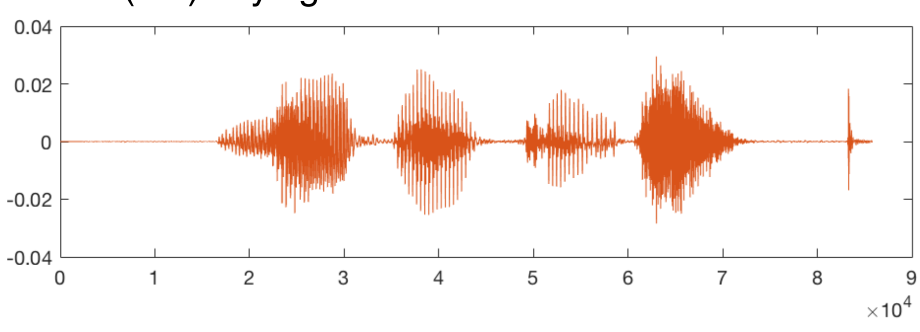
Signal Processing and Linear Systems1

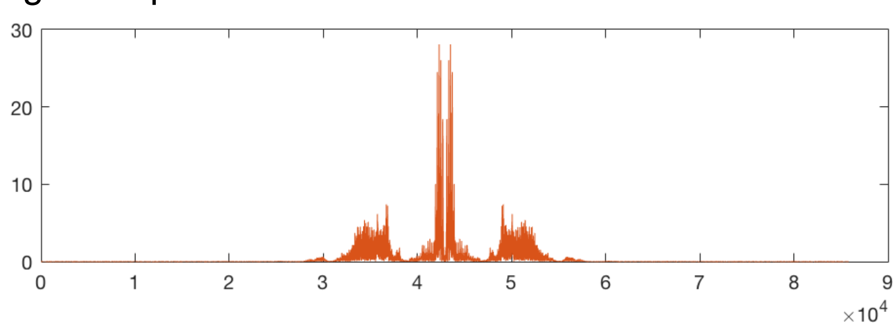
Lecture 14: Short-Time Fourier Transform

Nicholas Dwork
www.stanford.edu/~ndwork

Here is someone (me) saying “Zebra Finch.”



Here is the signal’s spectrum.

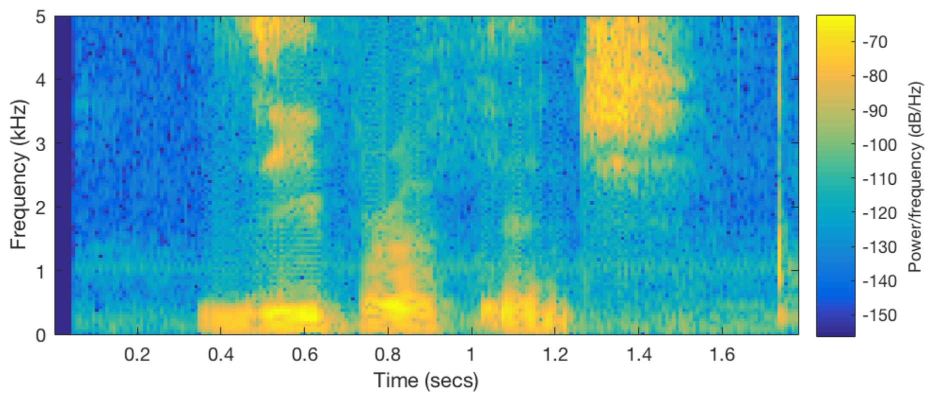


We saw a spectrum of someone saying “Zebra Finch”

We see which frequencies were significant. But we know that with speech, the spectrum changes over time.

