Digital Signal Analysis and Applications
Lecture 13: Short Time Fourier Transform

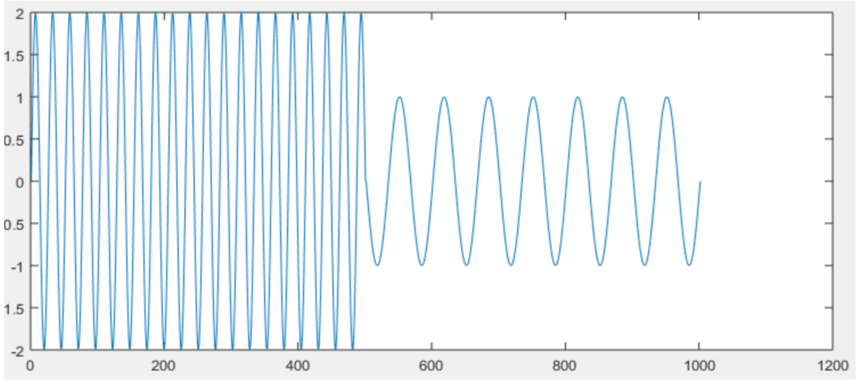
Vineet Gandhi

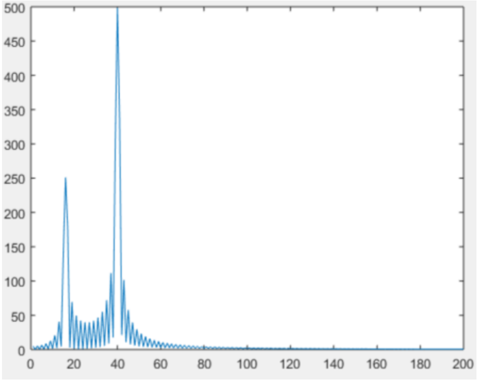
Center for Visual Information Technology (CVIT), IIIT Hyderabad

Consider a signal

- Consider an audio signals on $t \in [0,1]$

$$g(t) = \begin{cases} 2 * \sin(2\pi \cdot 39t)\,, 0 \le t \le 1/2 \\ \sin(2\pi \cdot 15t)\,, 1/2 < t \le 1 \end{cases}$$

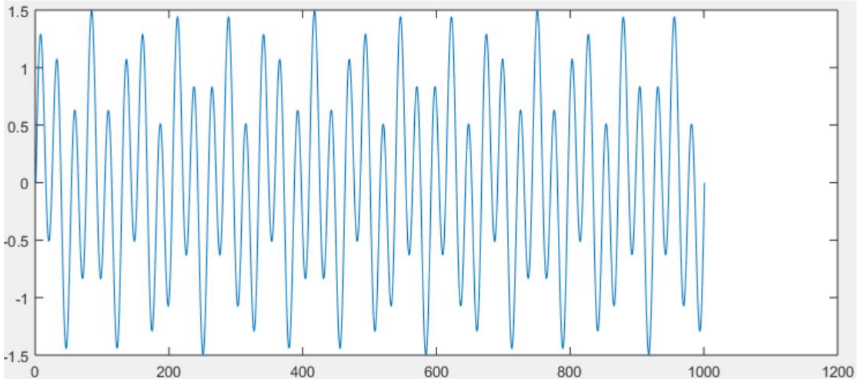


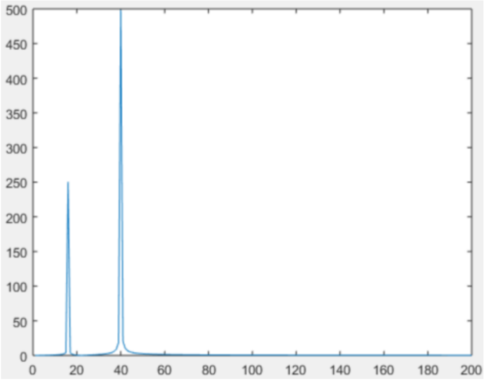


Consider another signal

- Consider two different audio signals on $t \in [0,1]$

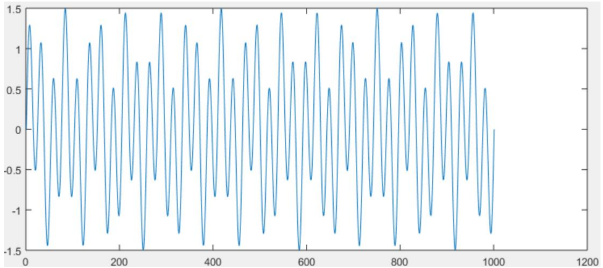
$$f(t) = \sin(2\pi \cdot 39t) + 0.5\sin(2\pi \cdot 15t)$$

