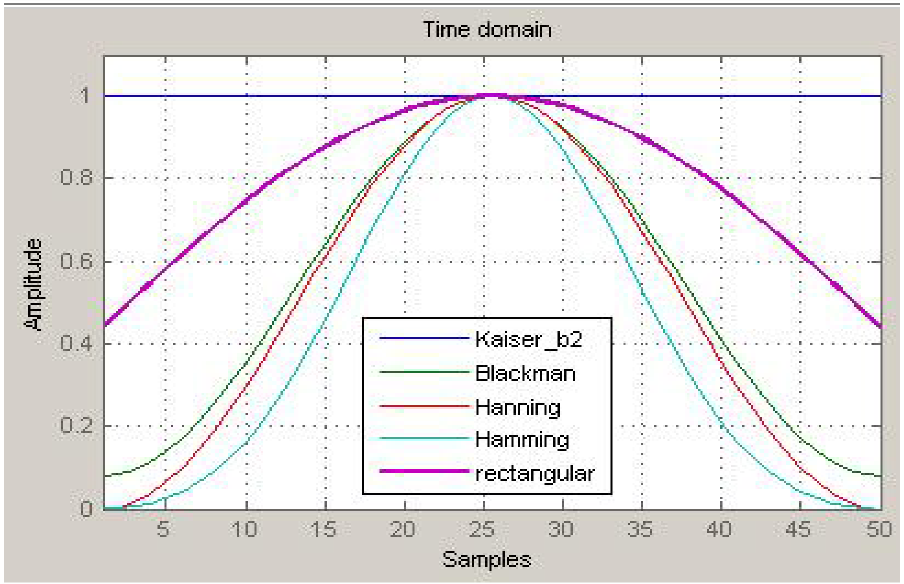
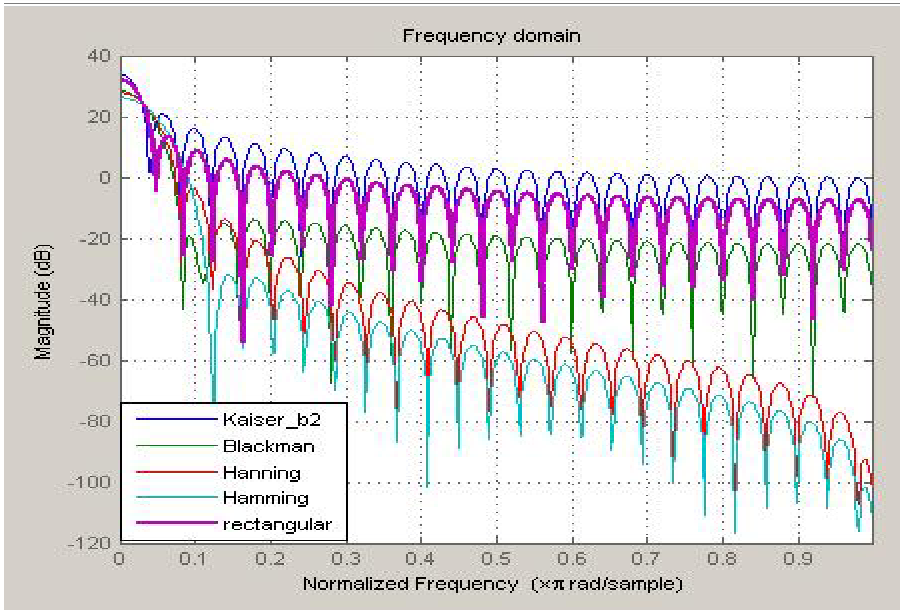
II. Spectrograms

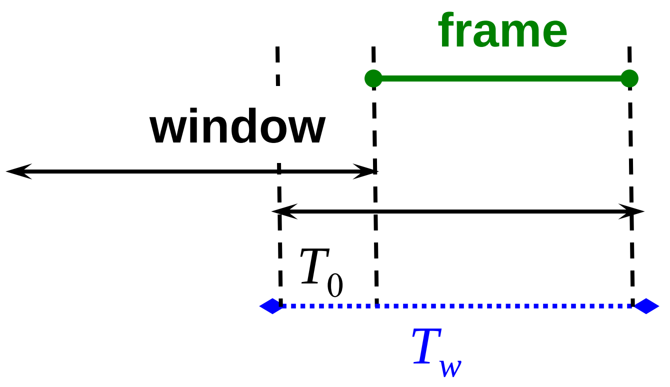
❖ Windows/Frames Effects





- **Purpose of windows**
- **Stationarity issues**
  Time-frequency resolution trade-off

Definition of window/frame

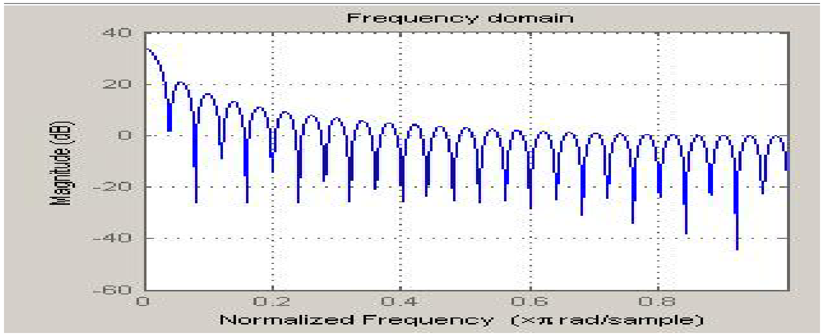


❖ Example: for speech applications, frame duration usually between 10-20 ms, and frame duration less than 8 ms usually not used.

