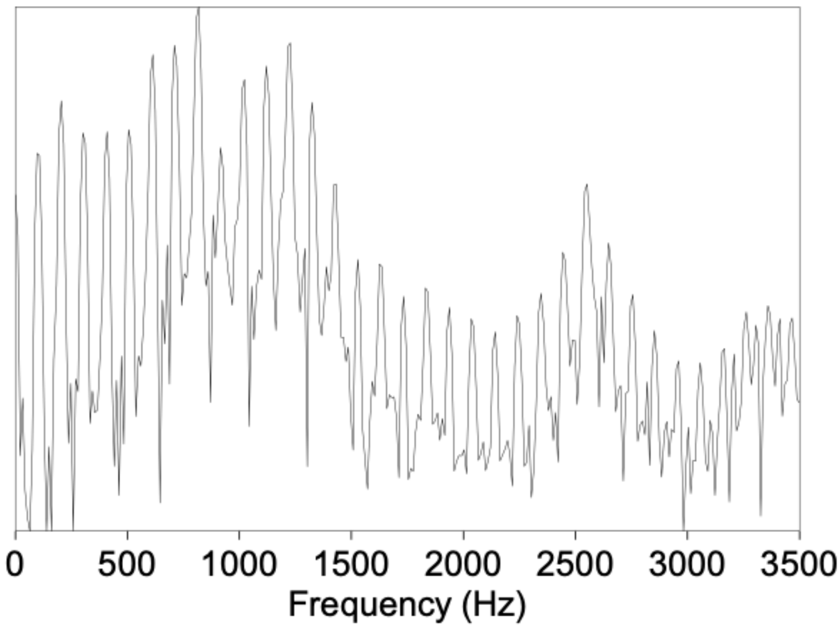
Because speech sound is highly complex, the spectrum is also more complex than the line spectra we have seen thus far. This structure is the result of the complex sound source (the voice) passing through and being shaped by the filter (the vocal tract).

These two components together result in the structure of the spectrum. The voice source contributes the **harmonic** structure, while the shape of the vocal tract contributes to the overall shape of the spectral envelope, resulting in what are known as **formants**.

Harmonics

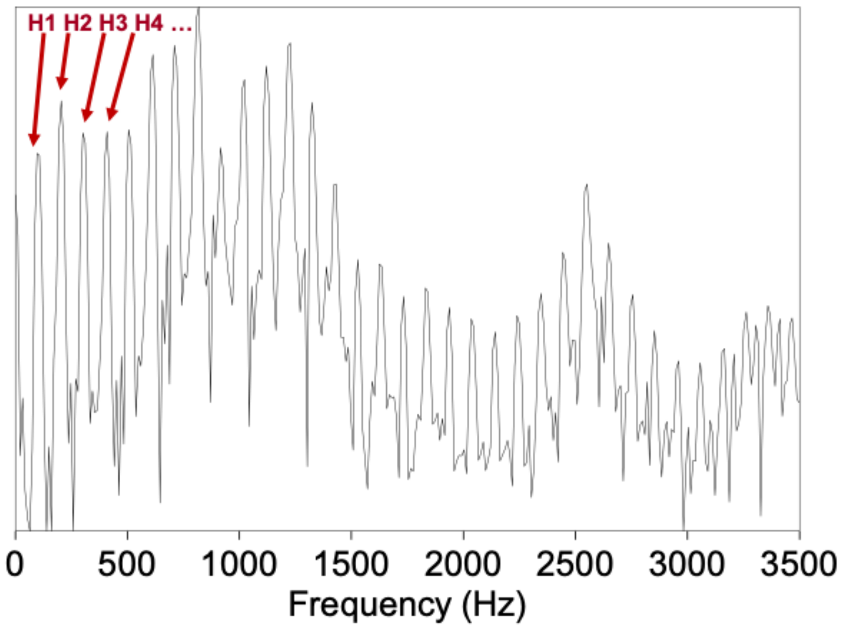
Source properties

The spectral shape of a periodic speech sound such as a vowel is the result of two different components interacting with each other. These components are the voice, also known as the source in the source filter model, and the vocal tract, also known as the filter.



The source characteristics are the fundamental frequency and its associated harmonics. These harmonics are represented as an array of frequency components which are equidistant from one another in the spectrum.

Figure. Spectrum of a section of the vowel [a].