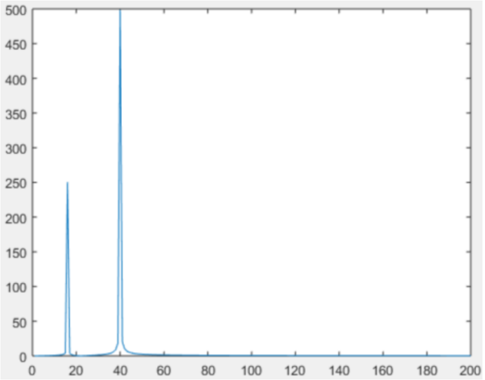




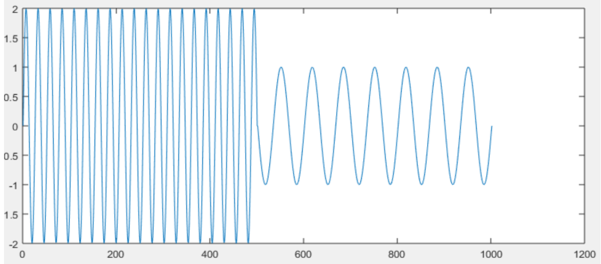
Non Locality of DFT

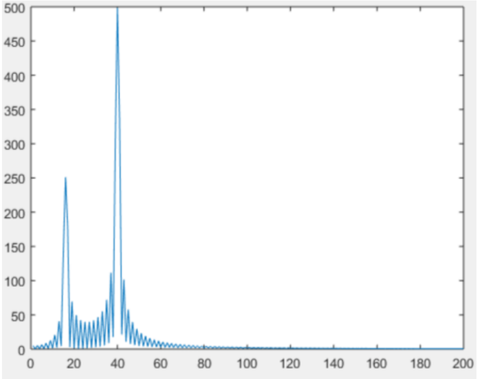
$$f(t) = \sin(2\pi \cdot 39t) + 0.5\sin(2\pi \cdot 15t)$$





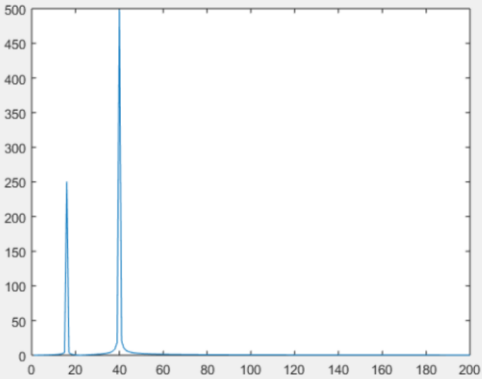
$$g(t) = \begin{cases} 2 * \sin(2\pi \cdot 39t)\,, 0 \le t \le 1/2 \\ \sin(2\pi \cdot 15t)\,, 1/2 < t \le 1 \end{cases}$$

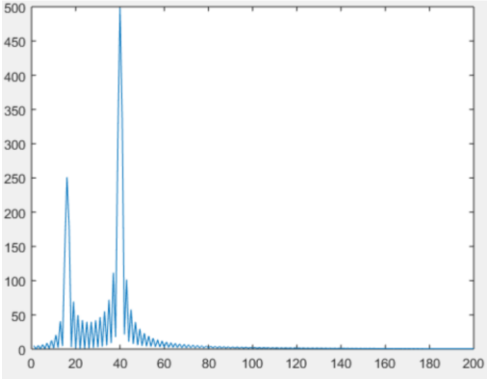




Non Locality of DFT

- Each signal contains dominant frequencies at 15 and 39 Hz
- Magnitudes of DFT are otherwise fairly similar





DFT is global in nature!

The fact that f(t) and g(t) are quite different in time domain is difficult to glean from the DFT graphs!

Non Locality of DFT

What can we do to solve this problem?