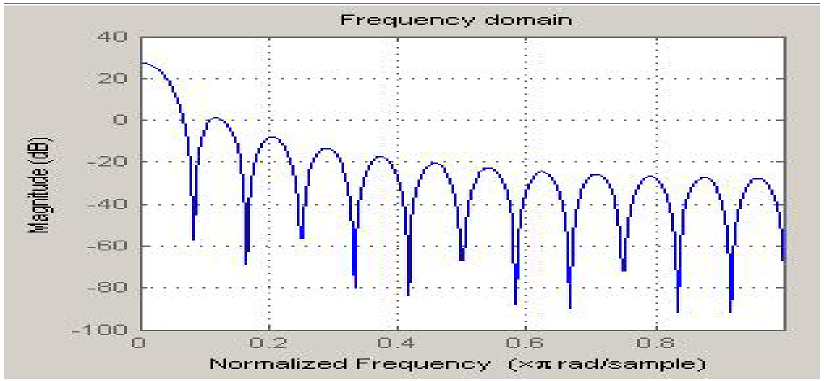
❖ Percentage overlap defined by

$$\boxed{\%\ \text{overlap} = \frac{T_0}{T_w} \times 100\%}$$

❖ Windows / Properties

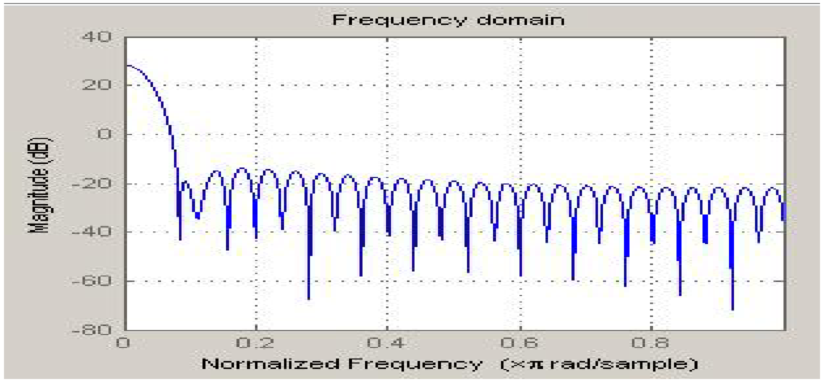


Rectangular window



Triangular window (Bartlett)

$$w(n) = \begin{cases} 2n/M & 0 \le n \le M/2 \\ 2 - 2n/M & M/2 \le n \le M \end{cases}$$

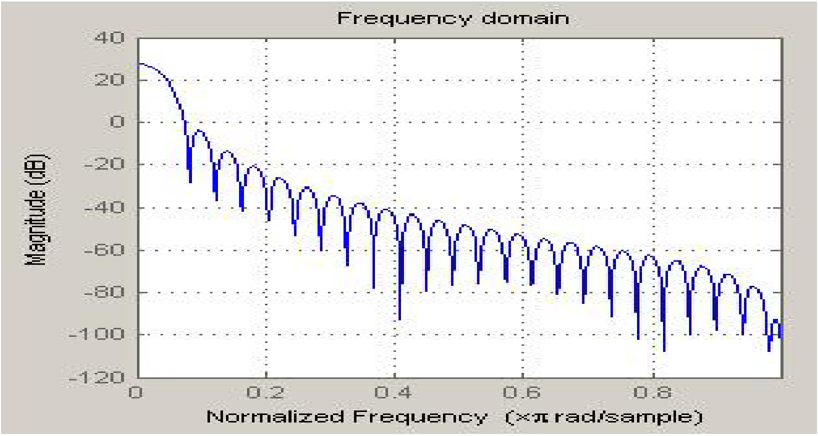


Hamming window

$$w(n) = 0.54 - 0.46\cos(2\pi n/M);\ 0 \le n \le M$$

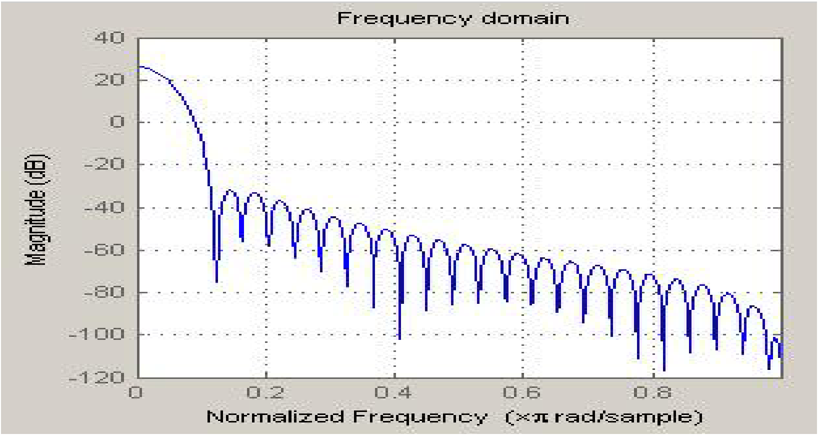
Window length *M=50*

❖ Windows / Properties, cont'



Hanning window

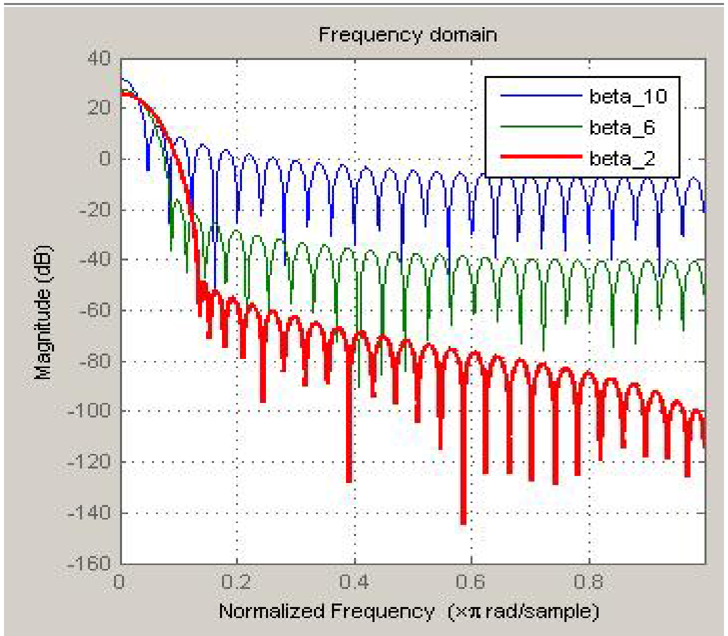
$$w(n)=\begin{cases}0.5-0.5\cos\left(2\pi n/M\right) & 0\le n\le M\\ 0 & ow\end{cases}$$