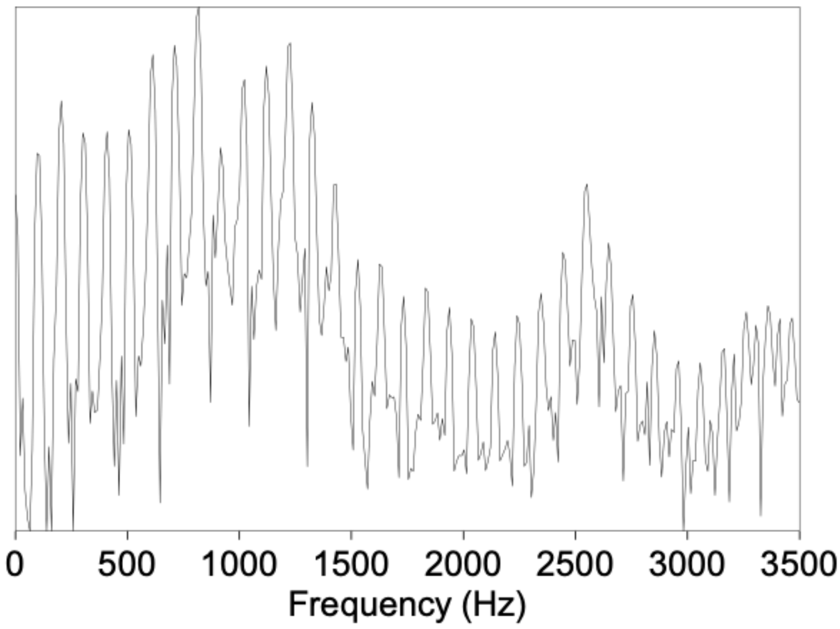
These are the harmonics. They are numbered sequentially starting with the first harmonic (H1). This lowest harmonic, so H1, is also the fundamental frequency (f0) of the harmonic series; the harmonics above H1, are found at every integer multiple of H1. So H2 is 2 times H1, H3 is 3 time H1 and so on.

Lower harmonics (on the left side of the frequency scale) tend to have a higher amplitude than higher harmonics (on the right side of the frequency scale), though this pattern is modulated by the presence or absence of formants, which amplify or dampen the harmonics.

Formants

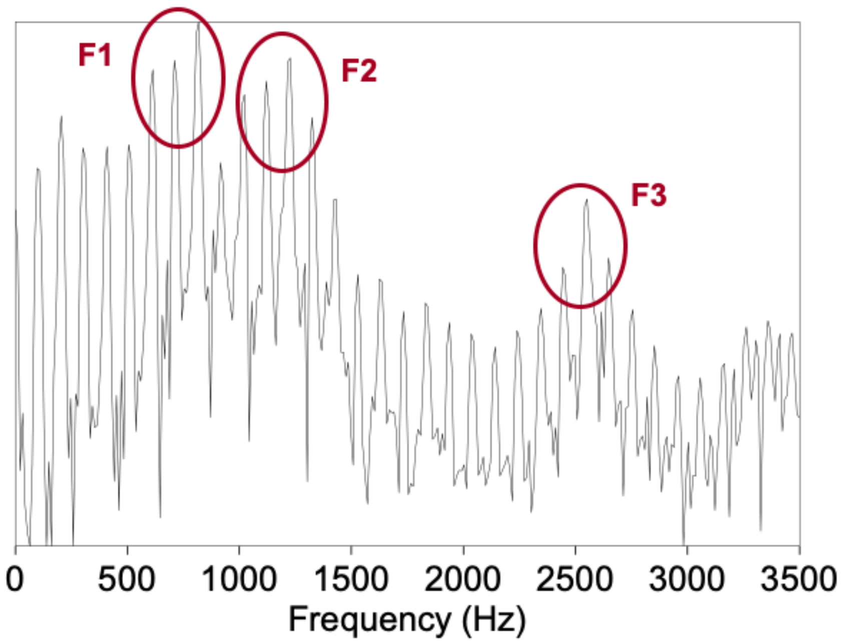
Filter properties

The filter shapes the complex wave that emerges from the larynx by amplifying or dampening frequency regions of the spectrum. Frequency ranges that are amplified by the filter are called formants and are important for the description of vowel quality.