Windowing

- Break the signal into blocks “windows” in time domain
- Certain frequencies may be present in some blocks and not in others
- Block size small enough so that frequency content is relatively stable over the window
- Apply DFT to each window independently
  - Adjacent blocks may overlap
- Represent the signal as a sequence of short-time DFT’s

Windowing

A rectangular window:

- Starting position m
- Length M Samples, m+M < N
- All samples x_j with j<m and j>m+M are zeroed out, others unchanged
- The resulting vector y has components $y_j = w_j x_j$
- Where the vector w defined as:

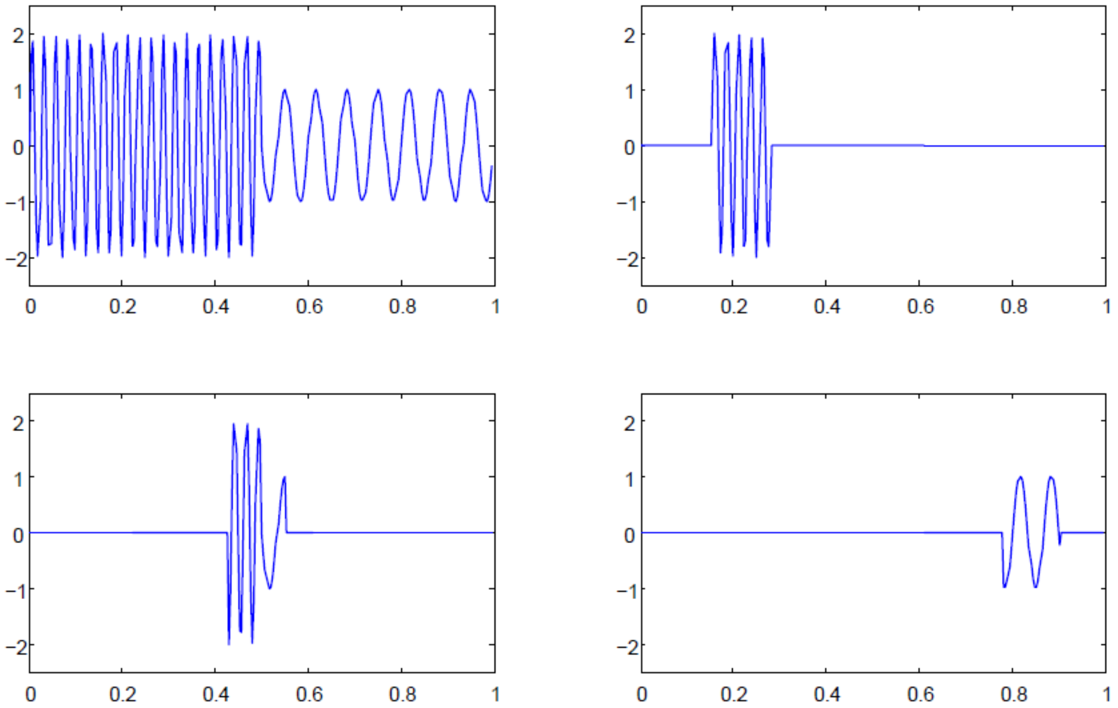
$$w_j = \begin{cases} 1, & m \leq j \leq m + M - 1 \\ 0, & otherwise \end{cases}$$

