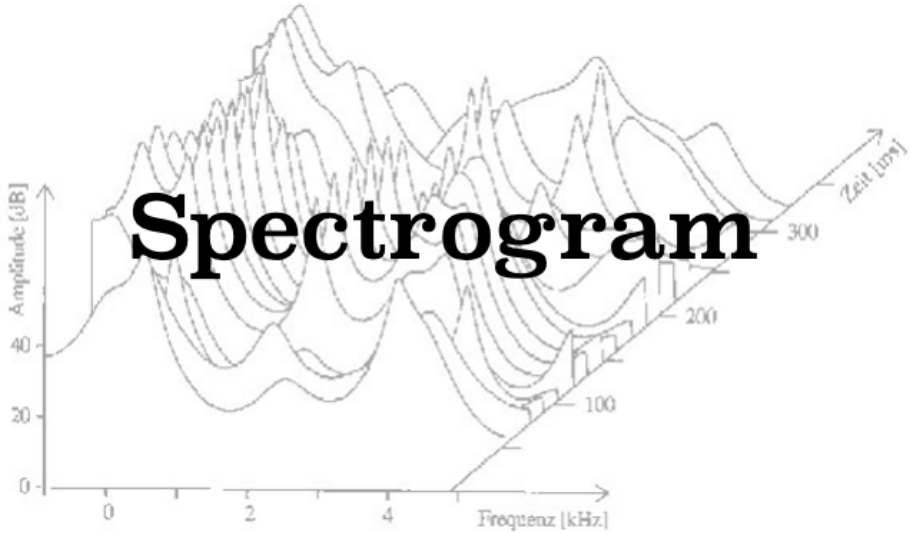
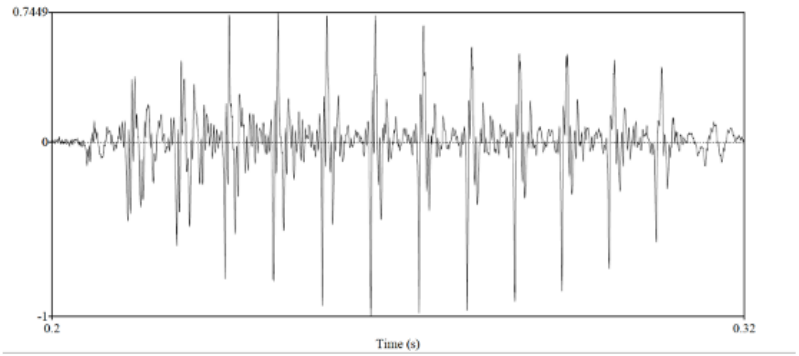
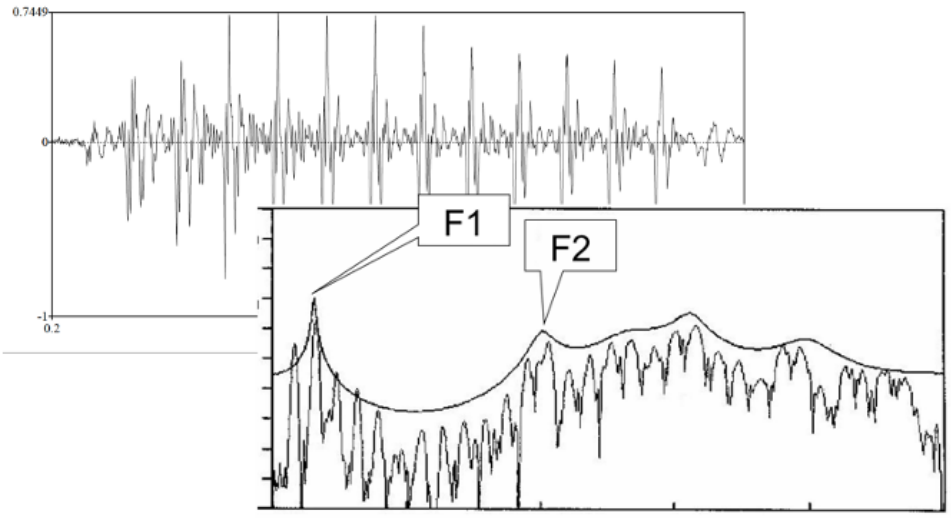
Spectrogram



In previous videos we've seen that sounds come in various forms, from simple, complex, periodic, aperiodic, etc, and we have looked at the ways we can represent these sounds.