Blackman window

$$w(n)=0.42-0.5\cos\left(\frac{2\pi n}{M}\right)$$
$$+0.08\cos\left(\frac{4\pi n}{M}\right),0\le n\le M$$

Window length M=50



Kaiser Window (actually, a family of windows)

Designed to be maximally concentrated around $\omega = 0$

$$w(n) = \begin{cases} \dfrac{I_0\left[\beta\left(1-\left[(n-\alpha)/\alpha\right]^2\right)^{1/2}\right]}{I_0(\beta)}, & 0 \le n \le M \\ 0, & ow \end{cases}$$

with $\alpha = \dfrac{M}{2}$

<u>Kaiser window:</u>

- Has 2 adjustable parameters:

$\rightarrow$ M and β

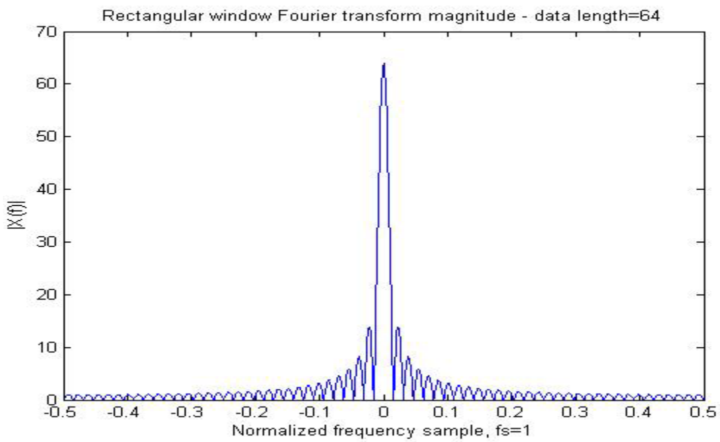
- Used to trade sidelobes for mainlobe width.

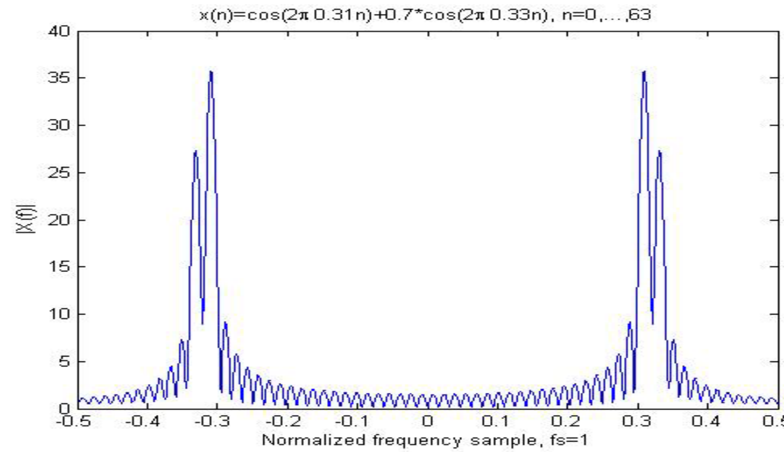
❖ Comparison of commonly used windows, from [1]

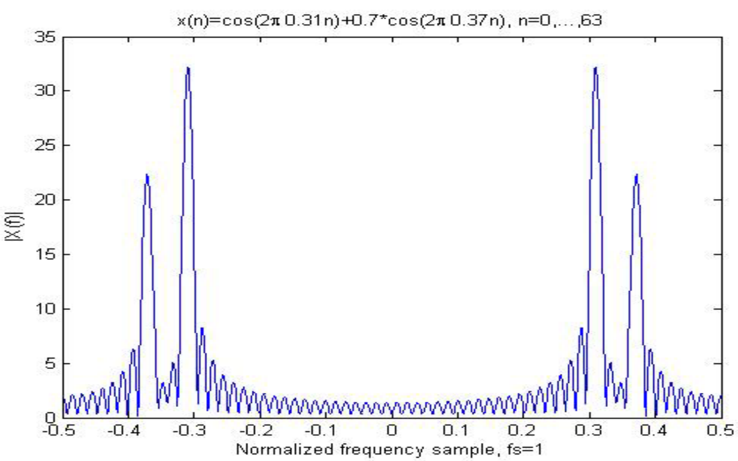
Window Type	Relative Peak Sidelobe Amplitude	Approximate Mainlobe Width	Equivalent Kaiser Window
Rectangular	-13	$4\pi/(M+1)$	0
Bartlett	-25	$8\pi/M$	1.33
Hamming	-41	$8\pi/M$	4.86
Hanning	-31	$8\pi/M$	3.86
Blackman	-57	$12\pi/M$	7.04

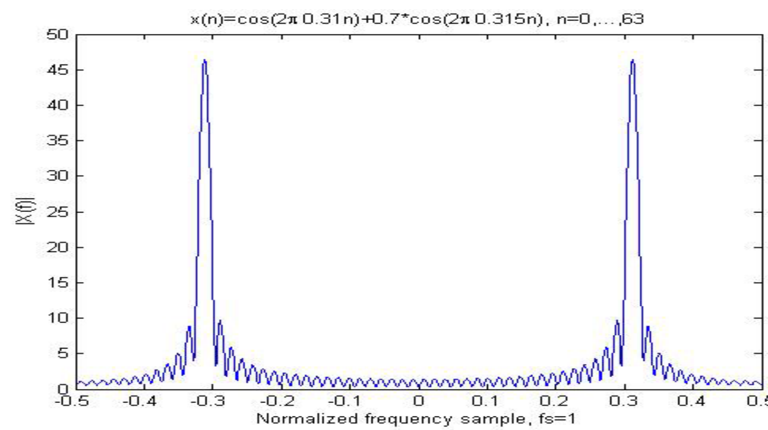
❖ Resolution issues

Controlled by length and type of window <u>NOT</u> by the FFT length.