is called a (rectangular) window

Windowing

Original (upper left) and windowed signals with N = 256, M = 32

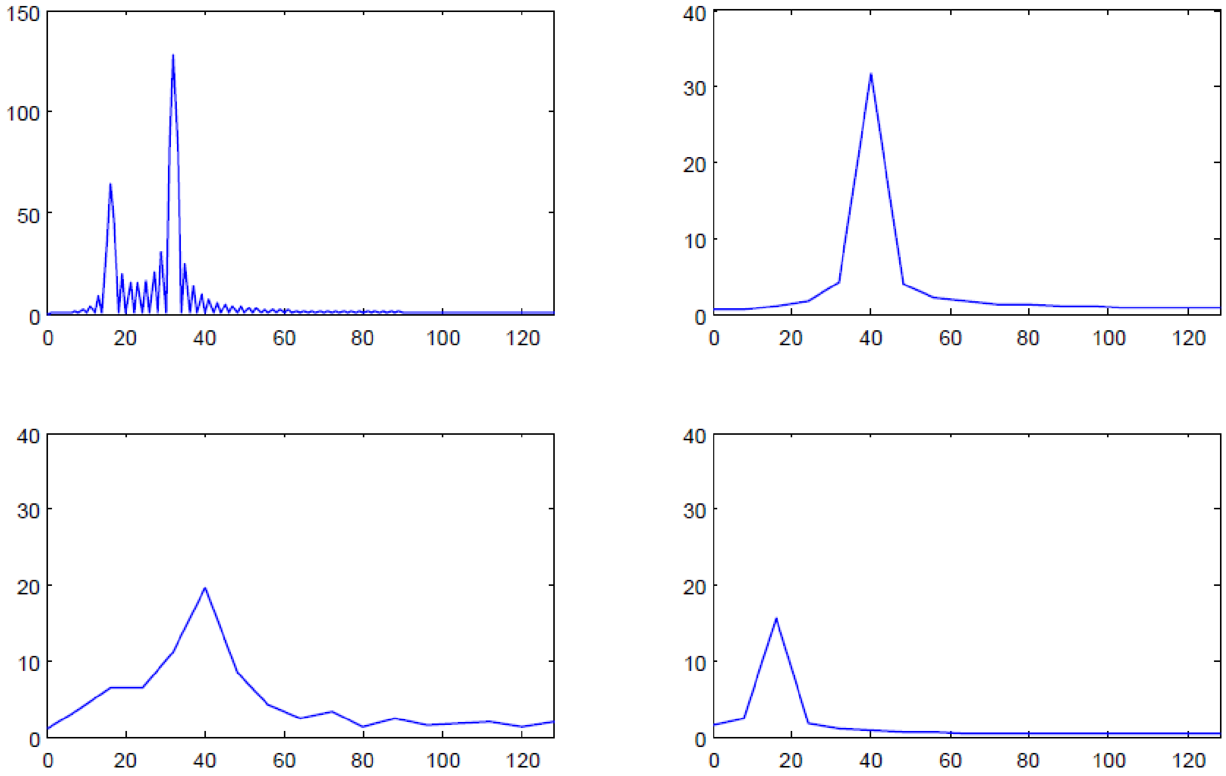
- m = 40
- m = 110
- m = 200



Windowing

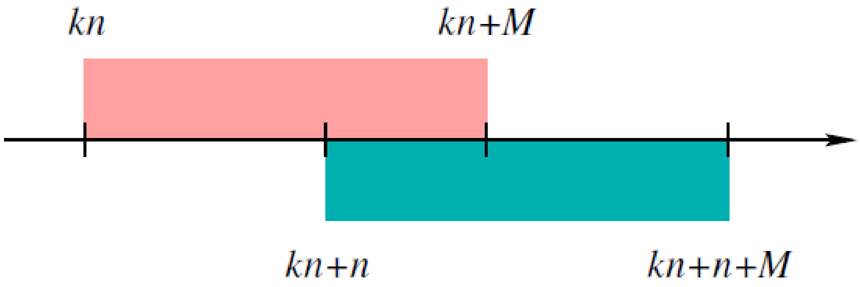
Global DFT spectrum and
M-point DFT spectra of
windowed signal
N = 256, M = 32

- m = 40
- m = 110
- m = 200



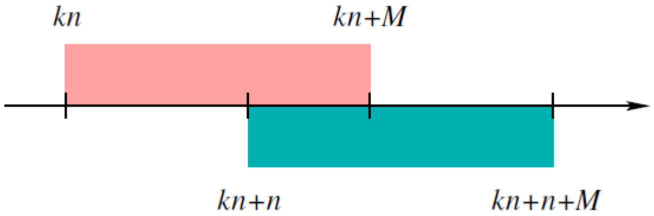
Windowing

- A collection of SFT’s computed over windowed portions of the signal, is called a short-time Fourier transform
- Adjacent windows may overlap
- Let m = k . n be the starting point of the kth window
- The integer n controls in overlap
- No overlap for n=M