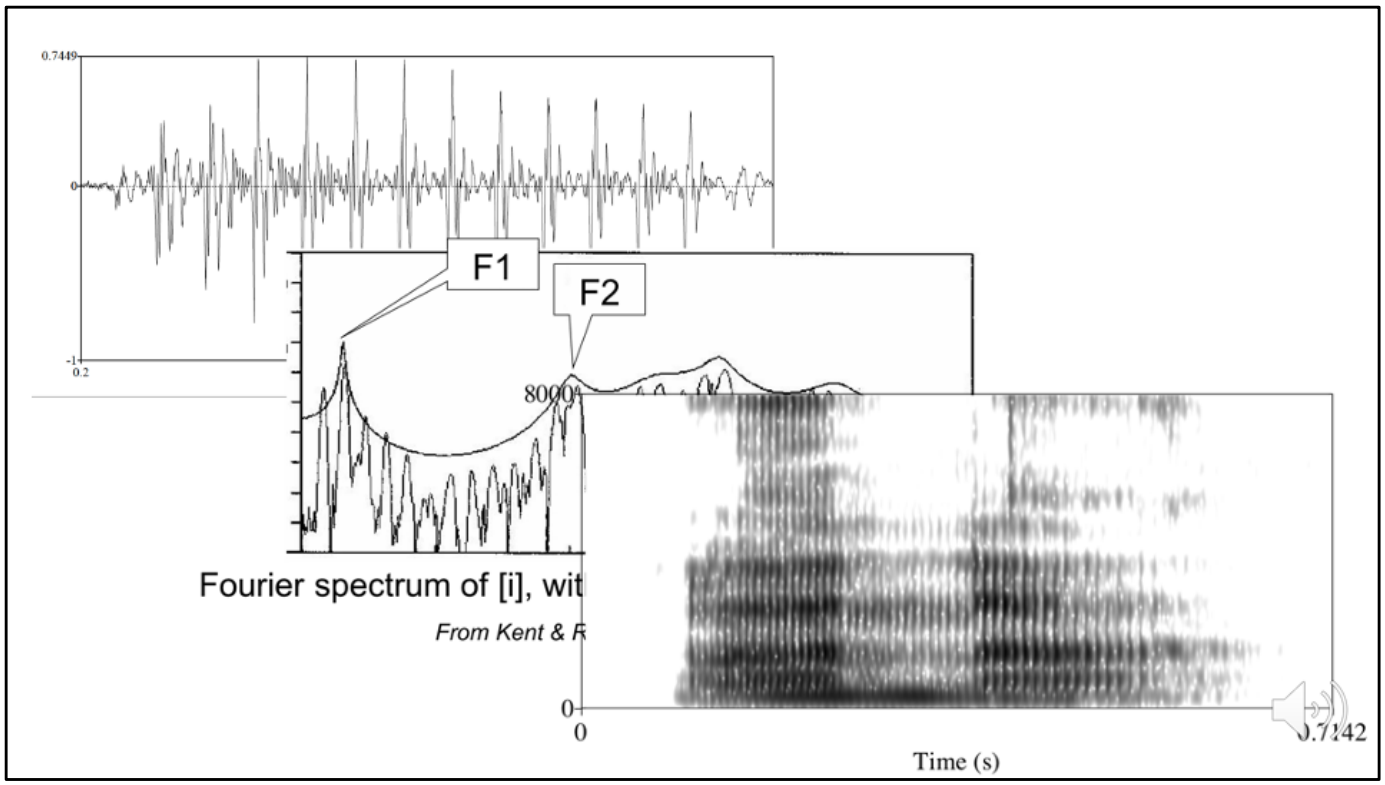
The waveform displays changes in amplitude over time, and from this we can calculate the fundamental frequency of periodic sounds.