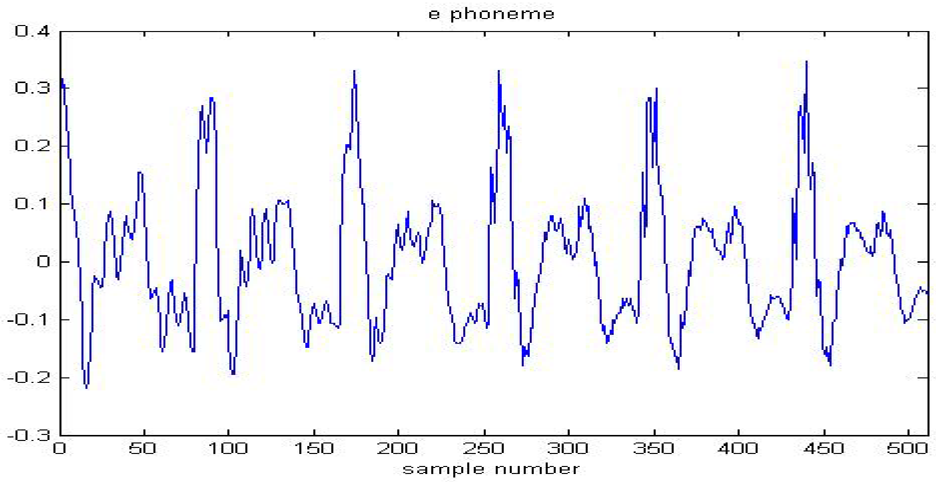




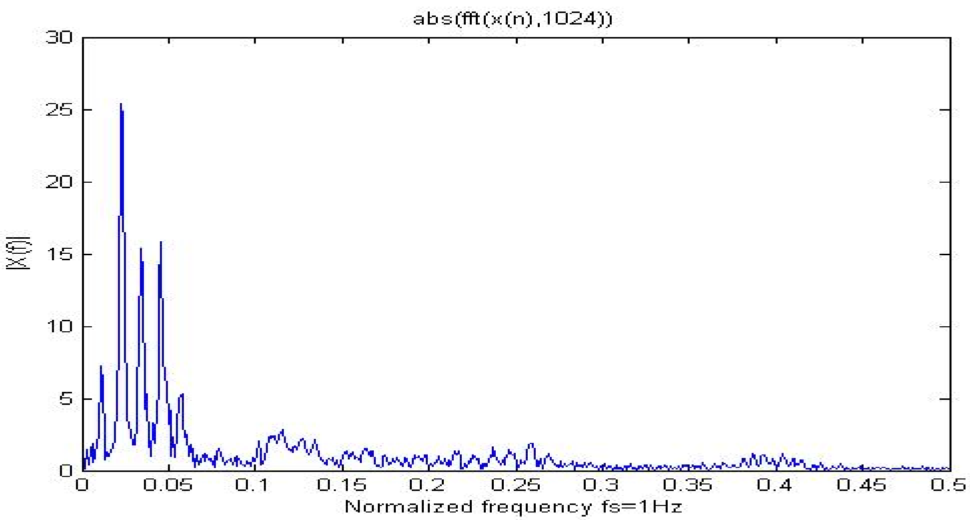




❖ Practical Applications of the Fourier Transform (length effects)

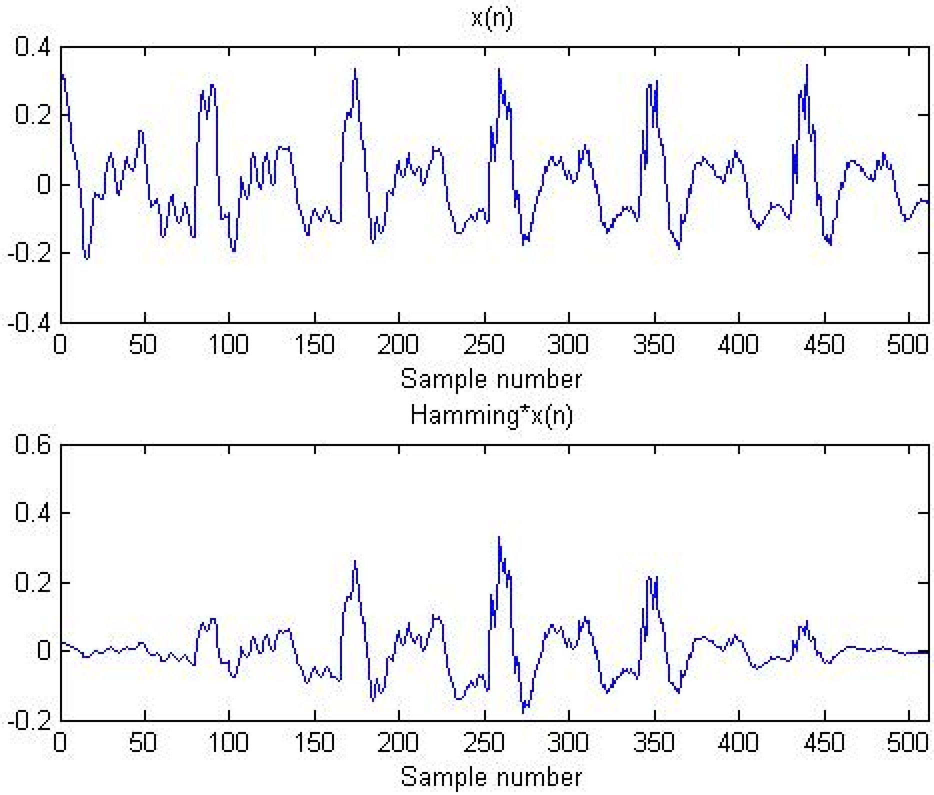


A segment of a vowel extracted with a rectangular window.