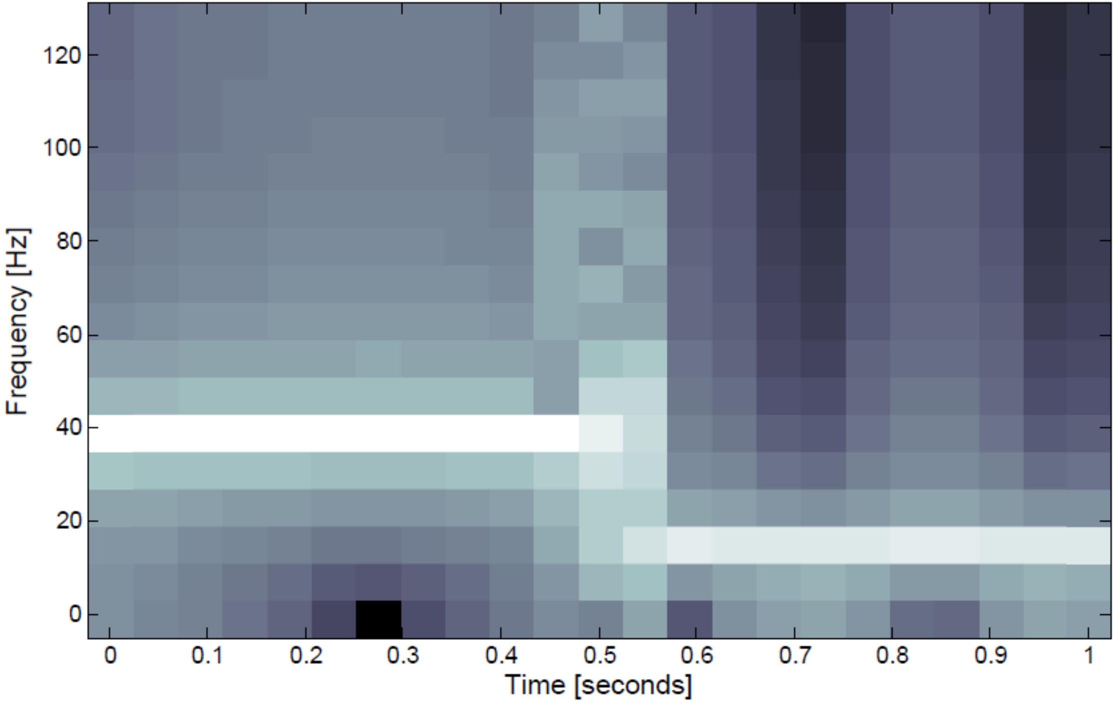
Windowing

- The kth block of data is: $(x_{kn}, x_{kn+1}, x_{kn+2}, \dots, x_{kn+M-1})$, for $k = 0,1, \dots, [N - M]/n$



- Compute the M-point DFT of each block, plot its magnitude as a kth column of the intensity image
- The resulting plot is called a *spectrogram*
- In next page illustration, we take M=32 and n=10
- How many total blocks?