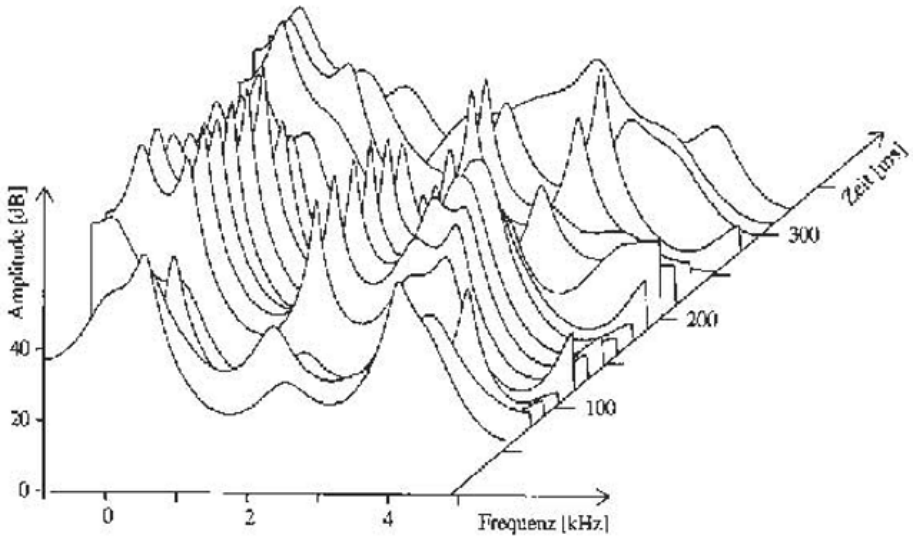
Fourier spectrum of [i], with an LPC spectrum overlaid.

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

The spectrum displays amplitudes of component wave frequencies in a complex wave, and reveals properties of both the source (known as harmonics) and the filter (known as formants)



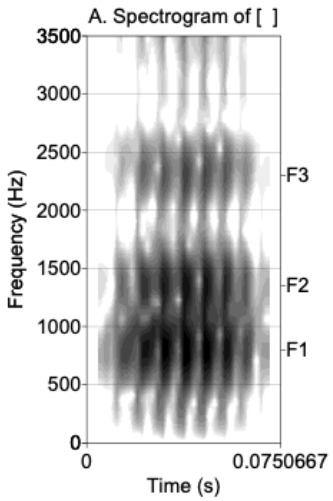
In this video we will look at a third representation of sound, known as the spectrogram,

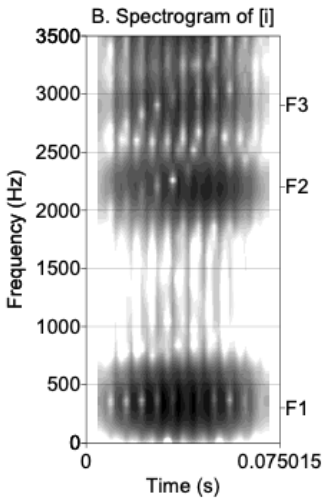


The spectrogram is a representation of sound that shows amplitude, frequency and time all in one figure. We can think of this as a series of spectra lined up at sequential points in time, somewhat like the figure shown here. This figure has a spectrum visible at the front, with amplitude on the y-axis and frequency on the x-axis. Time is shown here on the z-axis starting from zero at the front and moving forward in time in subsequent layers.

The sort of representation shown here gives us a bit more information about how the spectrum changes over time with varying speech sounds, but because we are looking at it on a 2-dimensionl screen, we still cannot see all of the information that we might need.