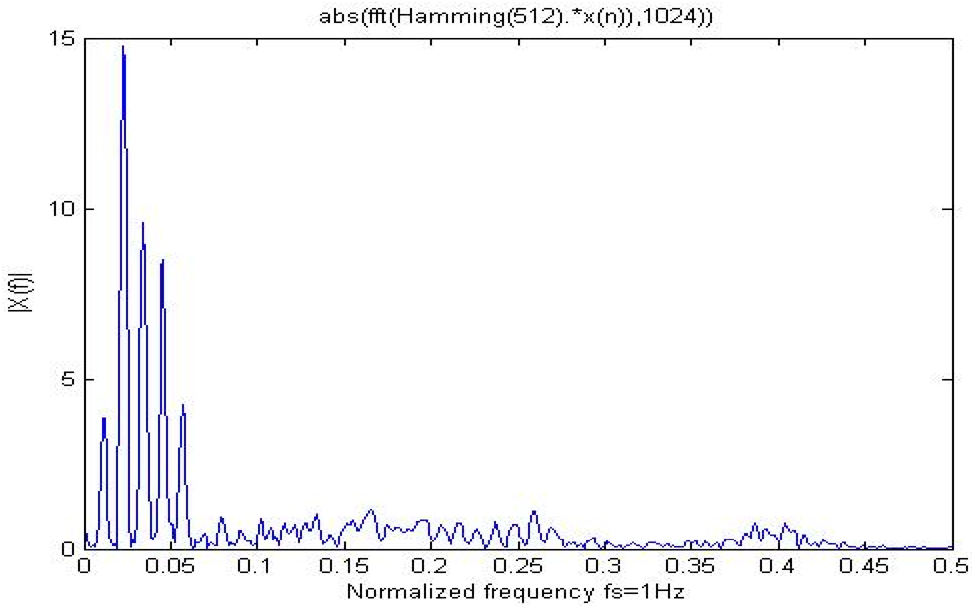
The amplitude spectrum using a rectangular window. Calculated using MATLAB: abs(fft(sig)).

- **Practical Applications of the Fourier Transform (length effects), cont’**

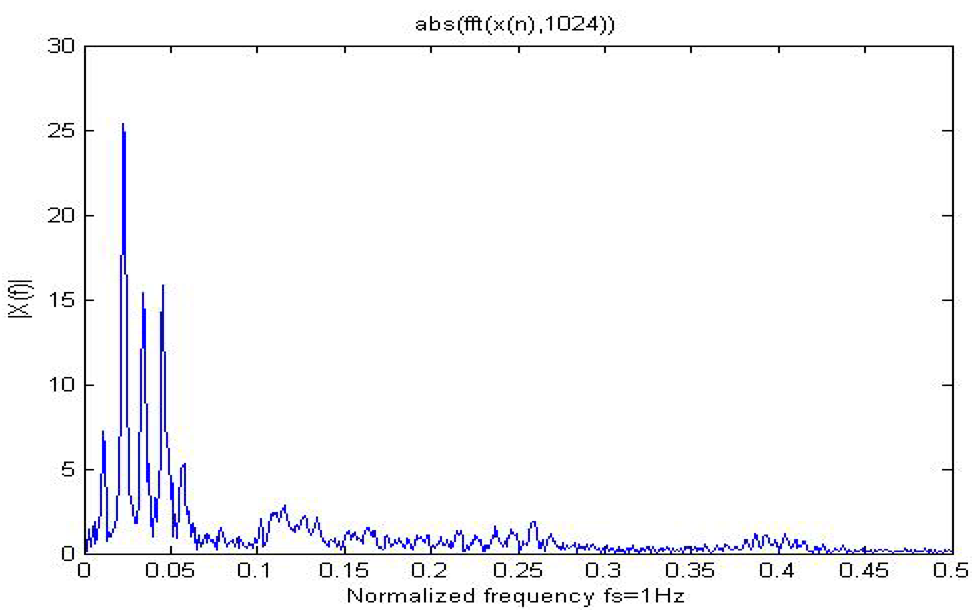


A segment of a vowel extracted with a Hamming window.

Calculated using MATLAB:
hamming(512) .* sig;