We can see the affect of the filter on the spectrum by looking for frequency ranges where the harmonics are higher or lower in amplitude than we would expect given the general pattern for lower frequencies to have higher amplitudes over all.

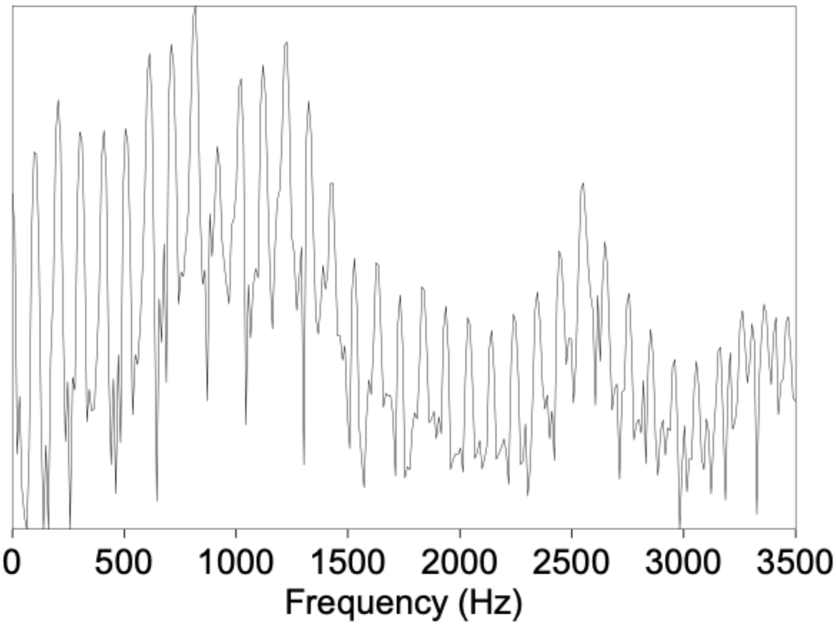
In this case, the frequency range of about 700-800 Hz is amplified, as is the 12-1300 Hz range. These are the first and second formant, respectively.



Similar to harmonics, formants are numbered sequentially, starting with F1. Unlike harmonics, however, there is no predictable relationship between the formant frequencies in the spectrum. Instead they are determined by the physical properties (such as the size, shape, and position of the articulators) of the vocal tract

It is also important to note here, that …

f_0 is not a formant



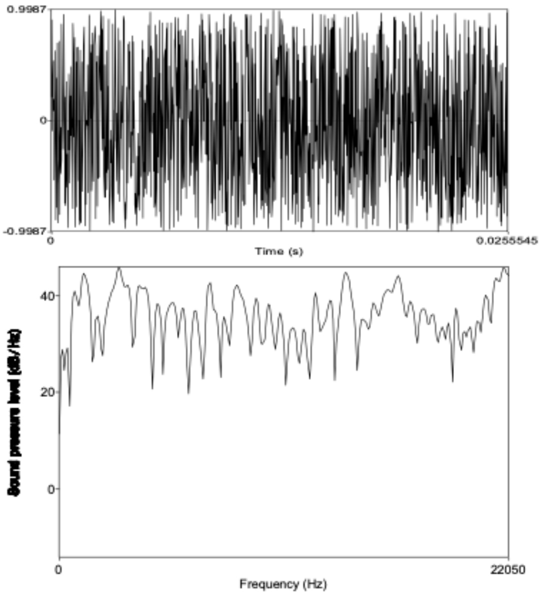
Despite the nomenclature, f_0 is not a formant
The fundamental frequency or f_0 is the same as the first harmonic, H1, and is the greatest common denominator of the components in a complex wave.
The higher harmonics are integer multiples of the this fundamental frequency.

f_0 = H1
H2 = 2*f_0…

The formants bear no direct relationship to f_0, and should not be confused with the harmonics.

Aperiodic spectrum

So far, we've been considering characteristics of the spectrum of periodic complex waves, but aperiodic waves also have a spectrum.



Unlike periodic sounds, the spectra of aperiodic sounds are flat
There are no harmonics, no fundamental frequency and the amplitudes of the various component waves do not decrease at higher frequencies. This is because, unlike periodic sounds, the vibratory pattern in aperiodic sounds is random. There is no pattern, and therefore, no predictable shape of the frequencies in the spectrum.