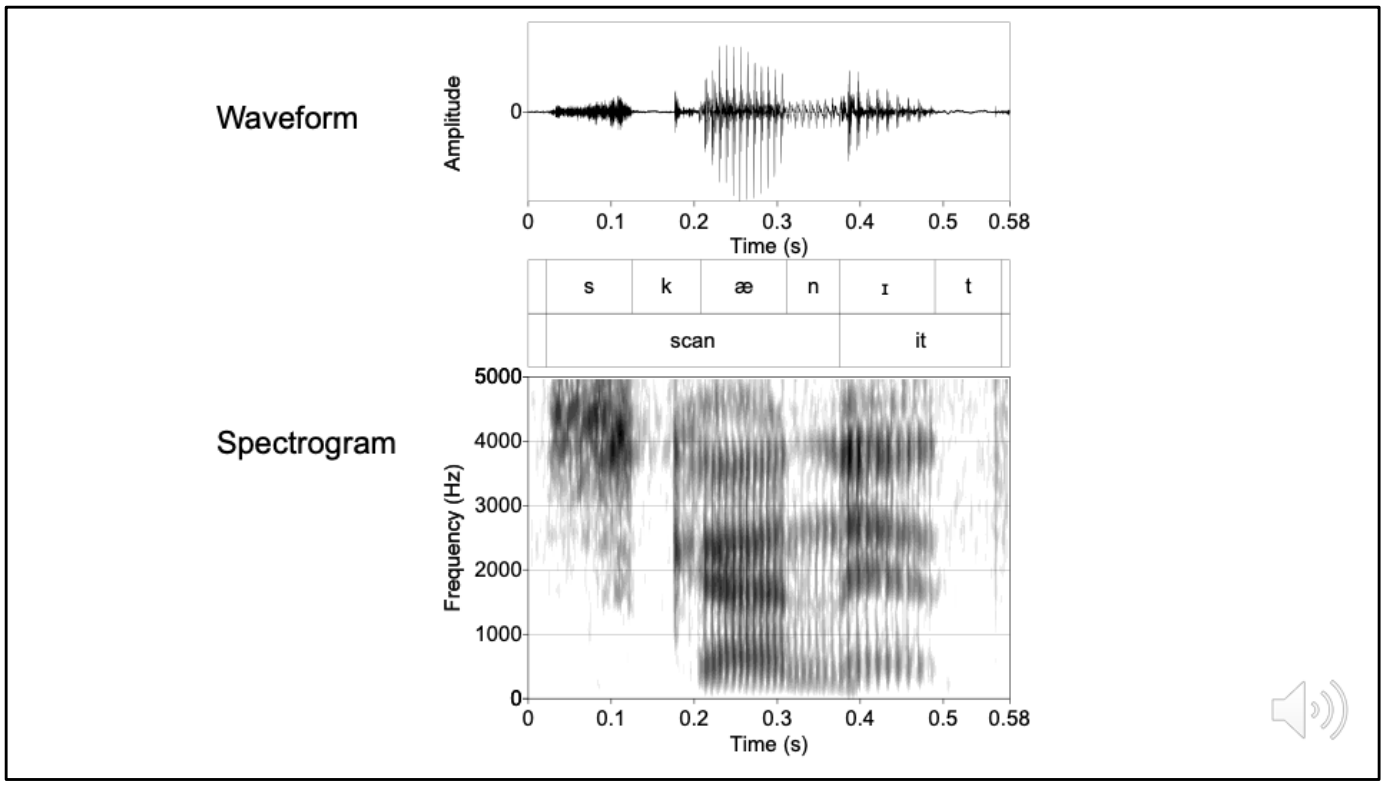
If we orient this figure so that time is on the x-axis, and frequency on the y-axis, we can then represent amplitude with grayscale shading.





Here we have examples of two spectrograms demonstrating its three-dimensional nature.
The spectrogram displays all three dimensions that are present in the waveform and the spectrum.

The x-axis represents time. The y-axis represents frequency, and the shading represents amplitude. In these two spectrograms we can see that on the left, the frequency range from about 500-1500 Hz has high amplitude (represented by darker shading), while on the right, the same range has lower amplitude (represented by lighter shading and even some areas of white or zero amplitude.)



Often in linguistic investigations of speech, the waveform and spectrogram will be presented together, with time-aligned phone and word level transcriptions. As we will see, the spectrogram will allow us to identify with more precision specific phones based on their acoustic characteristics.