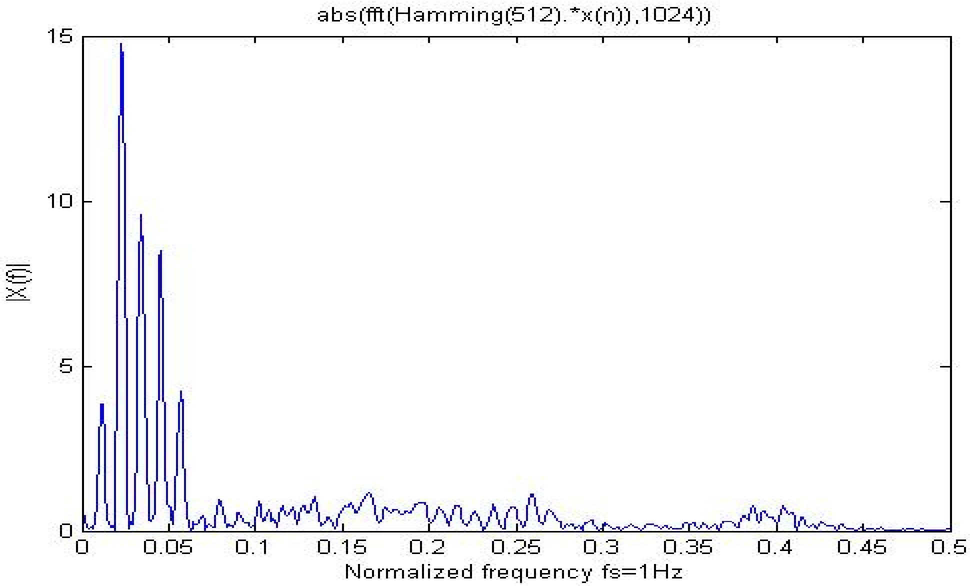


The amplitude spectrum using a Hamming window. Calculated using MATLAB

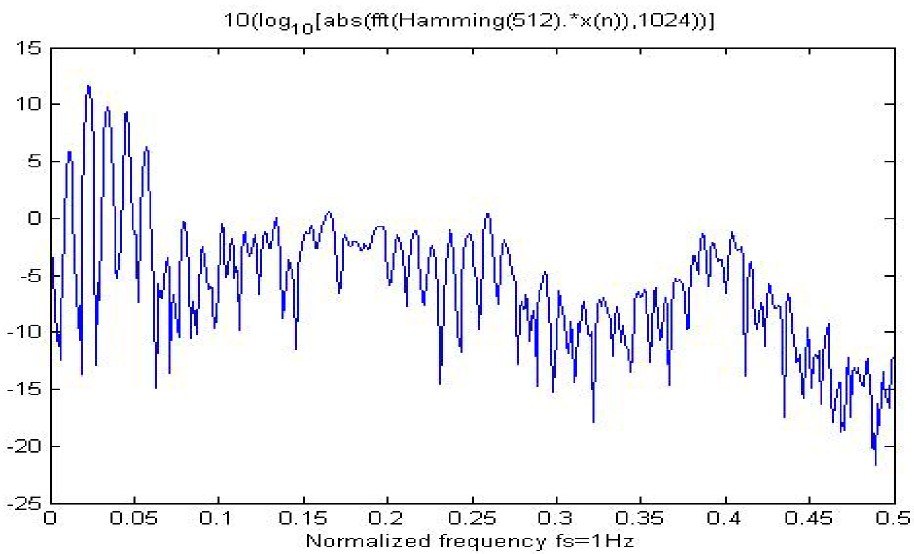


The amplitude spectrum using a rectangular window. Calculated using MATLAB

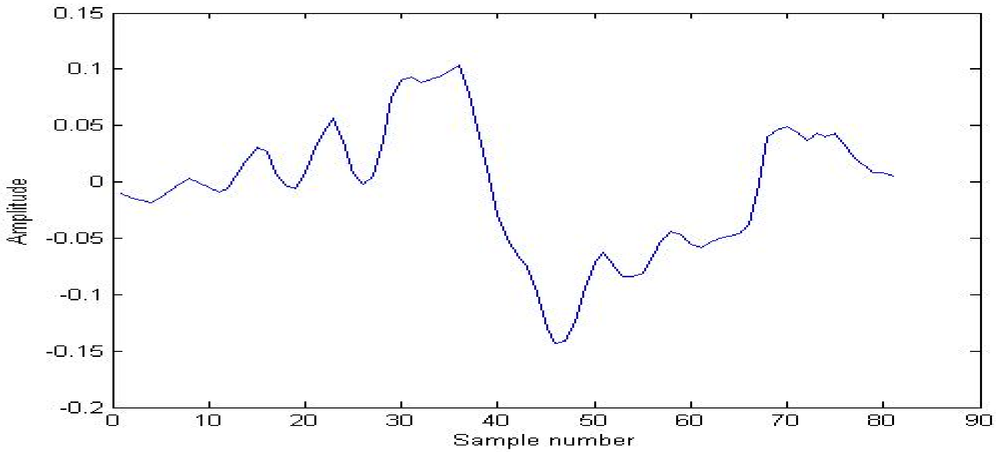
Impact of the DB scale: The power spectrum displayed on a dB scale. Calculated using MATLAB: 10 log10(abs(fft(hamming(512) .* sig,1024)))



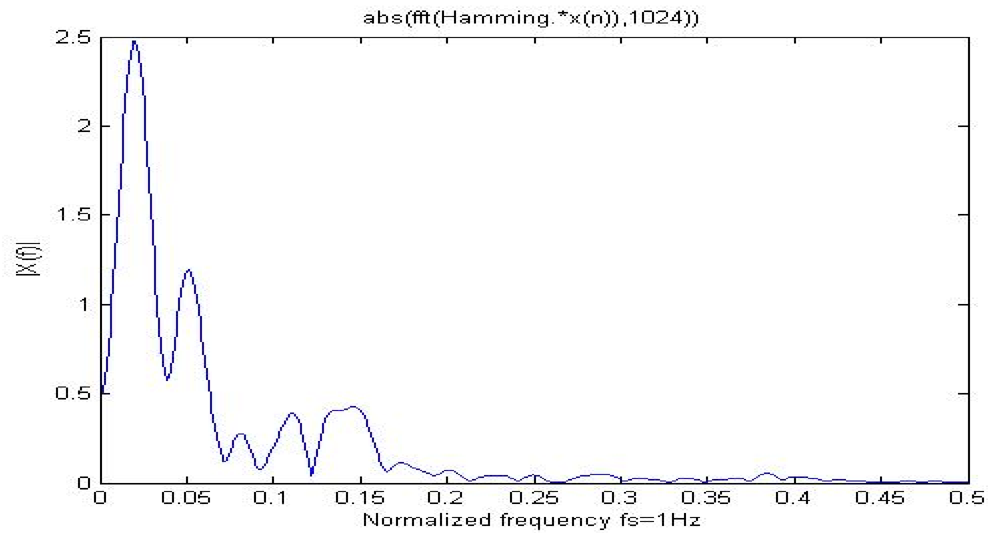
The amplitude spectrum using a Hamming window. Calculated using MATLAB: abs(fft(hamming(512) .* sig,1024)).