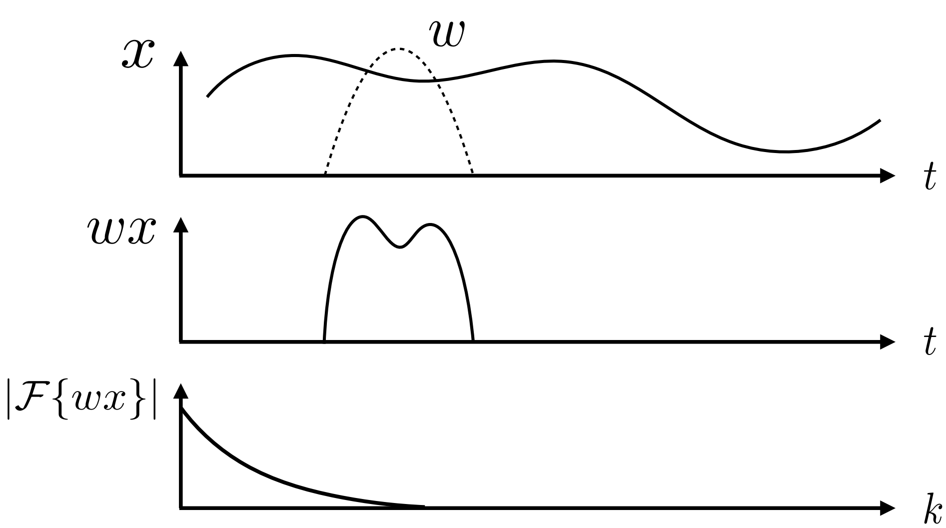
How did the spectrum change as a function of time?

This is what the Short-Time Fourier Transform will show us.



Approach

Find the spectrum for short segments of the signal.



Repeat for different window positions.

Short-Time Fourier Transform (STFT)

$$X(s,\tau) = \int_{-\infty}^{\infty} x(t)\, w(t-\tau)\, e^{-i\,2\pi\, st}\, dt$$

Discrete-Time STFT:

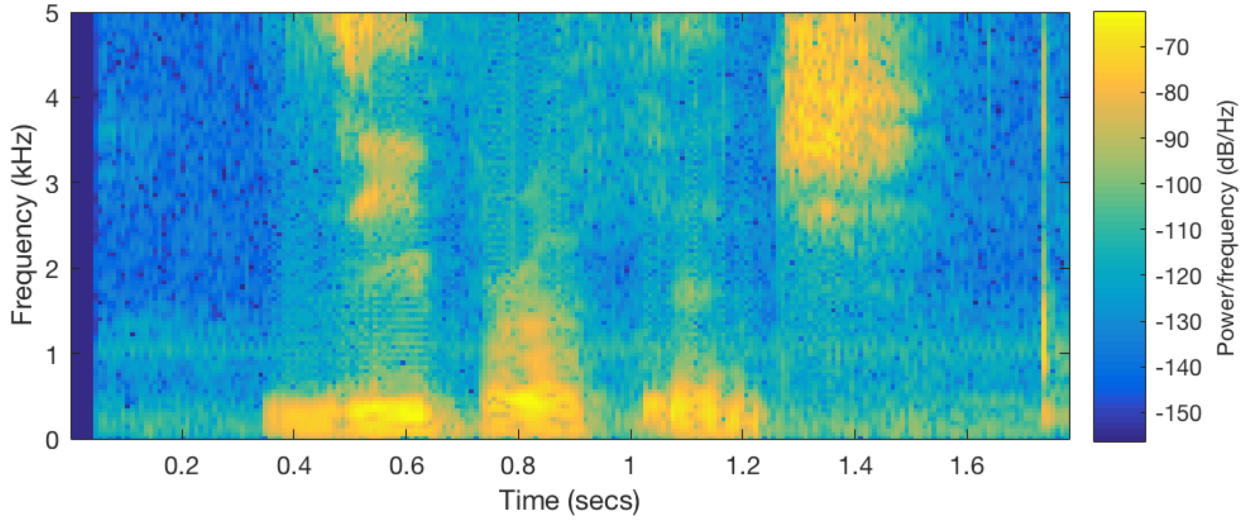
$$X(s,m) = \sum_{n=-\infty}^{\infty} x[n]\, w[n-m]\, e^{-i\,2\pi\, sn}$$

Discrete STFT:

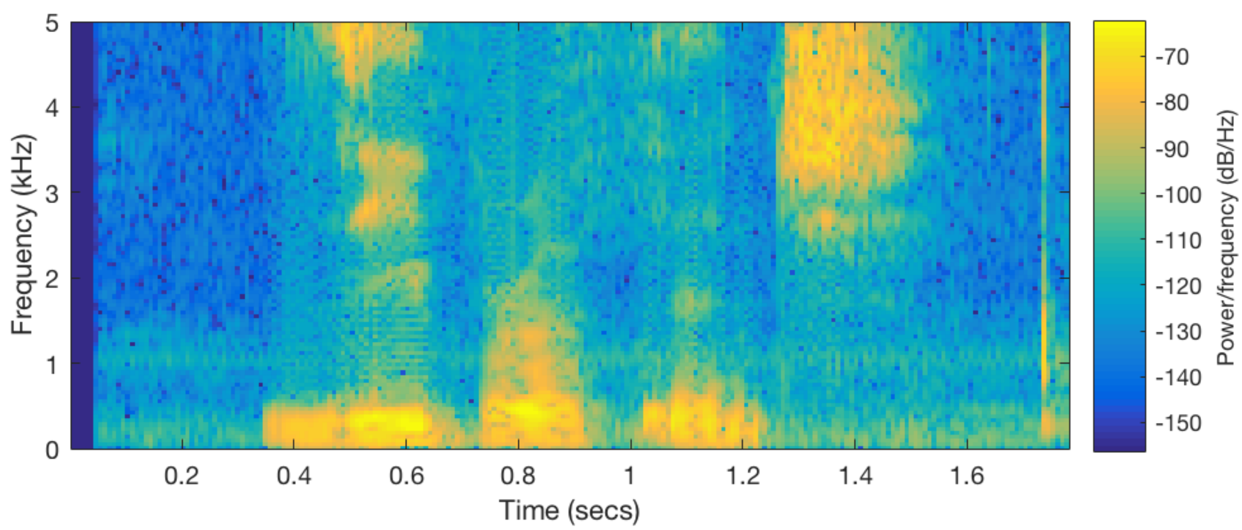
$$X[k,m] = \sum_{n=-\infty}^{\infty} x[n]\, w[n-m]\, e^{-i\,2\pi\,\frac{kn}{N}}$$

Spectrogram

Show the PSD for each windowed function.

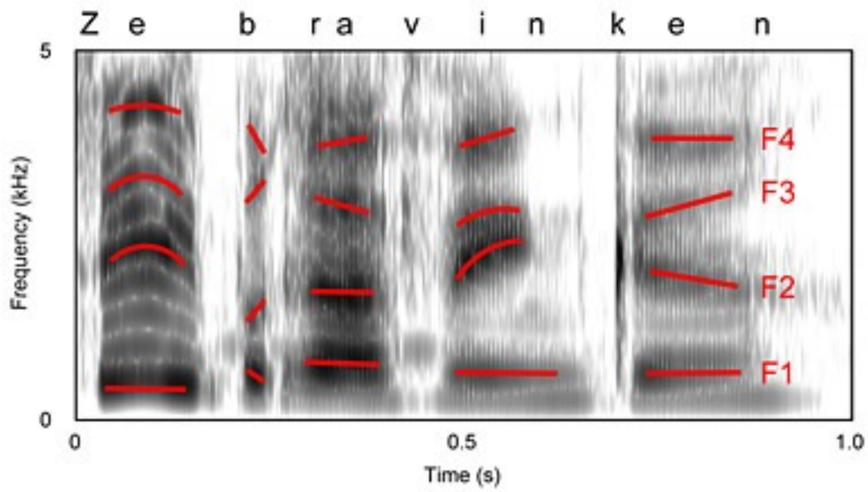


We get to choose to have good resolution in frequency or in time. We don't get both.



Formants

Characteristic spectrograms of speech.



http://www.news.leiden.edu/news-2011/birdsong-much-closer-to-human-speech-than-originally-thought.html

When computing the Discrete STFT, there are two parameters:

Which window should we use?

How much should each window overlap?

Inverse Discrete STFT

Goal: given the STFT, recover the original signal.

There is NOT always an inverse Discrete STFT.

This is because the window may distort the signal beyond recovery.

For many windows, recovery is possible.

There are several different methods.

IDFT Method of Inverse STFT

At each time point, compute the inverse DFT.

Divide by the window.

In those regions where the windows overlap, compute a weighted average.

This method works well when the data is stored on a computer.

This method is sensitive to errors in streaming data, so it’s not used in practice.

FBS Inverse STFT

$$y[n] = \frac{1}{N\,w[0]} \sum_{k=0}^{N-1} X[n,k]\, e^{i2\pi\frac{nk}{N}} = \frac{1}{w[0]} \mathrm{IDFT}\{X\}[n]$$

$$w[rN] = 0 \text{ for } r = -1, 1, -2, 2, \ldots$$

More stable than the IDFT method.