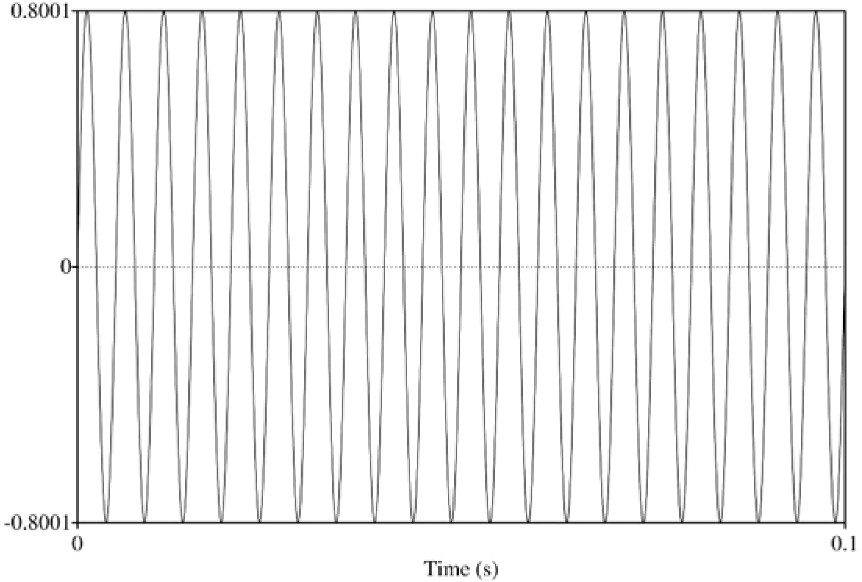
Windowing

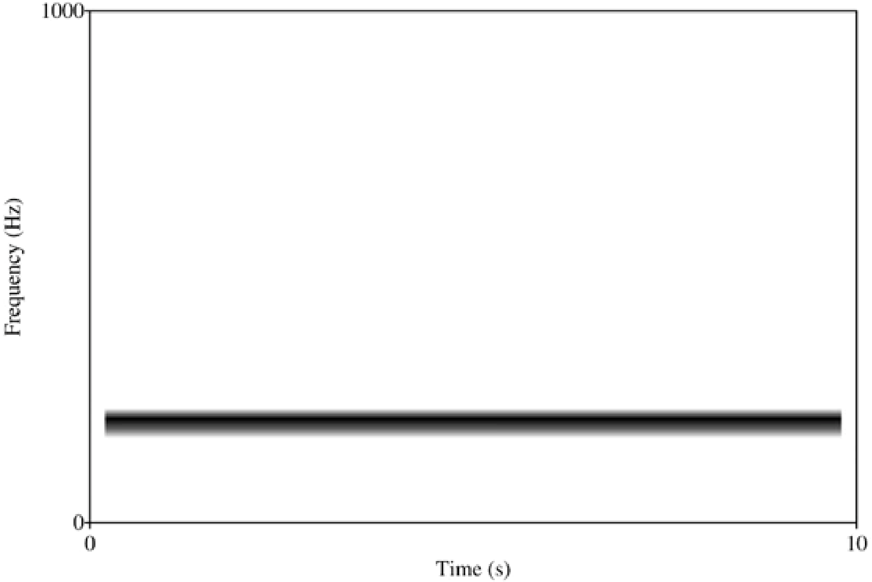


Spectrogram of a piecewise monochromatic signal.

Lighter color → greater DFT magnitude

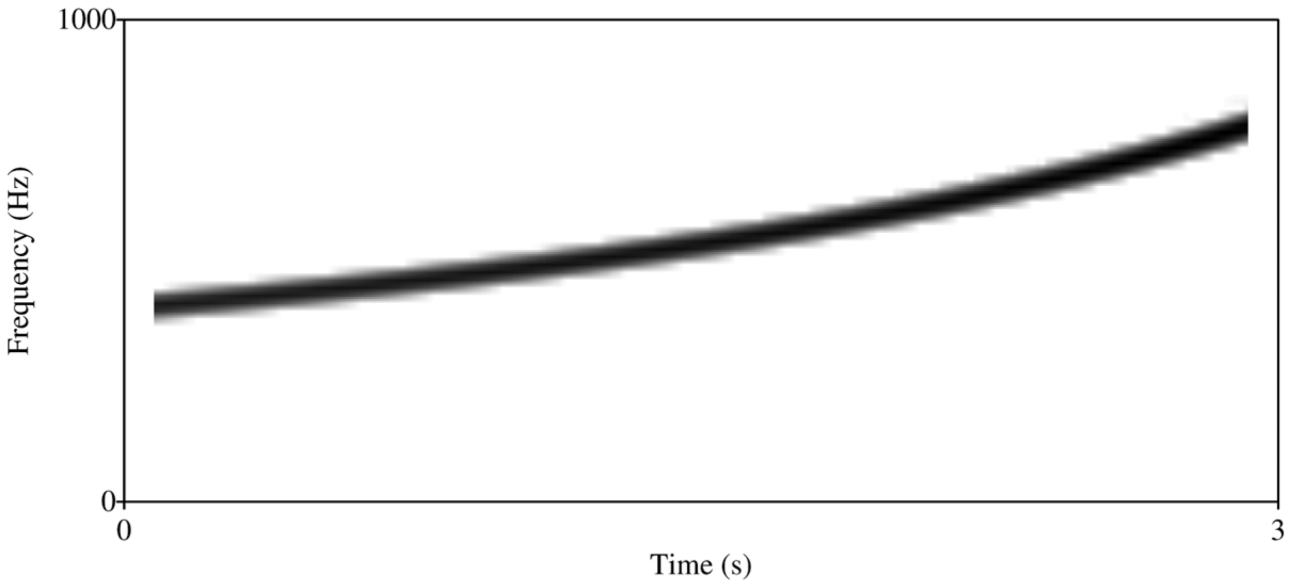
Some audio examples





Courtesy: floridalinguistics.com

Some audio examples



Courtesy: floridalinguistics.com

Some audio examples