Power spectrum displayed on a dB scale. Calculated using MATLAB: 10log10(abs(fft(hamming(512) .* sig)),1024)

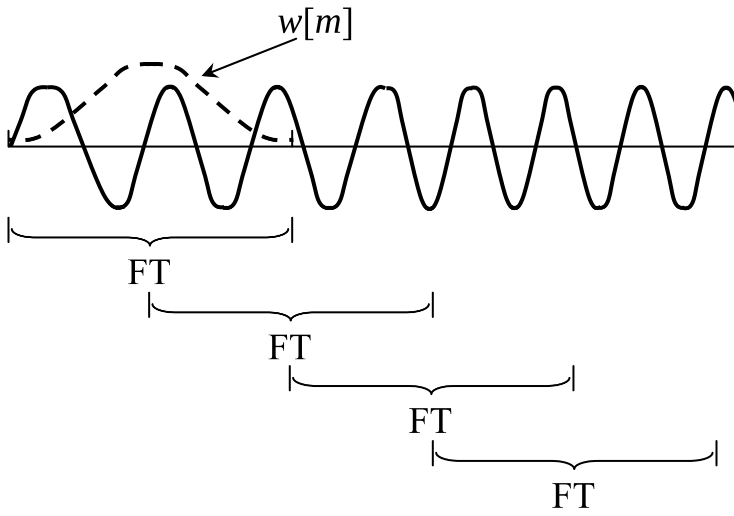


A segment of a vowel extracted with a Hamming window of length 81.

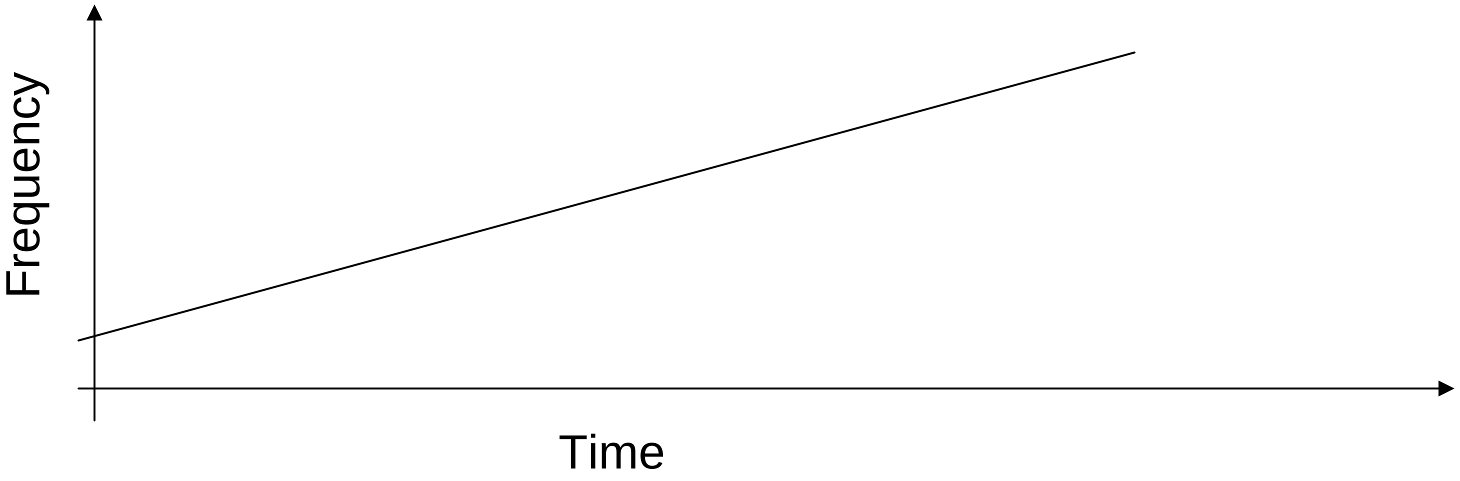


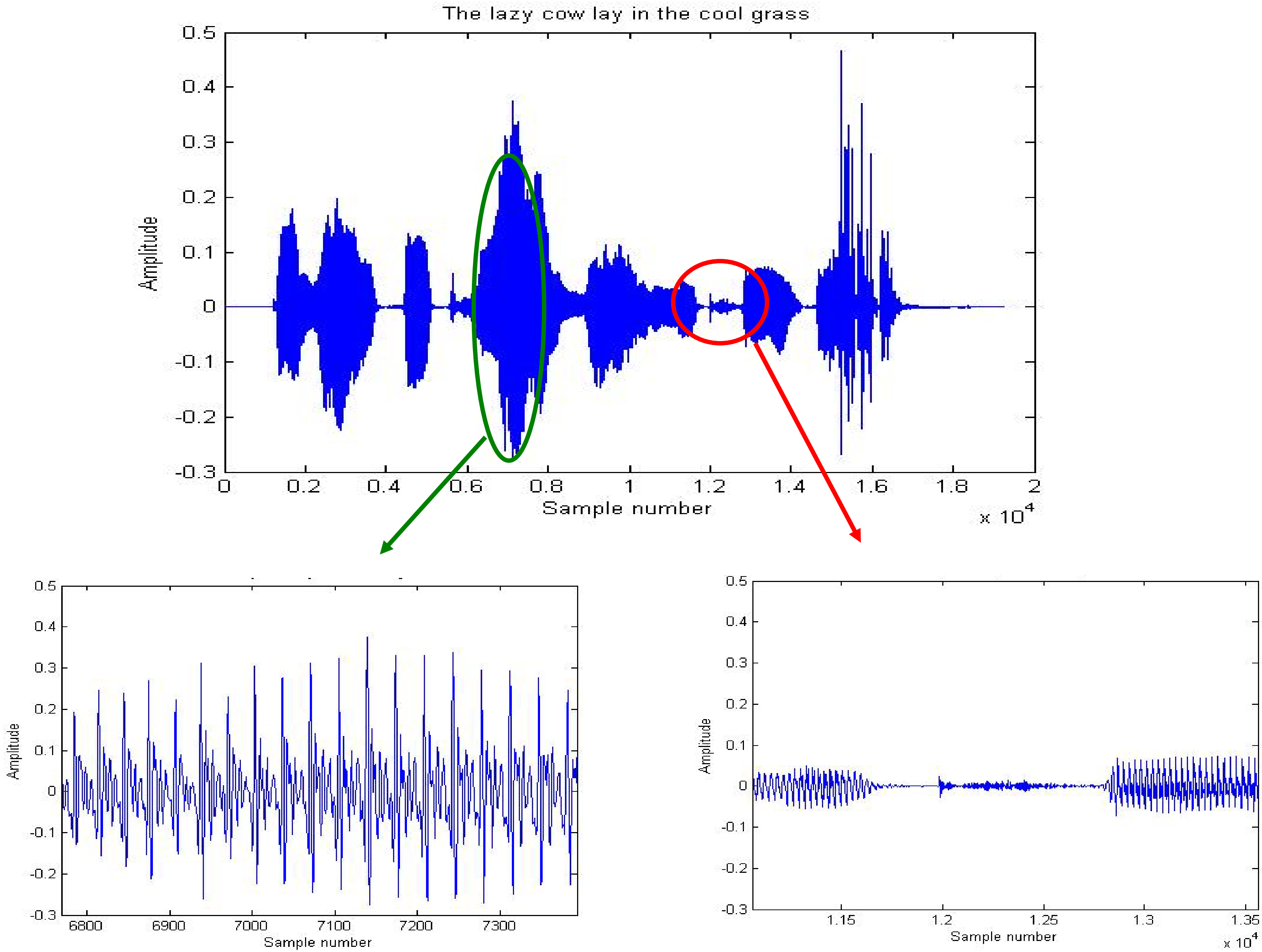
The amplitude spectrum using a Hamming window of length 81, fft zero-padded to 1024.

❖ Sliding Window FT (spectrogram)



Usually window incremented by a fraction (25 to 75%).




The lazy cow lay in the cool grass
Amplitude
Sample number
x 10^4
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
6800
6900
7000
7100
7200
7300
1.15
1.2
1.25
1.3
1.35

❖ Spectrogram – Windowing Effects

– Window type affects the spectrum shape on each window.

– Window size affects the spectrum smoothness.

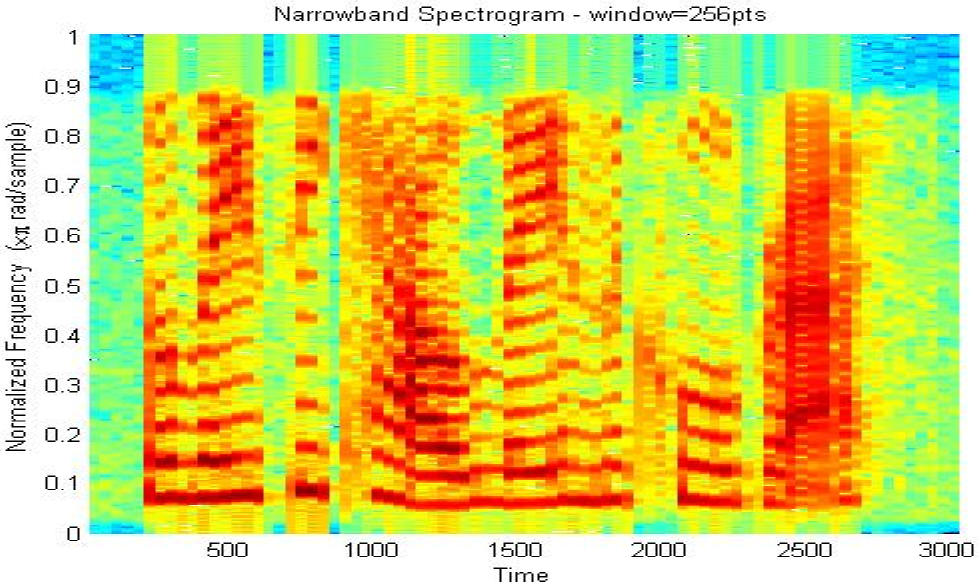
$\rightarrow$ see next pages

- $x(t) \rightarrow X(t, f)$

In speech processing applications:

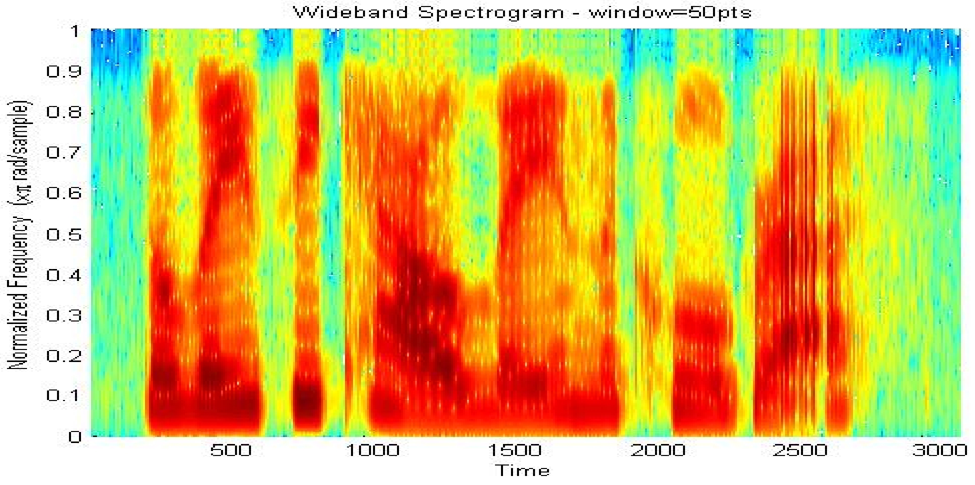
– Short window length (~ 15 ms) $\rightarrow$ <u>wideband spectrogram</u>

– Long window length (~ 50 ms) $\rightarrow$ <u>narrowband spectrogram</u>



Narrowband spectrogram

The lazy cow lay in the cool grass



Wideband spectrogram

References

[1] *Discrete-time Signal Processing*, A. Oppenheim & R. Schafer, Prentice Hall, 1989.

[2] *Discrete Time Processing of Speech Signals*, J. Deller et al, Macmillan, 1993.